Biased Learning under Ambiguous Information*

Jaden Yang Chen†

July 27, 2021

Abstract

This paper proposes a model of how biased individuals update beliefs in the presence of informational ambiguity. Individuals are ambiguous about the actual signal-generating process and interpret signals according to the model that can best support their biases. This paper provides a complete characterization of the limit beliefs under this rule. The presence of model ambiguity has the following effects. First, it destroys correct learning even if infinitely many informative signals can be observed. When the ambiguity is sufficiently high, individuals can justify their biases, leading to belief extremism and polarization. Second, an ambiguous individual can exhibit greater confidence than a Bayesian individual with any feasible model perception. This phenomenon comes from a novel complementary effect of different models in the belief set.

JEL Classification: D81, D83, C72
Keywords: Biased learning, model uncertainty, ambiguity, self-serving bias

*I am indebted to David Easley for his guidance throughout the project. I want to thank Larry Blume, Tommaso Denti, Ted O'Donoghue and Siguang Li for advice and discussions. I thank the editor and two anonymous referees for excellent suggestions. All errors are mine.

†Department of Economics, Cornell University. E-mail: yc2325@cornell.edu

1 Introduction

It is standard in the learning literature to assume that individuals are rational and update beliefs using Bayes rule. However, a growing body of evidence suggests that individuals often hold *biased* beliefs and process information in a self-serving manner. For example, individuals tend to overinterpret others' compliments in order to maintain a good self-image, leading to the famous "better-than-average" effect (e.g., Taylor and Brown 1988, Alicke et al. 2005). Citizens holding different political views and receiving identical information tend to process it differently to support their positions, which can lead to polarization in beliefs (e.g., Taber and Lodge 2006). In financial markets, many investors interpret evidence to support their current investment philosophy (e.g., Graham et al. 2009, Grinblatt and Keloharju 2009, Chang and Cheng 2015). Such biases also exist for analysts, and it is well-documented that analysts tend to overreact or underreact to new information and exhibit biases when issuing reports (e.g., Butler and Lang 1991, Brous and Kini 1993, Easterwood and Nutt 1999).

Beliefs of biased individuals typically exhibit non-Bayesian dynamics, so standard techniques in Bayesian learning cannot be directly applied to this situation.[1] It remains unanswered how beliefs evolve and how learning outcomes differ when we account for individuals' biases. This paper provides a framework that enables analysis of biased learning. In the paper, I investigate a standard learning problem where individuals observe a sequence of signals and update beliefs about the true state of the world. Different from the standard framework, (i) individuals are ***ambiguous*** about the true data-generating process and consider a set of models as possible, and (ii) individuals are ***biased*** toward some states of the world and interpret signals according to the model that best supports their biases. The ***biased updating rule*** captures the idea that individuals process information in a self-serving manner. The major deviation from Bayes rule is that under the biased rule, individuals can use different models for different signals. This feature allows the biased rule to exhibit interesting non-Bayesian dynamics. Below is one example.

Example 1. [Good News or Bad News] The state of the world is summarized by $\Theta = \{H, L\}$. Individuals receive signals from the signal space $S = \{S_H, S_L\}$. Signals can either be informative or uninformative in the sense that

$$\frac{P(S_H|H)}{P(S_L|H)} = \frac{P(S_L|L)}{P(S_H|L)} \in \{1, \alpha\} = \mathcal{A} \quad \text{where } \alpha > 1.$$

If the likelihood ratio is α, signals are informative; if the likelihood ratio is 1, signals are uninformative. Individuals are biased toward one of the two states. After receiving signals, they update beliefs by Bayes rule according to the signal interpretation that maximizes the probability of their biased states. It can be verified that individuals will process good news and bad news in an *asymmetric* way. Consider an individual with bias H; if he received a signal S_H (i.e., the good news), he would view this signal as informative hence adopt model α. On the other hand, if he

[1] see Blume and Easley (1998) for a survey on Bayesian learning in economics.

received a signal S_L (i.e., the bad news), he would treat it as uninformative by adopting model 1. In summary, individuals tend to value good news but overlook bad news. This "good-news effect" has been documented by many experimental findings (e.g., Eil and Rao 2011, Ertac 2011, Möbius et al. 2011, Coutts 2019, Barron 2020) but is not well reconciled by a standard Bayesian framework that requires *consistent* interpretations for all signal realizations.

In this paper, I examine a more general learning environment, where individuals face a general state space and signal structures and can be biased toward a weighted set of states. Individuals observe a sequence of i.i.d. signals, but they are ambiguous about the true data-generating process. In other words, individuals perceive a set of feasible models, $\mathcal{A}$, and do not assign probabilities to models. After observing the signals, individuals update beliefs via Bayes rule according to the feasible model that can best justify their endowed biases. This learning framework satisfies Bayes rule model-wise, but it can exhibit non-Bayesian features at the aggregate level. As a result, the model nests Bayes rule as a special case, and it can also explain some behavioral patterns inconsistent with Bayesian framework.

In addition to its new implications, this rule has the advantage of being tractable. As information accumulates, posteriors approach a limit that can be characterized in a simple form. This paper provides characterizations of limit beliefs under the biased rule (in Theorem 1 and Theorem 2). The characterizations are based on the seminal work of Berk (1966), and they can be summarized as follows. As the number of signals approaches infinity, (i) individuals will eventually adopt models that minimize a normalized version of relative entropy of their biased states, and (ii) limit beliefs will concentrate on the states that minimize the relative entropy under these models.

Based on these characterizations, I then discuss some effects of ambiguity on learning. First, ambiguity can lead to incomplete learning and polarization. Under sufficient ambiguity, individuals have enough flexibility in interpreting information to justify their biases or settle on states far from the true state in the limit. On the contrary, as ambiguity diminishes, correct learning can be restored in a weak sense; when the state space is finite, correct learning can be achieved under sufficiently small ambiguity. Second, ambiguity can lead to overconfidence. Under ambiguity, perceived models can complement each other during the learning, so a biased individual can become strictly more confident toward their biases than a Bayesian individual with any feasible model perception. The intuition is similar to Example 1. Model ambiguity enables individuals to exploit good news and hedge against the bad news by choosing to interpret signals asymmetrically, so individuals could justify their biases to a larger extent than if they only perceived a specific model. Moreover, this paper shows that the overconfidence can persist even with arbitrarily many signals, which highlights that the biased rule can be different from Bayesian rule even in the limit.

Related Literature and Contributions. The key contribution of the paper is to propose an updating rule under ambiguity that accounts for people's biases. First, in terms of topics, this paper belongs to the literature on learning under behavioral biases. The most relevant bias is self-serving bias, which means that individuals interpret information in a self-serving manner. This type of bias has been investigated in many experimental studies (e.g., Babcock et al. 1995, Babcock et al.

1996, Haisley and Weber 2010, and Deffains et al. 2016), but its effects on learning have not been systematically studied in theoretical works. This paper contributes to the literature by proposing a framework to study the biased learning. Methodologically, this paper employs techniques from misspecified learning literature, along with the idea of model uncertainty, to model biased learning. This approach allows us to derive complete characterizations of limit beliefs for a variety of biases, which are both applicable and new to the literature. In addition to self-serving bias, another relevant bias is confirmatory bias, which is more widely studied along with learning (e.g., Rabin and Schrag 1999, Fryer et al. 2017). This thread of literature is similar to this paper in that information is processed in a biased manner, but it is different in how the bias affects learning. Under confirmatory bias, individuals interpret information toward the most likely state according to their *current* beliefs, and they are not intrinsically attached to any state. Under the bias in this paper, individuals attach intrinsic value to some state, and they update beliefs to justify it regardless of their current beliefs. The former bias is more motivated by the fact that people exhibit some inertia to their prior judgments, whereas the latter is more motivated by the fact that people seek to maintain self-esteem. Due to these differences, they often lead to different belief dynamics and have different asymptotic properties.

Second, this paper also adds to the literature on learning under ambiguity, especially learning under model uncertainty, by suggesting a new approach to belief updating.[2] The biased updating rule is different from two well-studied updating rules under ambiguity, the full Bayesian rule and the maximum likelihood rule, in the following aspects. The full Bayesian rule updates all models indiscriminately and leads to a set of posteriors, whereas both the maximum likelihood rule and the biased updating rule only update models that satisfy some criterion, and they often lead to a unique posterior. Their main difference is that the maximum likelihood rule follows an objective criterion and selects models according to the probability of generating observed information, whereas the biased rule follows a subjective criterion and selects models according to whether they can maximally support the endowed bias. Under appropriate decision rules, these differences can also lead to different actions as shown in Section 7.

Finally, under model uncertainty, individuals will inevitably perceive some incorrect models, so this paper also contributes to the growing literature on learning with misspecified models. [3] This thread of literature mostly adopts Bayes rule, so this paper differs from that literature in the same way it differs from Bayes rule. The most significant difference is that under the biased updating rule, individuals can switch models to tailor their biases, whereas in the misspecification literature, individuals mostly adhere to a fixed model perception. Due to this difference, this paper's framework can produce phenomena that are inconsistent with misspecified Bayesian learning.

[2] Gilboa and Marinacci (2016), Machina and Siniscalchi (2014) provided exhaustive surveys on ambiguity-related topics. Below is an incomplete list of research on model uncertainty. Marinacci (2002,2015), Marrinaci and Massari (2019), Battigalli et al. (2015), Battigalli, Catonini, et al. (2019), Battigalli, Francetich, et al. (2019), Chen (2019).

[3] An incomplete list of learning under incorrect model include Blume and Easley (1982), Nyarko (1991), Bohren (2016), Fudenberg, Romanyuk and Strack (2017), Heidhues, Kőszegi and Strack (2018), Bohren and Hauser (2019), Frick, Iijima and Ishii (2019,2020), Fudenberg, Lanzani and Strack (2020).

This paper is organized as follows. Section 2 presents the benchmark model. Section 3 presents some examples that illustrate applications of the biased rule. Section 4 discusses major modeling assumptions. Section 5 and 6 characterize limit beliefs under the biased rule. Section 7 discusses some examples where actions are involved. Section 8 discusses some extensions. The Appendix includes the omitted proofs and supplementary materials.

2 Benchmark Model

The state space Θ is a compact subset of a Polish space which is endowed with Borel σ-algebra $\mathcal{B}_\Theta$ and a finite measure m. The true state is denoted by $\theta^* \in \Theta$. Individuals do not know the true state and hold a prior that is specified by some density function $\mu_0 : \Theta \to \mathbb{R}_+$ with respect to m, where μ_0 is continuous and has full support. Individuals receive a sequence of i.i.d. signals $\{s_t\}$, where s_t is a random variable taking values in the signal space S. The signal space S is a Polish space associated with Borel σ-algebra $\mathcal{B}_S$ and a σ-finite measure v. Signals are generated by a ***model*** that belongs to the model space $\mathbb{A}$, where $\mathbb{A}$ is also a Polish space. Conditional on state θ, each model $\alpha \in \mathbb{A}$ induces a signal distribution $f(s|\theta, \alpha)$, which is a density function on S with respect to v. Denote by α^* the true model, so $f(s|\theta^*, \alpha^*)$ represents the true distribution of s_t. Denote by $\mathbb{P}^*$ and $\mathbb{E}^*$as the true probability measure and expectation operator induced by the true signal distribution. Below are some technical assumptions on f:

Assumption 1. *The mapping $f : S \times \Theta \times \mathbb{A} \to \mathbb{R}_+$ is jointly continuous, and for all $(\theta, \alpha) \in \Theta \times \mathbb{A}$, the support of $f(s|\theta, \alpha)$ is S.*

Assumption 2. *For all $\theta \in \Theta$ and $\alpha \in \mathbb{A}$, there exists an open set U containing (θ, α), such that $\mathbb{E}^* \sup_{(\theta,\alpha)\in U} \log^2 f(s|\theta, \alpha) < \infty$.*

Assumption 3. *For all $\theta, \theta' \in \Theta$ and $\alpha \in \mathbb{A}$, $\mathbb{P}^*(\{s \in S : f(s|\theta, \alpha) \neq f(s|\theta', \alpha)\}) > 0$.*

These assumptions are standard and can accommodate a wide class of signal distributions. The purpose of Assumption 2 is to impose some boundedness conditions, which allow us to apply the dominated convergence theorem to establish continuity. Assumption 3 requires that every model induces different signal distributions under different states, so signals are informative under every perceived model.[4]

Under the standard learning framework, individuals are certain about the true model, and they update beliefs according to their perceived true models, where the perceived models can be either correctly or incorrectly specified. Different from the standard model, this paper assumes that individuals are ***ambiguous*** about the true model. More precisely, individuals consider a compact set of models $\mathcal{A} \subset \mathbb{A}$ as possible and are unable to assign probabilities on $\mathcal{A}$. Intuitively, this means

[4]Without this assumption, we may encounter a trivial situation where learning cannot occur on a subset of states, in which case how the prior μ_0 is formed on this set becomes vital. It needs to be noted that Assumption 3 is not required for the characterizations of limit beliefs in Theorem 1 and 2 but is needed to establish the overconfidence effect in Theorem 3 in the Appendix.

that individuals are uncertain about how to interpret signals, and they only know that the correct interpretation belongs to $\mathcal{A}$. The model set is *correctly specified* if $\alpha^* \in \mathcal{A}$, and is *misspecified* if $\alpha^* \notin \mathcal{A}$.

Every individual is endowed with some bias over Θ, where the ***bias*** is represented by a payoff function $\tau : \Theta \to \mathbb{R}_+$. The support of τ, denoted by $\operatorname{supp}(\tau) = \overline{\{\theta : \tau(\theta) > 0\}}$, is referred to as the set of ***biased states***, which are states with strictly positive payoff.[5] Denote by $\mathcal{T}$ the set of all possible biases. I assume that every $\tau \in \mathcal{T}$ is a continuous function on its domain. Every individual derives utility from his belief, called ***belief utility***, which is equal to the expected payoff from τ based on his belief. If biases are confirmed by beliefs to a larger extent, individuals can derive higher utility. More precisely, for all $\alpha \in \mathcal{A}$, denote by $U_t(\tau|\alpha)$ the time-t belief utility if individuals update according to model α by Bayes rule, that is,

$$U_t(\tau|\alpha) \equiv \int_\Theta \tau(\theta)\, \mu_t(\theta|\alpha)\, dm(\theta),$$

where $\mu_t(\theta|\alpha)$ denotes the Bayes update of $\mu_0(\theta)$ according to model α at time t. When updating beliefs, individuals adopt the ***biased updating rule*** in the sense that they will update according to the model which generates the highest belief utility, denoted by $U_t(\tau)$. More formally, denoting by μ_t^τ the time t belief held by an individual with bias τ, we have:

$$\forall \theta \in \Theta : \quad \mu_t^\tau(\theta) = \mu_t(\theta|\alpha_t^\tau) \quad \text{where } \alpha_t^\tau \in \arg\max_{\alpha \in \mathcal{A}} U_t(\tau|\alpha) \tag{1}$$

Here, α_t^τ is referred to as the ***model perception*** of an individual with bias τ at time t. Notice that it is possible that multiple models can maximize the belief utility. In this case, I assume that individuals adopt some rule to break the tie, so our biased updating rule always generates a unique belief. Consider again the case in Example 1, the bias-H individual's bias can be described by $\tau(\theta) = \delta_G(\theta)$, so he will update his beliefs to maximize the probability of state H. By appropriate definition of τ, this model can accommodate a variety of biases. Section 3 discusses some examples.

3 Illustrative Examples

In this section, I discuss some examples which illustrate various applications of the biased rule in economics. Some basic properties of the biased rule are also discussed in these examples.

Example 2. [Overconfidence and Excess Entry] A group of candidates $N = \{1, ..., n\}$ decide whether to apply for a prize that will be awarded to the top candidate. Candidates do not know their rankings and they start with a common prior. An ambiguous feedback is then generated by a model in $\mathcal{A}$ and is observed by all candidates. Every candidate is biased and trying to justify himself as the best. Denote by $\mu^i(i)$ as the probability that candidate i thinks that he is the best

[5] For simplicity, this paper focuses on positive-valued τs , but the discussion can be extended to negative-valued τs, where individuals can also deflate the probability of some states.

candidate. In non-trivial cases, we have:

$$\sum_{i\in N}\mu^{i}\left(i\right)=\sum_{i\in N}\max_{\alpha\in\mathcal{A}}\mu\left(i|\alpha\right)>\max_{\alpha\in\mathcal{A}}\left[\sum_{i\in N}\mu\left(i|\alpha\right)\right]=1.$$

The sum of the self-perceived probability of being the top candidate is *strictly* greater than 1, which features overconfidence or the "better-than-average" effect. In the case of Bayesian updating with a common model perception, this effect cannot be naturally reconciled due to the fact that the beliefs always sum up to 1. [6]As a result, ambiguous feedback may contribute to a larger volume of applications than non-ambiguous case. This effect can be used to explain excess entry into competitive markets triggered by overconfidence, especially when the performance feedback is ambiguous. For related experimental or empirical evidence, see Heath and Tversky (1991), Camerer and Lovallo (1999) and Karelaia and Hogarth (2010).

Example 3. [Fairness and Preferences for Redistribution-I] A worker has an ability level $\theta\in\{H,L\}$ and is working in some profession. The fairness of the profession is represented by $\alpha\in\{fair,unfair\}$. The worker receives feedback $x_t\in\{G,B\}$ generated by $f\left(x_t|\theta,\alpha\right)$, where

$$\begin{cases}f\left(G|H,fair\right)=p & f\left(G|L,fair\right)=1-p\\ f\left(G|H,unfair\right)=1-p & f\left(G|L,unfair\right)=p\end{cases},\quad\text{where } p>1/2.$$

In other words, a high-ability worker is more (less) likely to receive the good feedback G than a low-ability worker when the profession is fair (unfair). Suppose that the true fairness is $\alpha^*=fair$. Assume that $\tau=\delta_H$, so the worker is trying to defend the conjecture that he has high ability. This worker updates beliefs according to the biased updating rule. The following results are easy to verify:

(i) At each period t, if the worker received more good feedback than bad feedback, he would perceive the profession as fair (i.e., $\alpha_t^\tau=fair$), otherwise, he would perceive the profession as unfair (i.e., $\alpha_t^\tau=unfair$).

(ii) As $t\to\infty$, the worker with low ability will almost surely perceive the profession as unfair; the worker with high ability will almost surely perceive the profession as fair.

In this example, high income groups believe that their wealth is due to their ability and low income groups attribute their lower wealth to unfairness. Both groups are biased and may disagree about redistribution plans due to their biases. Therefore, this simple example provides a possible explanation of the phenomenon that low-income groups tend to support redistribution more than high-income groups. Similar effects are supported by recent experimental evidence (e.g., Deffains, Espinosa and Thöni 2016, Casser and Klein 2019).

[6]Even with heterogeneous model perceptions, the effect is not always true. It is easy to construct an example such that for all possible model-perception combinations $(\alpha_1,...,\alpha_n)$ in $\mathcal{A}$, the sum of self-perceived probability is less than 1 for some signals, that is, the group may exhibit under-confidence in some cases.

Example 4. [Optimistic and Pessimistic Investors-I] The true state of the market can either be good G or bad B. The market consists of two types of investors $\mathcal{T} = \{o, p\}$, where the o-type (i.e., optimists) are biased toward the good state, and the p-type (i.e., pessimists) are biased toward the bad state. Both types of investors hold full-support priors. Investors receive i.i.d. signals overtime. Signals take values in $S = \{g, m, b\}$ with the true data-generating process being

$P(s\|\theta)$	g	m	b
G	$\frac{\lambda}{1+\lambda}(1-\varepsilon)$	ε	$\frac{1}{1+\lambda}(1-\varepsilon)$
B	$\frac{1}{1+\lambda}(1-\varepsilon)$	ε	$\frac{\lambda}{1+\lambda}(1-\varepsilon)$

,

where $\lambda > 1$ and $\varepsilon \in (0,1)$. Investors know how to interpret signals g and b in the sense that they know that

$$\frac{P(g|G)}{P(g|B)} = \frac{P(b|B)}{P(b|G)} = \lambda,$$

but they are ambiguous in the interpretation of signal m in the sense that they perceive a set of likelihood ratios induced by m,

$$\frac{P(m|G)}{P(m|B)} \in \left[\frac{1}{1+\delta}, 1+\delta\right],$$

where $\delta \geq 0$ describes the degree of ambiguity. When $\delta = 0$, investors are certain about how to interpret signal m. As δ grows, the uncertainty in interpretation also expands. When δ is sufficiently large, *polarization* of opinions arises almost surely, and we have $\mu_t^o \to \delta_G$ and $\mu_t^p \to \delta_B$ almost surely. In this case, investors become confident in their biased states in the limit, so presenting them with the same information increases instead of decreasing their disagreements. On the contrary, when δ is sufficiently small, polarization disappears and both types of investors will learn the true state correctly.

The intuition is straightforward. When there is sufficient uncertainty, individuals have adequate flexibility to interpret the signals in their most preferred way. Therefore, they can successfully convince themselves that their biased states are true states, which leads to belief polarization. On the contrary, when the degree of uncertainty is low, the room for signal interpretations is also restricted, which leads to complete learning. This example suggests that informational uncertainty can exacerbate biased beliefs, which is also supported by empirical findings (e.g., Ackert and Athanassakos 1997, Das et al. 1998, Athanassakos and Kalimipalli 2003, Chang and Choi 2017).

Example 5. [Inferring the market fundamental-I] The market price of a commodity is determined by the following demand-and-supply system:

$$\begin{cases} \log Q_t^D = -\alpha^* \log P_t + \varepsilon_t \\ \log Q_t^S = \log P_t + \eta_t \\ Q_t^D = Q_t^S \end{cases} \quad \text{with } \varepsilon_t \sim \mathcal{N}\left(\theta^*, \sigma^2(\theta^*)\right) \text{ and } \eta_t \sim \mathcal{N}(0,1),$$

where $\alpha^* > 0$ represents the *demand elasticity*, and $\theta^* \in \Theta = [\underline{\theta}, \overline{\theta}]$ represents the *market fundamental*— a higher θ^* means that the market demand is stronger. An agent can observe the full price history $\{P_1, ..., P_t\}$ up to some period t, and he knows the demand-and-supply relation. Solving for the equilibrium prices, he knows that:

$$\log P_t \sim \mathcal{N}\left(\frac{\theta^*}{\alpha^* + 1}, \frac{1 + \sigma^2(\theta^*)}{(\alpha^* + 1)^2}\right).$$

The agent lacks sufficient information to figure out the demand elasticity α^*, and he only knows that $\alpha^* \in \mathcal{A}$, so the true state can not be perfectly detected. If the agent observes a high price, it may that the demand is strong, but it may also be that the high price comes from a low demand elasticity. The agent is endowed with some bias $\tau : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}_+$ and adopts the biased updating rule to update beliefs. At time t, the agent takes an action to match the true market fundamental based. [7] For example, he selects an r_t to maximize

$$V_t(r) = -\int_{\underline{\theta}}^{\overline{\theta}} (r - \theta)^2 \times \mu_t(\theta)\, d\theta.$$

The optimal action is $r_t = E_t\theta$, where the expectation is taken under the subjective belief at time t. For instance, an analyst wants to issue a report r_t that can best reflect the current economic condition, but how he evaluates the economy may be unconsciously influenced by his bias, so he may end up issuing a report that is also biased. The biased reporting of analysts has been documented in many research works, for example, Brous and Kini (1993) and Easterwood and Nutt (1999). By flexibly defining the τ function, we can accommodate a variety of biases. For example, (i) if $\tau = \delta_\theta$, where δ denotes the Dirac delta function, it means that the agent seeks to confirm a specific state θ. (ii) if τ is a constant function, it implies that there is no bias, so the agent adopts the Bayes rule to update his beliefs. (iii) if τ is an increasing (or decreasing) function, the agent is upward (or downward) biased and updates his beliefs to justify a state as higher (or lower) as possible. Unlike previous examples, it is not immediately clear how different biases will lead to different model perceptions and limit beliefs. This example will revisited in Section 6.

4 Discussion of the Model

In this section, I discuss some the critical assumptions about biased learning given in section 2. This section focuses on the two most important elements, the model ambiguity and the bias, and discusses their motivations and interpretations. This section also provides a brief discussion of other modeling assumptions.

[7] A brief discussion of the decision rule is in the next section.

Discussion of the Ambiguity

One key element is that individuals are ***ambiguous*** about the true signal-generating process. The presence of model ambiguity can come from multiple sources. From a frequentist perspective, ambiguity can be a result of identification problems. If every model in $\mathcal{A}$ can match the true signal distribution under some state, that is,

$$\forall \alpha \in \mathcal{A}, \quad \exists \theta \in \Theta \text{ s.t. } f\left(s|\alpha, \theta\right) = f\left(s|\alpha^*, \theta^*\right) \quad \text{for all } s \in S,$$

then it is impossible for individuals to figure out the correct model based on the long-run signal frequency. [8]Below is a simple example.

Example 6. [Identification Problem] In Example 3, models $fair$ and $unfair$ induce the following signal distributions

$fair$	G	B
H	p	$1-p$
L	$1-p$	p

$unfair$	G	B
H	$1-p$	p
L	p	$1-p$

, where $p > \frac{1}{2}$.

Suppose that the true signal distribution is $f^* = (p, 1-p)$. Then both models are consistent with the true distribution. For model $fair$, the distribution under state H matches the true distribution; for model $unfair$, the distribution under state L also matches. Therefore, a frequentist cannot identify the true model based on the long-run signal frequency.

From a Bayesian perspective, even if $\mathcal{A}$ is incorrectly specified, but every model in $\mathcal{A}$ can match the signal distribution equally well in terms of relative entropy, a Bayesian individual with prior on $\mathcal{A}$ is still unable to eliminate any model in the limit.[9] In the previous example, suppose instead that the true model is $f^* = (1/2, 1/2)$, so neither model can match the true distribution. We notice that both models are symmetric with respect to f^*, in other words, they are "equally wrong", so a Bayesian individual is still unable to exclude either model. Another justification is that individuals remain ambiguous simply because they are unable to pin down a specific prior, hence every feasible model can be correct as long as it generates the observed history with a strictly positive probability. In this paper, I am agnostic about how the model set $\mathcal{A}$ is formed. Instead of imposing specific restrictions on the model set $\mathcal{A}$, this paper treats $\mathcal{A}$ as part of the learning environment and aims to derive a general characterization for *all* possible forms of $\mathcal{A}$ (e.g., correctly or incorrectly specified, matching the long-run frequency or not). It is worth noting that in addition to just being general, this flexibility accommodates situations where individuals will adopt a less-likely (but still possible) explanation. These situations may be less interesting in environments with "rational" individuals, but they seem more prevalent for biased individuals who have incentives to distort information.

[8]Similar justifications are seen in Battigalli et al (2015). They introduced a self-confirming equilibrium under model uncertainty, in which players remain ambiguous about strategies that match the true long-run frequency.

[9]In generic cases, beliefs on the model space will not converge, and *any* belief on $\mathcal{A}$ can be an accumulation point. Bunke and Milhaud (1998) provided one example, and the idea of Berk-Nash equilibrium proposed by Esponda and Pouza (2016) is built on a similar spirit.

Discussion of the Bias

Another key element is that individuals are endowed with some ***bias*** when processing new information. First, in terms of **motivation**, the bias is motivated by the idea that people attach intrinsic values to some states of the world and hence will process information to justify those states. For example, individuals enjoy feeling good about themselves, so they tend to listen to good news and overlook bad news to maintain positive self-esteem. Second, in terms of its **impact on updating**, the bias pins down a posterior. Under model uncertainty, there are two approaches to pin down a posterior: (i) individuals can assign a prior over models, and (ii) individuals can employ a mechanism to determine which model to update. This paper follows the second approach, and the bias works as a model-selection mechanism. To draw a parallel, under maximum likelihood updating, individuals select a model to maximize the likelihood function; similarly, under biased updating, individuals select a model to maximize the bias, which is represented by a belief-utility function.[10]

This paper focuses on the dynamics of beliefs, but given that individuals hold biased beliefs, it is also natural to ask about **actions**. One possibility is that individuals are *naive* and evaluate all choices according to their current beliefs and model perceptions as in Example 5. The naivety is reflected in the following aspects: (i) when forming beliefs, individuals only seek to justify their bias, but do not consider the impacts on actions; (ii) when making decisions, individuals inherit beliefs from the learning stage and make decisions accordingly, but do not manipulate their biases or act in a strategic manner. In Example 5, it implies that the analyst's biased reporting only reflects his interpretation of the information and does not embed any strategic concerns. One natural question is that what do we gain from use of the naivety assumption? First, it has the advantage of being highly tractable, since we can determine all relevant actions by keeping track of the agent's beliefs. Second, it seems more in line with how the bias normally works compared with a sophisticated assumption. Many experimental findings suggest that people do not get biased in a conscious or a strategic manner, so it is difficult to reconcile the idea that individuals are biased but are also sufficiently sophisticated to manipulate their own biases. [11] Due to these reasons, it seems natural to use the naive rule as a benchmark if we want to talk about actions. Under this assumption, Section 7 analyzes some examples to illustrate how biased individuals make decisions.

Discussion of Other Assumptions

Some ***other assumptions*** also merit discussion. First, individuals face the constraint that beliefs are updated by **Bayes rule model-wise**. Notice that when maximizing their belief utilities, individuals must face some constraint, since otherwise they can perfectly self-deceive, which makes the

[10]For references on the belief-utility function, see Caplin and Leahy (2001), Benabou and Tirole (2002), Brunnermeier and Parker (2005), and Kőszegi (2006)

[11]Pronin et al (2002) showed that people often fail to recognize their own biases (e.g., the self-serving bias, the "better-than-average" effect) even if they can recognize that the bias exists for others. Babcock et al (1995) showed that the self-serving bias still exists even if bargainers have incentives to process the information objectively, which implies that "the bias does not appear to be deliberate or strategic" (Babcock and Lowenstein 1997). Lowenstein and Adler (1995) implied that people do not have good awareness of the endowment effect, that is, they tend to underestimate that they would become biased toward an object once they received it.

problem trivial and unrealistic. The model-wise Bayes constraint can be regarded as a "rationality constraint", which requires that individual would still update in a "rational" way if they knew how signals are generated. Admittedly, it remains debatable whether this is a sensible constraint to impose, but it serves as a natural first-step. Second, individuals have **perfect recall** so that they can go back and revise how to interpret all previous signals. It also remains controversial whether perfect or bounded memory (usually with one-period) seems most appropriate, but the perfect recall assumption has the advantage of enabling a simple characterization of limit beliefs. Third, individuals may **change their world views** (i.e., model perceptions) each time they receive a new signal. This assumption implies some "discontinuity" of their world views, which needs justifications. First, in many situations, the perceived models will converge, hence a sudden shift of the world view is less frequent in the long run. Second, the shifting of world views is not new to the literature. One prominent example is the maximum likelihood updating, where individuals can modify their model perceptions after observing new information. [12] Third, we can re-interpret the model as that individuals observe all signals first, and then decide their model perceptions and beliefs (assuming that there is no later information). This re-interpretation circumvents the problem and results in the same belief as in the benchmark model. Some other assumptions are worth discussing, e.g., individuals hold a fixed model set and a fixed bias, and Section 8 will discuss some extensions of the benchmark model by relaxing these assumptions.

5 Relative Entropy and Related Concepts

This section introduces concepts that will be used in later sections. One preliminary characterization of limit beliefs is also presented.

Definition 1. Define the ***relative entropy*** of state θ under model α as

$$\mathcal{R}(\alpha,\theta) \equiv \int_S \log\left[\frac{f(s|\alpha^*,\theta^*)}{f(s|\alpha,\theta)}\right] f(s|\alpha^*,\theta^*)\,dv(s)$$

Further denote $r(\alpha,\theta) \equiv \mathcal{R}(\alpha,\theta) - \min_{\theta'\in\Theta}\mathcal{R}(\alpha,\theta')$, called ***information potential*** of state θ under model α.[13]

The *relative entropy* of state θ under model α provides a measure of closeness between distribution $f(s|\alpha,\theta)$ and the true signal distribution, $f(s|\alpha^*,\theta^*)$. Fixing a model α, if the relative entropy of θ is lower, it means that state θ induces a distribution that is closer to the true distribution. The left graph of Figure 1 provides an illustration. The solid curve represents the true distribution, and the dashed and dotted curves represent the distributions of states θ and θ_1 under

[12] Here are some other examples. Ortoleva (2012) provided a axiomatic foundation for a non-Bayesian updating rule, where individuals can change their paradigms (i.e., the priors) if they observe some event to which their original paradigms assign a small probability. Galperti (2019) studied a persuasion problem under a similar assumption that individuals can shift their paradigms abruptly.

[13] This concept is well-defined since it is easy to verify that $\mathcal{R}(\alpha,\theta)$ is a continuous function, so the minimum can be obtained.

an alternative model, model α_1. In this graph, the true distribution is "closer" to $f(s|\alpha_1, \theta_1)$ than to $f(s|\alpha_1, \theta)$, and this corresponds to the relation $\mathcal{R}(\alpha_1, \theta_1) < \mathcal{R}(\alpha_1, \theta)$. [14] *Information potential* is a normalized version of relative entropy, where every entropy-minimizing state is normalized to have zero potential. The right graph of Figure 1 provides an illustration. This graph plots the relative entropy under α_1 over a continuum of states, where state θ_1 has the minimum relative entropy. The information potential of state θ is equal to its relative entropy minus a normalized term, which is the minimum relative entropy under α_1. As will be seen later, such normalization is useful since it enables comparison *across* models. In the rest of this paper, I will use the term "*zero-potential state*" and "(relative) *entropy-minimizing state*" interchangeably.

Below is a simple example showing how to find zero-potential states.

Example 7. (Fairness and Preferences for Redistribution-II) Consider the case in Example 3. Suppose that the true signal distribution $f^* = (q, 1-q)$ with $q > 1/2$, where the first coordinate represents the probability of feedback G, and the second represents the probability of feedback B. In this example, it is possible that q is different from p, so there may exist model misspecification. Considering $\alpha = fair$, by Definition 1, we have

$$\begin{aligned}\mathcal{R}(fair, H) &= f^*(G) \times \log\left(\frac{f^*(G)}{f(G|fair, H)}\right) + f^*(B) \times \log\left(\frac{f^*(B)}{f(B|fair, H)}\right) \\ &= q \times \log\left(\frac{q}{p}\right) + (1-q) \times \log\left(\frac{1-q}{1-p}\right).\end{aligned}$$

Analogously, $\mathcal{R}(fair, L) = q \times \log\left(\frac{q}{1-p}\right) + (1-q) \times \log\left(\frac{1-q}{p}\right)$. Recalling that $p, q > 1/2$, it follows that

$$\mathcal{R}(fair, H) - \mathcal{R}(fair, L) = (2q-1) \times \log\left(\frac{1-p}{p}\right) < 0,$$

so state H is the entropy-minimizing state, hence zero-potential state, under model $\alpha = fair$. This fact can be explained more intuitively. Under $\alpha = fair$, it is more likely to receive good feedback (G) than bad feedback (B) if the worker has high ability (H), and the opposite if the worker has low ability (L). Recall that the true distribution is $(q, 1-q)$ with $q > 1/2$, so it is better approximated by the distribution under state H under $\alpha = fair$. Therefore, state H is the zero-potential state under $\alpha = fair$. It can be further shown that if the worker updates according to $\alpha = fair$, he will hold a degenerate belief on the zero-potential state, state H, in the limit.

Previous examples have simple structures such that limit beliefs can be calculated explicitly. To characterize limit beliefs for more general cases, I introduce the following definition.

Definition 2. A sequence of probability measures $\{\mu_t\}$ is ***asymptotically carried*** on set $\mathcal{U}$ if $\lim \mu_t(U) = 1$ for all open sets $U \supset \mathcal{U}$.

This concept was also used by Berk (1966) to characterize limit beliefs under incorrect models. Roughly speaking, if a sequence of probability measures are asymptotically carried on $\mathcal{U}$, then only

[14] A more concrete example is the class of normal distribution, $f(s|\alpha, \theta) = \mathcal{N}(\theta, \alpha^2)$. To generate the relation in Figure 1, we simply set θ_1 very close to θ^* and θ very far from θ^*.

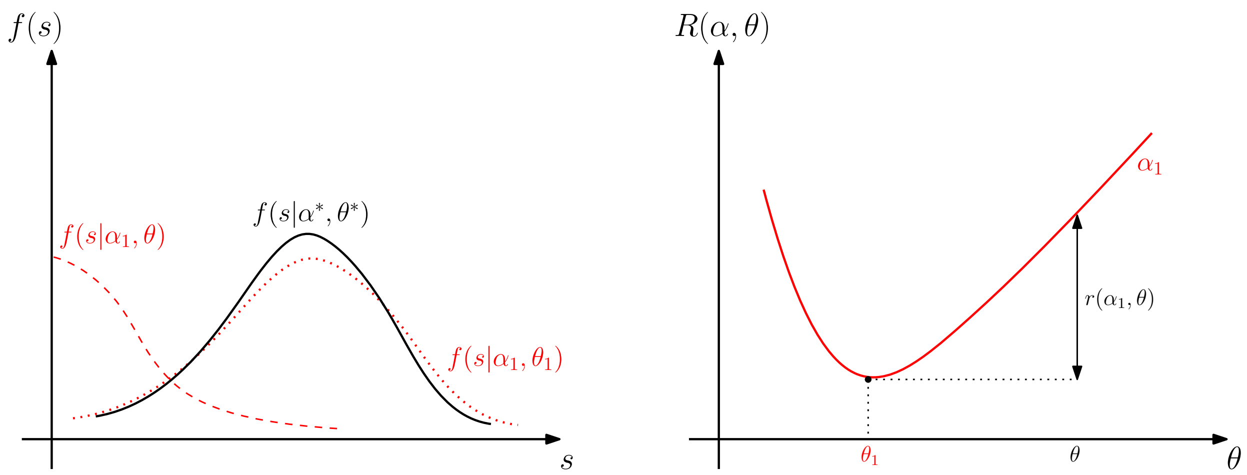

Figure 1: Relative Entropy & Information Potential

states in $\mathcal{U}$ will have non-zero weights in the limit, and states outside of $\mathcal{U}$ will be attached zero weight. Lemma 1 says that limit beliefs are asymptotically carried on zero-potential states.[15]

Lemma 1. *Let $\mathcal{U}_\mathcal{A}$ be the zero-potential states under models in $\mathcal{A}$, that is,*

$$\mathcal{U}_\mathcal{A} = \{\theta \in \Theta : r(\alpha, \theta) = 0 \text{ for some } \alpha \in \mathcal{A}\}.$$

Then all individuals' beliefs are asymptotically carried on set $\mathcal{U}_\mathcal{A}$ $\mathbb{P}^*$-almost surely.

To understand the intuition, it is useful to review the case in Berk (1966). Berk showed that if individuals update beliefs according to a possibly incorrect model, limit beliefs will settle on zero-potential states under that model. If individuals perceive a model α, and if limit beliefs attach positive weights to a state θ_0, we must have $r(\alpha, \theta_0) = 0$, or equivalently,

$$\theta_0 \in \arg\min_{\theta \in \Theta} \mathcal{R}(\alpha, \theta) \quad \mathbb{P}^* - a.s..$$

Intuitively, relative entropy measures the distance between state θ's induced distribution under α and the true distribution. Therefore, Berk's result captures the idea that limit beliefs will settle on states that generate the closest distributions to the true signal distribution. As in Example 7, under $\alpha = fair$, state H's induced distribution best approximates the true distribution, so the worker will settle on state H in the limit. Back to our model, when individuals can perceive *a set of* models, $\mathcal{A}$, the intuition still holds, and limit beliefs will settle on the set of zero-potential states under models in $\mathcal{A}$. When $\mathcal{A}$ consists of finite number of models, Lemma 1 comes from a simple union. When $\mathcal{A}$ consists of infinitely many models, the main difficulty is to establish that beliefs will converge *uniformly* for all models in $\mathcal{A}$, which is discussed in the Appendix.

[15] It is conceivable that all results still hold in a stronger sense of convergence. Bunke and Mihaud (1998) used a stronger notion of concentration which is expressed in terms of the expected distance to set $\mathcal{U}$ under measures μ_ts. They proved a stronger version of Berk's (1966) result using this notion.

6 A Complete Characterization of Limit Beliefs

Lemma 1 provides a coarse characterization of beliefs, which applies to all possible biases but does not address how biases influence model perceptions and beliefs. This section provides a fuller characterization of limit beliefs and model perceptions. This section first discusses a baseline case where individuals have one biased state. The discussion is then extended to the case with general biases.

6.1 Baseline Case: Single Biased State

Consider first the case in which individuals have only one biased state, that is, $\tau = \delta_\theta$ for some state $\theta \in \Theta$. Individuals whose biased state is θ are referred to as bias-θ individuals. Their updating rule is:

$$\forall \theta' \in \Theta: \quad \mu_t^\theta (\theta') = \mu_t \left(\theta' | \alpha_t^\theta\right) \qquad \text{where } \alpha_t^\theta \in \arg\max_{\alpha \in \mathcal{A}} \mu_t (\theta | \alpha),$$

where μ_t^θ denotes the time-t belief of a bias-θ individual, and α_t^θ denotes the model adopted by the individual at time t. As discussed previously, under this updating rule, individuals seek to confirm one state, so they update beliefs according to the model that delivers the highest possible likelihood for the biased state. The lemma below characterizes limit model perceptions.

Lemma 2. *For all $\theta \in \Theta$, let $\mathcal{A}_\infty^\theta$ denote the set of limit points of $\left\{\alpha_t^\theta\right\}$, we have:*

$$\mathcal{A}_\infty^\theta \subset \arg\min_{\alpha \in \mathcal{A}} r (\alpha, \theta) \quad \mathbb{P}^* - a.s.$$

that is, individuals will asymptotically update according to the models that minimize the information potential of their biased states.

The intuition behind Lemma 2 can be summarized as follows. When maximizing the likelihood of some bias θ, individuals will eventually select a model α that minimizes the "distance" between θ and the states that limit beliefs concentrate on, where the "distance" is measured by the difference in relative entropy. From Lemma 1, limit beliefs will concentrate on zero-potential states. By definition 1, minimizing the "distance" to zero-potential states is equivalent to minimizing the information potential. As such, the bias-justifying behavior will lead individuals to select a model that minimizes the information potential of the biased state θ, which implies Lemma 2.

If $\mathcal{A}$ satisfies the condition that every model has the same minimum relative entropy,[16] we can further replace the $r(\alpha, \theta)$ in Lemma 2 with $\mathcal{R}(\theta, \alpha)$. In this case, individuals will adopt models that minimize the *relative entropy* of their biases in the limit. One simple example is when every model in $\mathcal{A}$ can match the true signal distribution. However, for a general $\mathcal{A}$, the limit model may not be the one that minimizes the relative entropy (e.g., when individuals are not ambiguous in a frequentist or a Bayesian manner). Below is one simple example.

[16] Formally, it means that $\forall \alpha, \alpha' \in \mathcal{A}$ we have $\min_{\theta \in \Theta} \mathcal{R}(\theta, \alpha) = \min_{\theta \in \Theta} \mathcal{R}(\theta, \alpha')$.

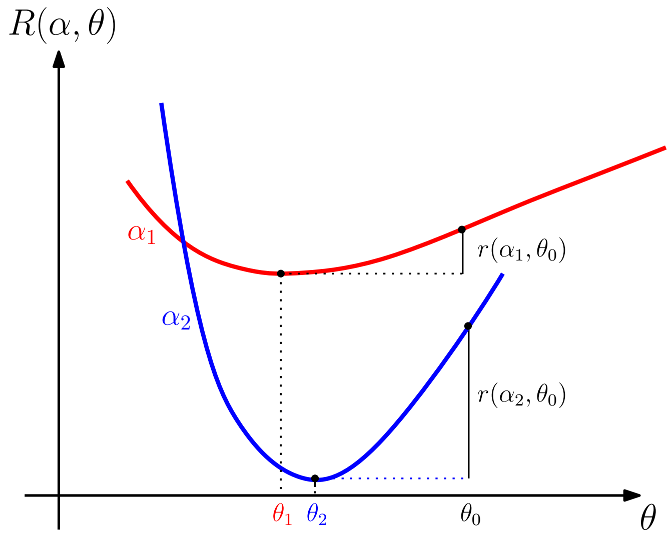

Figure 2: Limit beliefs

Example 8. Suppose that $\Theta = \{\theta_1, \theta_2\}$, $S = \{s_1, s_2\}$ and $\mathcal{A} = \{\alpha_1, \alpha_2\}$, where

α_1	s_1	s_2
θ_1	$\frac{1}{2}$	$\frac{1}{2}$
θ_2	$\frac{2}{5}$	$\frac{3}{5}$

α_2	s_1	s_2
θ_1	$\frac{1}{4}$	$\frac{3}{4}$
θ_2	$\frac{3}{4}$	$\frac{1}{4}$

,

The true signal distribution is $f^* = (1/3, 2/3)$, where the first coordinate denotes the probability of s_1 and the second coordinate denotes the probability of s_2. In this example, individuals are endowed with a *misspecified* model set (i.e., neither model matches the long-run frequency). It can be seen that θ_2 and θ_1 minimize the relative entropy under α_1 and α_2 respectively. Suppose that an individual has a biased state θ_1. If he chooses to believe in model α_1, he will form a belief δ_{θ_2} in the limit. If he chooses to believe in model α_2, he will form a belief δ_{θ_1} in the limit, which leads him to justify his bias perfectly. As a result, the bias-θ_1 individual will almost surely interpret signals according to model α_2 in the limit. However, it can be verified that model α_1 induces a lower relative entropy than model α_2 at state θ_1.

Based on Lemma 1 and Lemma 2, limit beliefs can be characterized by the theorem below.

Theorem 1. *For all $\theta \in \Theta$, define a set $\mathcal{U}_{\mathcal{A}}^{\theta}$ as follows*

$$\mathcal{U}_{\mathcal{A}}^{\theta} = \left\{\theta' \in \Theta : r\left(\alpha', \theta'\right) = 0 \text{ where } \alpha' \in \arg\min_{\alpha \in \mathcal{A}} r\left(\alpha, \theta\right)\right\}.$$

Then bias-θ individuals' beliefs are asymptotically carried on $\mathcal{U}_{\mathcal{A}}^{\theta}$ $\mathbb{P}^*$-almost surely.

To summarize, limit beliefs will concentrate on zero-potential states under the potential-minimizing models of the biased state. Figure 2 provides an illustration. Suppose that individuals only consider two models as possible $\mathcal{A} = \{\alpha_1, \alpha_2\}$. From Lemma 2, a bias-$\theta_0$ individual will asymptotically adopt the model that minimizes the potential for state θ_0, which is model α_1. From Theorem 1, the bias-θ_0 individual will eventually settle on the zero-potential state under α_1, that is, state θ_1. In other words, individuals with bias θ_0 will hold a degenerate belief on state θ_1 almost surely.

Remark 1. If $\mathcal{U}_{\mathcal{A}}^{\theta}$ contains a unique state, bias-θ individuals will almost surely hold a degenerate belief in the limit. When $\mathcal{U}_{\mathcal{A}}^{\theta}$ contains multiple states, there is no guarantee that beliefs will converge to a well-defined limit. In most interesting cases, beliefs of bias-θ individuals will oscillate on set $\mathcal{U}_{\mathcal{A}}^{\theta}$ in the limit. This comes from the fact that the log likelihood ratio between any two entropy-minimizing states constitutes a zero-mean random walk, so it oscillates between $-\infty$ and $+\infty$ with probability 1 (see Berk 1966 for a concrete example).

Below is a numerical example based on Theorem 1.

Example 9. [Inferring the market fundamental-II] Consider the case described by Example 5. I assume that $\Theta = [-10, 10]$ with the true state $\theta^* = 2$, the true demand elasticity $\alpha^* = 1$, and the market volatility $\sigma(\theta) = 1$. Therefore, the true distribution of market price is:

$$\log P_t \sim \mathcal{N}\left(1, \frac{1}{2}\right)$$

The perceived set of models is $\mathcal{A} = \left[\frac{1}{1+\delta}, 1+\delta\right]$ where a higher δ corresponds to a higher degree of ambiguity. Suppose that every agent has one biased state $\theta \in \Theta$. For example, investors try to persuade themselves of some market condition in order to support some investment tendency; policy makers want to justify some economic policy by persuading people (or even themselves) of a specific market condition. For simplicity in exposition, I assume that the degree of ambiguity δ is not too large (to avoid a kink in the expression of relative entropy). It can be verified that

$$r(\alpha, \theta) = \left(\frac{\theta}{\alpha+1} - 1\right)^2.$$

Letting $\theta(\alpha)$ denote the *zero-potential state* under model α, we have:

$$\theta(\alpha) = \alpha + 1 \quad \text{for all } \alpha \in \mathcal{A}$$

Further defining $\alpha(\theta)$ to be the *potential-minimizing model* of state θ, we have:

$$\alpha(\theta) = \begin{cases} \frac{1}{1+\delta} & \theta \in \left[-10, \frac{2+\delta}{1+\delta}\right] \\ \theta - 1 & \theta \in \left[\frac{2+\delta}{1+\delta}, 2+\delta\right] \\ 1+\delta & \theta \in [2+\delta, 10] \end{cases},$$

which gives the limit model perception for each possible bias from Lemma 2. Using results from Theorem 1, the limit belief carrier for each bias is:

$$\mathcal{U}_{\mathcal{A}}^{\theta}=\begin{cases}\left\{\frac{2+\delta}{1+\delta}\right\} & \theta\in\left[-10,\frac{2+\delta}{1+\delta}\right]\\ \{\theta\} & \theta\in\left[\frac{2+\delta}{1+\delta},2+\delta\right]\\ \{\delta+2\} & \theta\in[2+\delta,10]\end{cases}$$

For all individuals with bias $\theta\in\left[\frac{2+\delta}{1+\delta},2+\delta\right]$, they will successfully justify their biased states and hold a Dirac belief δ_θ in the limit. For individuals with bias too high (i.e., $\theta\geq 2+\delta$) or too low (i.e., $\theta\leq\frac{2+\delta}{1+\delta}$), they are unable to justify their biased states. Individuals with a very high bias will interpret price information according to the highest possible demand elasticity, $1+\delta$, and believe in the corresponding zero-potential state, $2+\delta$. Here is the intuition. Recall that a higher demand elasticity implies a lower price, and a higher market fundamental implies a higher price. When faced with the same price sequence, a higher demand elasticity enables individuals to justify a higher market fundamental, so individuals with high (or low) bias will update beliefs according to the model with high (or low) demand elasticity.

Remark 2. As the degree of ambiguity $\delta\to 0$, the limit belief carrier $\mathcal{U}_{\mathcal{A}}=\left[\frac{2+\delta}{1+\delta},2+\delta\right]\to\{\theta^*\}$,which means that limit beliefs become more concentrated around the true state. However, for any degree of ambiguity $\delta>0$, polarization arises on states in $\mathcal{U}_{\mathcal{A}}$. Depending on their biases, individuals can hold any limit posterior δ_θ with $\theta\in\mathcal{U}_{\mathcal{A}}$, so they can totally disagree on the correct state. Unlike Example 4, in this example, by reducing ambiguity, we only decrease the range of disagreement but can never eradicate the disagreement itself.

6.2 Comparative Statics: Asymptotic Learning and Polarization

Based on the remark under Example 9, this subsection discusses what happens as ambiguity varies. The degree of ambiguity $d(\mathcal{A})$ is defined to be the diameter of $\mathcal{A}$. [17]For simplicity in exposition, this subsection assumes that the perceived model set $\mathcal{A}$ contains the true model α^*. First, as the degree of ambiguity approaches zero, all individuals will learn in a weak sense.

Proposition 1. *As ambiguity vanishes, limit beliefs converge **weakly** to the correct belief δ_{θ^*}, that is,*

$$\lim_{d(\mathcal{A})\to 0}\left[\lim_{t\to\infty}\left|\int_\Theta h\times\mu_t^\theta dm-\int_\Theta h\times\delta_{\theta^*}dm\right|\right]\to 0\quad\mathbb{P}^*-a.s.$$

for all bounded and continuous function $h:\Theta\to\mathbb{R}$ and for all $\theta\in\Theta$.

The concept of convergence is an adapted version of convergence in weak topology. Roughly speaking, it says that as ambiguity vanishes, the limit belief carrier approaches the true state. We can have a stronger version of learning if the true state, θ^*, is a "dominant" state for all models within a small neighborhood of the true model, α^*. More precisely, θ^* is *locally dominant* if there exists some non-degenerate closed neighborhood $C\ni\alpha^*$ such that θ^* is the unique zero-potential state for all $\alpha\in C$.

[17]More precisely, $d(\mathcal{A})=\max_{\alpha,\alpha'\in\mathcal{A}}||\alpha-\alpha'||$, where $||.||$ denotes the relevant metric on the model space $\mathbb{A}$.

Proposition 2. *If θ^* is locally dominant, limit beliefs converge* ***strongly*** *to the correct belief when ambiguity is sufficiently small, that is,*

$$\sup \left\{ \left| \mu_t^{\theta}(E) - \delta_{\theta^*}(E) \right| : \theta \in \Theta, E \in \mathcal{B}_\Theta \right\} \to 0 \quad \mathbb{P}^* - a.s.$$

for all $\mathcal{A}$ such that $d(\mathcal{A}) < \varepsilon$ for some $\varepsilon > 0$.

The concept of convergence in Proposition 2 is strong convergence, which requires convergence on all measurable sets. It is a direct implication of Lemma 1. If the true state θ^* is locally dominant, $\mathcal{U}_{\mathcal{A}}$ becomes a singleton $\{\theta^*\}$ when the degree of ambiguity is sufficiently small. Therefore, all beliefs converge to the correct belief δ_{θ^*} almost surely. The intuition is that if the true state is locally dominant, then it can be perfectly identified within a small neighborhood around the true model. One important case where the local dominance condition holds is when the state space is finite.

Corollary 1. *If $|\Theta| < \infty$, beliefs converge strongly to the correct belief when ambiguity is sufficiently small.*

This corresponds to the case in Example 4. Recall that in this example, $\Theta = \{G, B\}$, and both optimists and pessimists can learn the true state when ambiguity is sufficiently small. To illustrate the idea, a special case of Example 4 is discussed as follows.

Example 10. [Optimistic and Pessimistic Investors-II] Suppose that $\lambda = 2$ and $\varepsilon = 1/2$, and that the true state $\theta^* = B$. Let $\mathcal{A} := \left[\frac{1}{1+\delta}, 1+\delta\right]$ denote the set of likelihood ratios induced by signal m. Every model $\alpha \in \mathcal{A}$ represents a likelihood ratio. It is easy to verify that as $t \to \infty$,

$$\frac{1}{t} \log \left[\frac{\mu_t(G|\alpha)}{\mu_t(B|\alpha)} \right] \to \frac{1}{2} \log \alpha - \frac{1}{6} \log 2 \quad \mathbb{P}^*\text{-}a.s..$$

For pessimistic investors, we have:

$$\frac{1}{t} \log \left[\frac{\mu_t^p(G)}{\mu_t^p(B)} \right] = \min_{\alpha \in \mathcal{A}} \frac{1}{t} \log \left[\frac{\mu_t(G|\alpha)}{\mu_t(B|\alpha)} \right] \to \frac{1}{2} \log \left(\frac{1}{1+\delta} \right) - \frac{1}{6} \log 2 < 0 \quad \mathbb{P}^*\text{-}a.s.,$$

which implies that $\mu_t^p(B) \to 1$ almost surely, so pessimists will almost surely learn the true state, state B. For optimistic investors, we have:

$$\frac{1}{t} \log \left[\frac{\mu_t^o(G)}{\mu_t^o(B)} \right] = \max_{\alpha \in \mathcal{A}} \frac{1}{t} \log \left[\frac{\mu_t(G|\alpha)}{\mu_t(B|\alpha)} \right] \to \frac{1}{2} \log(1+\delta) - \frac{1}{6} \log 2 \quad \mathbb{P}^*\text{-}a.s.. \tag{2}$$

When the degree of ambiguity is small such that $\delta < 2^{1/3} - 1$, the RHS of (2) is strictly negative, which implies that $\mu_t^o(B) \to 1$ almost surely. In this case, correct learning arises for both optimists and pessimists.

The intuition for Corollary 1 is straightforward. When the state space is finite, the distance between the true signal distributions under any two states is bounded away from 0. Therefore, the true state can be identified when all models are sufficiently close to the true model, that is, when the degree of ambiguity is sufficiently small. However, when the state space is rich enough (e.g., uncountably infinite), correct learning may **not** occur. Moreover, polarization can occur for any positive degree of ambiguity as in Example 9. I refer to state θ^* as *singular* if θ^* is the unique zero-potential state under all model $\alpha \in \mathbb{A}$ such that $r(\alpha, \theta^*) = 0$.[18] One simple example satisfying this property is that every model has a unique zero-potential state. We have:

Proposition 3. *If θ^* is singular and not locally dominant,* ***polarization*** *occurs for any positive degree of ambiguity, that is,*

$$\sup\left\{\left|\mu_t^{\theta}(E) - \mu_t^{\theta'}(E)\right| : \theta, \theta' \in \Theta, E \in \mathcal{B}_\Theta\right\} \to 1 \quad \mathbb{P}^* - a.s.$$

for all $\mathcal{A}$ such that $d(\mathcal{A}) > 0$.

Proposition 3 says that with probability 1, there exist individuals with different biases who *totally disagree* on some event. More precisely, there exists some event where one individual assigns probability 0 whereas another individual assigns probability 1; that is, individuals' limit beliefs can be mutually singular. Proposition 3 shows that learning and belief consensus can be fragile with respect to ambiguity. When individuals are trying to defend their biases, it is possible that a slight degree of ambiguity suffices to destroy agreement. The proof of Proposition 3 is not difficult. The singular assumption can be thought of as a regularity assumption. The main assumption driving the result is the assumption that θ^* is not locally dominant, so there exists an identification problem in any small neighborhood around the correct model. As a result, limit posteriors of differently biased individuals can disagree—— they may or may not learn the true state, depending on their biases.

6.3 General Case: Multiple Biased States

This section characterizes limit beliefs in a more general framework where individuals have more than one biased state. The characterizations are parallel to the benchmark case, so readers may choose to skip this part. Recall that when each individual only has one biased state, the limit models are those that minimize the potential of *the* biased state. The following lemma extends this result to the case with multiple biased states.

Lemma 3. *For all $\tau \in \mathcal{T}$, let $\mathcal{A}_\infty^\tau$ denote the set of limit points of $\{\alpha_t^\tau\}$, we have:*

$$\mathcal{A}_\infty^\tau \subset \arg\min_{\alpha \in \mathcal{A}} \left[\min_{\theta \in supp(\tau)} r(\alpha, \theta)\right] \quad \mathbb{P}^* - a.s.$$

that is, bias-τ individuals will asymptotically update according to the models that minimize the minimum information potential of their biased states.

[18] In mathematical language, $\{\theta \in \Theta : r(\alpha, \theta) = 0 \text{ where } \alpha \text{ solves } r(\alpha, \theta^*) = 0\} = \{\theta^*\}$

Lemma 2 says that in the case with multiple biased states, individuals seek to minimize the *minimum* potential of their biased states. This comes from the fact that the biased states with the minimum potential will eventually dominate all other biased states, so only the minimum-potential biased states are relevant in the limit. Following the same logic as in Lemma 2, individuals will seek to minimize the potential of these relevant biased states, or equivalently, they will minimize the minimum potential of biased states in the limit. Based on this lemma, limit beliefs can be characterized as below.

Theorem 2. *For all $\tau \in \mathcal{T}$, define a set $\mathcal{U}_{\mathcal{A}}^{\tau}$ as follows,*

$$\mathcal{U}_{\mathcal{A}}^{\tau} = \left\{ \theta' \in \Theta : r\left(\alpha', \theta'\right) = 0 \text{ where } \alpha' \in \arg\min_{\alpha \in \mathcal{A}} \left[\min_{\theta \in supp(\tau)} r\left(\alpha, \theta\right) \right] \right\}.$$

Then bias-τ individuals beliefs are asymptotically carried on $\mathcal{U}_{\mathcal{A}}^{\tau}$ $\mathbb{P}^*$-almost surely.

Theorem 2 is a parallel statement of Theorem 1. It says that limit beliefs will settle on zero-potential states under models that minimize the minimum potential of biased states. Below is an example example that explains how to find limit beliefs.

Example 11. Suppose that $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, $S = \{s_1, s_2\}$ and $\mathcal{A} = \{\alpha_1, \alpha_2\}$, where

α_1	s_1	s_2
θ_1	1/10	9/10
θ_2	7/8	1/8
θ_3	2/5	3/5
θ_4	3/4	1/4

α_2	s_1	s_2
θ_1	7/9	2/9
θ_2	1/10	9/10
θ_3	3/4	1/4
θ_4	3/5	2/5

,

and the true signal distribution is $f^* = (1/2, 1/2)$. An individual is endowed with bias $\tau = \frac{1}{2}1_{\theta_1} + \frac{1}{2}1_{\theta_2}$, which means that this individual seeks to confirm either state θ_1 or state θ_2, and he assigns equal weights to these two biased states. To calculate his limit beliefs, we follow the steps below.

First, *find the zero-potential state under each model.* It can be verified that θ_3 is the zero-potential state under α_1, and θ_4 is the zero-potential state under α_2.[19] In this case, neither model can justify the individual's biased states in the limit.

Second, *find the minimum-potential biased state under each model.* Under model α_1, the

[19]Intuitively, under model α_1, θ_3 induces a distribution $(2/5, 3/5)$ that is closer to the true distribution, $(1/2, 1/2)$, than any other state, so θ_3 is the zero-potential state under α_3. Analogously, we can see that θ_4 is the zero-potential state under α_4.

information potential of θ_1 and θ_2 is

$$r(\alpha_1, \theta_1) = \mathcal{R}(\alpha_1, \theta_1) - \mathcal{R}(\alpha_1, \theta_3) = \frac{1}{2}\log\left(\frac{2/5}{1/10}\right) + \frac{1}{2}\log\left(\frac{3/5}{9/10}\right)$$
$$r(\alpha_1, \theta_2) = \mathcal{R}(\alpha_1, \theta_2) - \mathcal{R}(\alpha_1, \theta_3) = \frac{1}{2}\log\left(\frac{2/5}{7/8}\right) + \frac{1}{2}\log\left(\frac{3/5}{1/8}\right).$$

It is easy to see $r(\alpha_1, \theta_1) > r(\alpha_1, \theta_2)$, so state θ_2 is the minimum-potential biased state under α_1. Intuitively, it means that θ_2 is "closer" to the zero-potential state (i.e., state θ_3) than θ_1 in terms of their induced distributions. Similarly, we have

$$r(\alpha_2, \theta_1) = \frac{1}{2}\log\left(\frac{3/5}{7/9}\right) + \frac{1}{2}\log\left(\frac{2/5}{2/9}\right) \qquad r(\alpha_2, \theta_2) = \frac{1}{2}\log\left(\frac{3/5}{1/10}\right) + \frac{1}{2}\log\left(\frac{2/5}{9/10}\right).$$

Since $r(\alpha_2, \theta_1) < r(\alpha_2, \theta_2)$, state θ_1 is the minimum-potential biased state under α_2. The intuition follows analogously.

Third, *compare the information potential of the minimum-potential biased state under each model.* That is, we need to compare $r(\alpha_1, \theta_2)$ and $r(\alpha_2, \theta_1)$. It is easy to see that

$$r(\alpha_1, \theta_2) = \frac{1}{2}\log\left(\frac{2/5}{7/8}\right) + \frac{1}{2}\log\left(\frac{3/5}{1/8}\right) > \frac{1}{2}\log\left(\frac{3/5}{7/9}\right) + \frac{1}{2}\log\left(\frac{2/5}{2/9}\right) = r(\alpha_2, \theta_1),$$

so α_2 minimizes the minimum potential of the biased states. Lemma 3 and Theorem 2 further imply that this individual will almost surely adopt model α_2 and hold a degenerate belief on the zero-potential state under α_2 in the limit. As a result, his limit belief is δ_{θ_4}.

In some situations, we can further refine the characterization in Theorem 2 by comparing the belief utility $\tau(\theta)$ among all states in $\mathcal{U}^\tau_\mathcal{A}$.

Corollary 2. [Refinement of Theorem 2] *Suppose that every $\alpha \in \mathcal{A}$ has a unique zero-potential state and denote*

$$\mathcal{V}^\tau_\mathcal{A} = \arg\max\{\tau(\theta) : \theta \in \mathcal{U}^\tau_\mathcal{A}\}.$$

For all possible bias $\tau \in \mathcal{T}$, bias-τ individuals beliefs are asymptotically carried on $\mathcal{V}^\tau_\mathcal{A}$ $\mathbb{P}^*$-almost surely.

In other words, if the uniqueness of zero-potential state is assumed, limit beliefs will only accumulate states in $\mathcal{U}^\tau_\mathcal{A}$ with the highest belief utility.[20] Without the uniqueness assumption, it is possible that beliefs may also accumulate on states that do not have the highest weight (see Example 12).

[20] Actually, we only need a weaker assumption that for each τ, there exists a model that has a unique zero-potential state with the highest $\tau(\theta)$ in $\mathcal{U}^\tau_\mathcal{A}$

6.4 The Overconfidence Effect under Ambiguity

In Examples 1 and 2, we can see that the biased rule is different from Bayes rule (i.e., both correctly and incorrectly specified). In these examples, individuals only observe finite number of signals, so it is natural to ask: will the biased rule be identical to Bayes rule in the limit? The answer is no. This section discusses a novel overconfidence effect that can persist even in the limit. Even with arbitrarily many signals, biased individuals can exhibit strictly greater confidence than *any* Bayesian agent (i.e., with arbitrary feasible model perception) in the limit.

Example 12. [Overconfidence Effect under Ambiguity] Suppose that state space $\Theta = \{1, 0, -1\}$ with the true state $\theta^* = 0$. Signals take values in $S = \{g, m, b\}$. Individuals only consider two possible models α_1 and α_2, where

α_1	g	m	b
1	$1-\varepsilon$	$\frac{\varepsilon}{2}$	$\frac{\varepsilon}{2}$
0	$1/4$	$1/2$	$1/4$
-1	$\frac{\varepsilon}{2}$	$\frac{\varepsilon}{2}$	$1-\varepsilon$

α_2	g	m	b
1	$1/2$	$1/4$	$1/4$
0	$\frac{\varepsilon}{2}$	$1-\varepsilon$	$\frac{\varepsilon}{2}$
-1	$1/4$	$1/4$	$1/2$

,

where $\varepsilon > 0$ is sufficiently small (e.g., $\varepsilon = 0.01$). For simplicity, I assume a symmetric true model $f^* = (1/3, 1/3, 1/3)$. Individuals have an upward bias such that $\tau(1) > \tau(0) > \tau(-1)$. At each time t, each individual reports his estimation of the state $r_t \in \{-1, 0, 1\}$ which corresponds to the state he assigns the highest probability to. [21]

Fact 1: If the individual is a Bayesian, the expected report $\mathbb{E}^* r_t$ converges to 0 almost surely for every possible model perception.

If the individual adopts model α_1, his beliefs will almost surely settle on state 0, the unique entropy-minimizing state under α_1, which implies $r_t \to 0$ almost surely (hence $\mathbb{E}^* r_t \to 0$ almost surely). If the individual adopts model α_2, beliefs will oscillate between state 1 and -1, which are both entropy-minimizing states under α_2. Furthermore, it can be shown that limit beliefs will concentrate around 1 and -1 each with probability 1/2. Therefore, $\mathbb{E}^* r_t$ converges 0 almost surely (the details are shown in the Appendix).[22]

Consider a third party who collects reports from a large group and wants to infer the true state based on the average report. When individuals are Bayesian, the third-party will asymptotically conclude that the true state is 0 for *all* possible model perceptions within the group. From this perspective, the model uncertainty in this example seems harmless. However, the following fact contradicts this conjecture.

[21] If two or more states have the highest probability, individuals adopt some tie-breaking rule (or randomize) when issuing their reports.

[22] Here is an intuitive way to understand the result. Under model α_2, the distribution in state 1, $(1/2, 1/4, 1/4)$, and the distribution in state -1, $(1/4, 1/4, 1/2)$, are symmetric with respect to the true distribution $f^* = (1/3, 1/3, 1/3)$. Intuitively, these two states are "equally" likely to be the true state, so limit beliefs will accumulate around each state with probability $1/2$. Here, I only use the word "accumulate" since beliefs will not converge (see the remark 1).

Fact 2: If the individual is ambiguous about the true model, the expected report $\mathbb{E}^* r_t$ converges to a *strictly* positive number almost surely.

This fact comes from the following arguments. In the limit, (1) if the belief under α_2 concentrates around state -1, this individual will update according to model α_1 (since he "prefers" state 0 to state 1), which leads to a report equal to 0; (2) if the belief under α_2 concentrates around state 1, this individual will adopt model α_2, which leads to a report equal to 1. Each case occurs with probability 1/2, so the expected report $\mathbb{E}^* r_t$ converges to 1/2. Consequently, this individual becomes strictly more optimistic than the case where he is not ambiguous.

The argument above shows an interesting result of overconfidence that cannot be accommodated by the Bayesian framework. It illustrates how ambiguity and bias-defending work together to produce the overconfidence effect. The basic intuition is that different models may have **complementary effects**, which in a similar spirit to the "good-news effect" as in Example 1. Whenever the "good news" occurs such that model α_2 leads to a high state, state 1, individuals can *exploit* this good news by interpreting signals according to α_2. Whenever the "bad news" occurs such that when model α_2 leads to a low state, state -1, individuals can *hedge against* this bad news by interpreting signals according to α_1. As can be seen, ambiguity accommodates the asymmetric treatment of signals under different signal paths, enabling individuals to exhibit greater confidence than any Bayesian individual.

Remark 3. Combining the results in this example and examples in previous sections, one observes that model ambiguity produces two related effects that can lead to overconfidence (relative to Bayesian case). The first is what I call "flexible effect". The ambiguity accommodates the flexibility in signal interpretations, so individuals with different biases can choose to interpret signals differently, which leads to greater confidence than the case with a common model perception (see Example 2 and Example 4). The second effect is the "complementary effect" in this section. This effect utilizes another feature of ambiguity: ambiguity allows every individual to adopt different interpretations under different signal realizations. Under appropriate conditions, models can complement each other, which enables a biased individual to exhibit strictly greater confidence than any Bayesian individual. In Appendix B, I provide detailed conditions under which the complementary effect occurs.

7 Examples: Decisions under the Biased Updating Rule

Previous sections focus on the dynamics of beliefs, but it is also interesting to see how actions are different under the biased rule. To provide a fully satisfactory answer, we may need to build a decision theoretical foundation, which is beyond the scope of this paper. To shed light on how actions might be chosen, this section discusses some examples under the assumption that individuals are naive, that is, they choose what is optimal according to their current belief and model perception. This naivety assumption was discussed in Section 4 and seems to be a natural benchmark. Under this assumption, I present two examples—one static and one dynamic. The static example compares the biased rule with the full Bayesian rule and the maximum likelihood rule in their action implications. The dynamic example discusses the issue of dynamic (in)consistency.

7.1 Differences from the Full Bayesian and Maximum Likelihood Rule

One may wonder how the biased rule is different from two common updating rules under ambiguity—the full Bayesian updating (or FB) and maximum likelihood updating (or ML). Under FB, individuals keep the posteriors updated from all possible models. Under ML, individuals only keep the posteriors updated from the models that maximize the probability of generating the observed information. To facilitate the comparison, I consider a decision environment and examine how decisions differ under each of these updating rules. For the rest of this subsection, I assume that individuals select an action to maximize the minimum expected utility when there are multiple posteriors.

Example 13. [Difference among FB, ML & the Biased Rule] Suppose that $\Theta = \{L, M, R\}$ and $S = \{a, b\}$. An agent holds a flat prior over Θ and can observe a sequence of signals. After observing the signals, the agent chooses among three actions $A = \{l, m, r\}$ with the following payoffs

$$u(l, \theta) = 1_l(\theta), \quad u(r, \theta) = 1_r(\theta), \quad \text{and } u(m, \theta) = \frac{1}{2}.$$

Actions l and r are risky actions that only generate payoffs in specific states, and action m is a safe action that generates a constant payoff of 1/2. The agent is ambiguous over a set of models $\mathcal{A} = \{\alpha_1, \alpha_2\}$, where

$[\alpha_1]$	a	b
L	1/3	2/3
M	1/2	1/2
R	2/3	1/3

$[\alpha_2]$	a	b
L	3/4	1/4
M	1/2	1/2
R	1/4	3/4

.

Suppose that the agent observed 8 signals, which consist of 2 signal as and 6 signal bs. The observed signal frequency is hence denoted by $(1/4, 3/4)$. In the following discussion, we compare three different updating rules: the full Bayesian rule, the maximum likelihood rule, and the biased rule. To make the discussion not straightforward, for the biased rule, I assume that the agent is biased toward state M. This is because if the agent is biased toward state L or R, the optimal choice is just action l or r. Under this specification, it turns out that *each* updating rule induces a different optimal action:

updating rule	optimal action
Full Bayesian	$a^{FB} = m$
Maximum Likelihood	$a^{ML} = r$
Biased Rule	$a^{BS} = l$

.

(i) If the agent updates via the ***full Bayesian rule***, his optimal choice will be $a^{FB} = m$. To see that, we first notice that under the given signal structures, each signal can be flexibly interpreted as good news and bad news for any state. For example, model α_1 interprets signal a as bad news for state L, so receiving a signal a will decrease the likelihood of state L. On the contrary, model α_2 interprets signal a as "good news" for state L. If the agent adopts the full Bayesian rule, he is unwilling to take any risky action because he could always interpret the majority signals as bad news for the payoff relevant state. As such, the lowest probability on each state is less than $1/2$, so the agent will opt for the safe action, action m.

(ii) If the agent updates via the ***maximum likelihood rule***, his optimal choice will be $a^{ML} = r$. To see that, first note that under model α_2, state R induces a signal distribution $(1/4, 3/4)$, which perfectly matches the observed signal frequency; under model α_1, the best-matching distribution is $(1/3, 2/3)$, which is induced by state L and does not perfectly match the frequency. When there are sufficiently many signals (i.e., in this example, $n = 8$ suffices), posteriors will put sufficiently large probability on the state that induces the best match, so only the best-matching distribution matters for evaluating the likelihood of each model. As a consequence, model α_2 is more likely to generate the observed signals than model α_1, so the agent will update according to model α_2. It then follows that the agent will choose action r since state R induces the best-matching distribution.

(iii) If the agent updates via the ***biased rule*** and is biased toward state M, his optimal choice will be $a^{BS} = l$. The intuition is similar to the intuition behind Theorem 1. When trying to defend his bias, the agent will update according to the model which minimizes the "distance" between the signal distributions in his biased state and in the best-matching state of that model. In this example, the signal distribution in state M is $(1/2, 1/2)$ under both models. Intuitively, $(1/2, 1/2)$ is "closer" to $(1/3, 2/3)$, the best-matching distribution under α_1, than to $(1/4, 3/4)$, the best-matching distribution under α_2. Therefore, the biased agent will update according to model α_1. Under model α_1, the best-fitting state is state L, so the agent will choose action l.[23]

Remark 4. To sum up, the biased updating rule is different from the full Bayesian rule and maximum likelihood rule in the following aspects: (i) ML adopts a purely objective criterion and evaluates models according to their probability of generating the observed event, but BS adopts a purely subjective criterion and evaluates models according to their consistency with the bias; (ii) FB updates all models indiscriminately and leads to a set of posteriors, but BS only updates the bias-maximizing model and leads to a unique posterior. The difference between BS and FB is more evident if we compare the relevant posteriors for decisions. Under FB, individuals can use different posteriors to evaluate different actions, so the relevant posterior is choice-dependent; in contrast, under BS, individuals form a biased posterior first and then use that posterior to evaluate all choices, so the relevant posterior is choice-independent. These differences in beliefs can also lead to different actions as in Example 13. Though the example is built on a heuristic decision rule, it suggests that the biased rule has the potential of delivering new action implications.

[23] The detailed verification of these arguments are in the Appendix.

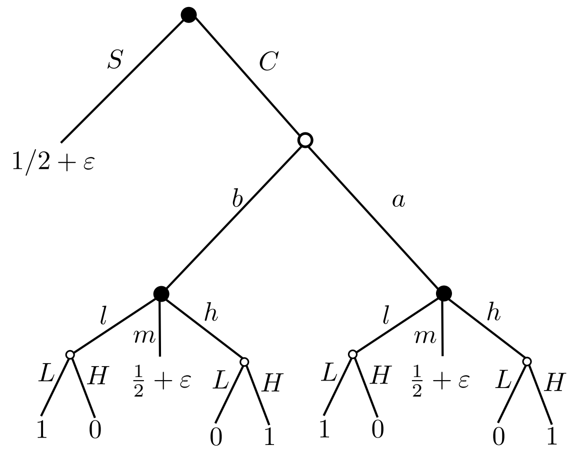

Figure 3: Dynamic Inconsistency

7.2 Dynamic (In)consistency of the Biased Agents

Another relevant question is what individuals would do if they had to choose at different points in time. This subsection looks at a dynamic decision environment, where individuals use the biased updating rule and make decisions according to the naive decision rule. One implication is that individuals may suffer from dynamic inconsistency. Below is a simple example.

Example 14. [Dynamic Inconsistency] The payoff relevant state space is $\Theta = \{H, L\}$. The signal space is $S = \{a, b\}$. An individual is faced with a simple problem of information acquisition as in Figure 3. Before the start of the game, the true state is drawn by the nature and is not known to the individual. At the initial decision node, the individual decides whether to observe a free signal—if he chooses not to observe (S), he gets a constant payoff of $1/2 + \varepsilon$, where $\varepsilon > 0$ and is sufficiently small; if he chooses to observe (C), the signal is then realized, and this individual needs to choose an action $x \in \{h, l, m\}$ based on the signal realization, where action h and l only generate a payoff of 1 in state H and L respectively, and action m generates a constant payoff of $1/2 + \varepsilon$. Finally, the true state is revealed, and the individual receives the payoff.

The individual is ambiguous about how to interpret the signal such that he perceives a set of likelihood ratios

$$\frac{P(a|H)}{P(b|H)} = \frac{P(b|L)}{P(a|L)} = \alpha \quad \text{where } \alpha \in \{1/4, 4\} \equiv \mathcal{A}.$$

Here, $\alpha = 4$ interprets signal a as more indicative of state H, whereas $\alpha = 1/4$ interprets signal a as more indicative of state L. The true model α^* belongs to $\mathcal{A}$. The individual starts with a prior $\mu_0 = (1/2, 1/2)$ and a model perception $\alpha_0 = 4$. The individual is biased toward state H and updates his belief according to the biased updating rule. The individual is risk neutral, so he evaluates every *plan* according to its expected payoff. Here, a plan is a complete description of the individual's action at each decision node. For example, plan Chl says that the individual will choose C at the initial node and then choose action h upon signal a and action l upon signal b.

(i) Suppose that the individual has the commitment power to be ***dynamically consistent***. That is, the individual is able to commit to a plan at the beginning and carry it out throughout the game. It is easy to see that the optimal plan is Chl, that is, the individual will accept the free signal and then choose what is optimal according to his initial signal interpretation.[24]

(ii) Suppose that the individual has no commitment power, then his choice will be ***dynamically inconsistent***. At the initial decision node, the individual naively evaluates all plans according to his current model perception, so the *ex-ante* optimal plan is still Chl, which requires him to observe the signal. Once the signal is realized, the individual will interpret the signal as good news for his biased state, state H. For example, if he observed a signal b, he would view b as more indicative of state H by changing his model perception to $\alpha_1 = 1/4$, which leads to a posterior $\mu_1 = \left(\frac{4}{5}, \frac{1}{5}\right)$. Given μ_1, action h generates the highest expected payoff, so the individual will choose action h instead of the planned action l. Therefore, even though Chl is *ex-ante* optimal, the individual is not willing to carry it out *ex-post*, which gives rise to dynamic inconsistency.

(iii) Suppose that the individual has no commitment power and follows a ***sophisticated*** decision rule. That is, the individual can correctly anticipate his future behavior, and he only chooses the best plan from the set of plans that can be actually carried out (see, Strotz 1955-1956, Siniscalchi 2011). We can think of the individual as consisting of two selves, the *ex-ante* self and the *ex-post* self. The *ex-ante* self evaluates a plan according to the ex-ante belief and model perception, but he understands how the *ex-post* self will react to information. The goal of the *ex-ante* self is to choose a plan that maximizes his own benefits. In this example, if he chooses to observe the signal, the only implementable plan is Chh. It is easy to see that $V(Chh) = 1/2 < 1/2 + \varepsilon = V(S)$, where $V(p)$ denotes the ex-ante expected payoff of plan p, so he will choose *not* to observe the signal at the very beginning. This leads to the seemingly paradoxical phenomenon of **information avoidance**. To explain in words, when the individual anticipates that he will handle information in a highly biased manner, he may find it optimal to reject the information in the first place. In this example, information avoidance can be viewed as the individual's commitment to "debias" himself.

8 Extensions

Some aspects of the model can be relaxed to incorporate more realistic concerns. For example, in the biased updating, individuals are trying to justify some *fixed* bias throughout the learning process. It would also be interesting to consider a **belief-dependent biased updating rule**, in which individuals can modify their bias to tailor their current beliefs. Below is an example.

[24] Recall that the initial signal interpretation is $\alpha_0 = 4$, so the individual will choose h after signal a and l after signal b.

Example 15. [Belief-Dependent Bias I] Suppose that $\Theta = \{G, B\}$. Consider an individual who is initially biased toward state θ. His bias process $\{b_t^\theta\}$ evolves according to the following rule

$$b_t^\theta = \begin{cases} G & \text{if } \mu_t(G) \geq 1/2 \\ B & \text{if } \mu_t(G) < 1/2 \end{cases}.$$

In other words, if state G is most likely, individuals will become biased toward G; if state B is most likely, individuals will become biased toward B. Given the time-t bias b_t^θ, the next-period belief μ_{t+1}^θ is updated according to the model that can best support the current bias, b_t^θ. This example in a similar spirit to the model of confirmatory bias in Rabin and Schrag (1999), where individuals are biased toward the most likely state according to their current beliefs.[25]

Remark 5. In Appendix C, I present a more general framework that accommodates multiple states and a large class of bias-evolving rules. It is worth mentioning that the belief-dependent biased rule is *not* identical to the maximum likelihood rule. Even though the bias is updated overtime, individuals are still biased at every period, so information is always processed in a non-objective manner.

Another interesting extension is that individuals also care about being correct when justifying their bias. In the benchmark model, individuals keep all models without narrowing them down, so it would be interesting to consider a setup where individuals can also discard some models during the learning. One possibility is to consider the **ρ-maximum likelihood biased updating rule**, where individuals apply the biased rule to a subset of models that pass some likelihood test as follows.

Example 16. [ρ-maximum likelihood biased updating rule] For all $\rho \in [0, 1]$, denote by

$$\mathcal{A}_\rho^t \equiv \left\{ \alpha \in \mathcal{A} : \mu(s_1, s_2, ..., s_t|\alpha) \geq (1-\rho) \times \max_{\alpha' \in \mathcal{A}} \mu(s_1, s_2, ..., s_t|\alpha') \right\}.$$

At time t, individuals only consider the models in $\mathcal{A}_\rho^t$ and apply the biased rule with these models to obtain the next-period belief. The ρ-maximum likelihood biased rule can be viewed as a combination of the biased updating rule and the ρ-maximum likelihood rule in Epstein and Schneidler (2007). Specifically, when $\rho = 0$, it refines the maximum updating rule; when $\rho = 1$, it corresponds to the benchmark biased rule. To characterize limit beliefs, we only need to replace $\mathcal{A}$ with $\mathcal{A}_\rho^\infty$ in the statement of relevant theorems, where $\mathcal{A}_\rho^\infty$ represents the set of models that will survive the likelihood test in the limit.

Some other extensions are worth pursuing. For example, we can consider a richer learning environment. This paper only deals with a simple learning environment where individuals receive

[25] In Rabin and Schrag (1999), if an agent received a signal that supports his current belief, he would correctly read the signal; if he received a signal that contradicts his current belief, he would misread the signal with a strictly positive probability.

exogenous i.i.d. signals. A natural extension is to introduce endogenous signals which could also depend on actions taken by individuals. It is conceivable that similar characterizations should still hold in the limit.

9 Concluding Remarks

This paper develops a framework to study biased learning and provides a comprehensive discussion of limit beliefs under this rule. The paper highlights the fact that informational ambiguity can contaminate learning and lead to overconfidence when individuals have bias-confirming incentives. This result mainly comes from two forces. First, ambiguity accommodates multiple interpretations of signals, so individuals have the flexibility to distort evidence toward their favorite directions. Second, ambiguous models have complementary effects, with which individuals can exploit favorable news and hedge against bad news, so biased individuals can become more confident than any Bayesian individual. Due to these forces, when there is sufficient ambiguity, individuals may end up learning incorrectly and more confidently. Some topics are not covered in this paper and are worth pursuing. First, it would be of interest to provide an axiomatic foundation for the biased rule suggested by this paper. Second, this paper focuses on the case where individuals are only driven by bias-justifying incentives, so it would be more realistic to consider other constraints (e.g., preferences for accuracy, consistency). Third, this paper only studies a passive learning problem and it would be interesting to investigate a problem where signals are endogenously driven. Lastly, the learning processes of individuals are independent of each other in this paper, so allowing for dependence across individuals seems a natural next-step (e.g., strategic interactions, social learning).

References

[1] Ackert, L. F., & Athanassakos, G. (1997). Prior uncertainty, analyst bias, and subsequent abnormal returns. *Journal of Financial Research*, 20(2), 263-273.

[2] Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The Self in Social Judgment*, 1, 85-106.

[3] Athanassakos, G., & Kalimipalli, M. (2003). Analyst forecast dispersion and future stock return volatility. *Quarterly Journal of Business and Economics*, 57-78.

[4] Babcock, L., & Loewenstein, G. (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives*, 11(1), 109-126.

[5] Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. The American Economic Review, 85(5), 1337-1343.

[6] Babcock, L., Wang, X., & Loewenstein, G. (1996). Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias. *The Quarterly Journal of Economics*, 111(1), 1-19.

[7] Barron, K. (2020). Belief updating: does the 'good-news, bad-news' asymmetry extend to purely financial domains?. *Experimental Economics*, 1-28.

[8] Battigalli, P., Catonini, E., Lanzani, G., & Marinacci, M. (2019). Ambiguity attitudes and self-confirming equilibrium in sequential games. *Games and Economic Behavior*, 115, 1-29.

[9] Battigalli, P., Cerreia-Vioglio, S., Maccheroni, F., & Marinacci, M. (2015). Self-confirming equilibrium and model uncertainty. *American Economic Review*, 105(2), 646-77.

[10] Battigalli, P., Francetich, A., Lanzani, G., & Marinacci, M. (2019). Learning and self-confirming long-run biases. *Journal of Economic Theory*, 183, 740-785.

[11] Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3), 871-915.

[12] Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics*, 51-58.

[13] Blume, L. E., & Easley, D. (1982). Learning to be rational. *Journal of Economic Theory*, 26(2), 340-351

[14] Blume, L. E., & Easley, D. (1998). Rational expectations and rational learning. *Organizations with Incomplete Information: Essays in Economic Analysis: A Tribute to Roy Radner*, 61-109.

[15] Bohren, J. A. (2016). Informational herding with model misspecification. *Journal of Economic Theory*, 163, 222-247.

[16] Bohren, J. A., & Hauser, D. N. (2019). Social Learning with Model Misspecification: A Framework and a Characterization.

[17] Brous, P. A., & Kini, O. (1993). A reexamination of analysts' earnings forecasts for takeover targets. *Journal of Financial Economics*, 33(2), 201-225.

[18] Brunnermeier, M. K., & Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4), 1092-1118.

[19] Bunke, O., & Milhaud, X. (1998). Asymptotic behavior of Bayes estimates under possibly incorrect models. *The Annals of Statistics*, 26(2), 617-644.

[20] Butler, K. C., & Lang, L. H. (1991). The forecast accuracy of individual analysts: Evidence of systematic optimism and pessimism. *Journal of Accounting Research*, 150-156.

[21] Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1), 306-318.

[22] Caplin, A., & Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1), 55-79.

[23] Cassar, L., & Klein, A. H. (2019). A matter of perspective: How failure shapes distributive preferences. *Management Science*, 65(11), 5050-5064.

[24] Chang, Y. C., & Cheng, H. W. (2015). Information environment and investor behavior. *Journal of Banking & Finance*, 59, 250-264.

[25] Chang, J. W., & Choi, H. M. (2017). Analyst optimism and incentives under market uncertainty. *Financial Review*, 52(3), 307-345.

[26] Chen, J. Y. (2019). Sequential Learning under Informational Ambiguity. *Available at SSRN 3480231.*

[27] Coutts, A. (2019). Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics*, 22(2), 369-395.

[28] Das, S., Levine, C. B., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecasts. *Accounting Review*, 277-294.

[29] Deffains, B., Espinosa, R., & Thöni, C. (2016). Political self-serving bias and redistribution. *Journal of Public Economics*, 134, 67-74.

[30] Easterwood, J. C., & Nutt, S. R. (1999). Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism?. *The Journal of Finance*, 54(5), 1777-1797.

[31] Eil, D., & Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2), 114-38.

[32] Epstein, L. G., & Schneider, M. (2007). Learning under ambiguity. *The Review of Economic Studies*, 74(4), 1275-1303.

[33] Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization*, 80(3), 532-545.

[34] Esponda, I., & Pouzo, D. (2016). Berk–Nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3), 1093-1130.

[35] Frick, M., Iijima, R., & Ishii, Y. (2019). Misinterpreting others and the fragility of social learning.

[36] Frick, M., Iijima, R., & Ishii, Y. (2020). Stability and Robustness in Misspecified Learning Models.

[37] Fryer Jr, R. G., Harms, P., & Jackson, M. O. (2019). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. *Journal of the European Economic Association*, 17(5), 1470-1501.

[38] Fudenberg, D., Lanzani, G., & Strack, P. (2020). Limits Points of Endogenous Misspecified Learning. Available at SSRN.

[39] Fudenberg, D., Romanyuk, G., & Strack, P. (2017). Active learning with a misspecified prior. *Theoretical Economics*, 12(3), 1155-1189.

[40] Galperti, S. (2019). Persuasion: The art of changing worldviews. *American Economic Review*, 109(3), 996-1031.

[41] Gilboa, I., & Marinacci, M. (2016). Ambiguity and the Bayesian paradigm. In *Readings in Formal Epistemology* (pp. 385-439). Springer, Cham.

[42] Graham, J. R., Harvey, C. R., & Huang, H. (2009). Investor competence, trading frequency, and home bias. *Management Science*, 55(7), 1094-1106.

[43] Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *The Journal of Finance*, 64(2), 549-578.

[44] Haisley, E. C., & Weber, R. A. (2010). Self-serving interpretations of ambiguity in other-regarding behavior. *Games and Economic Behavior*, 68(2), 614-625.

[45] Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, 4(1), 5-28.

[46] Heidhues, P., Kőszegi, B., & Strack, P. (2018). Unrealistic expectations and misguided learning. *Econometrica*, 86(4), 1159-1214.

[47] Karelaia, N., & Hogarth, R. M. (2010). The attraction of uncertainty: Interactions between skill and levels of uncertainty in market-entry games. *Journal of Risk and Uncertainty*, 41(2), 141-166.

[48] Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4), 673-707.

[49] Loewenstein, G., & Adler, D. (1995). A bias in the prediction of tastes. *The Economic Journal*, 105(431), 929-937.

[50] Machina, M. J., & Siniscalchi, M. (2014). Ambiguity and ambiguity aversion. In *Handbook of the Economics of Risk and Uncertainty* (Vol. 1, pp. 729-807). North-Holland.

[51] Marinacci, M. (2002). Learning from ambiguous urns. *Statistical Papers*, 43(1), 143.

[52] Marinacci, M. (2015). Model uncertainty. *Journal of the European Economic Association*, 13(6), 1022-1100.

[53] Marinacci, M., & Massari, F. (2019). Learning from ambiguous and misspecified models. *Journal of Mathematical Economics*, 84, 144-149.

[54] Mobius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2011). Managing self-confidence: Theory and experimental evidence (No. w17014). National Bureau of Economic Research.

[55] Nyarko, Y. (1991). Learning in mis-specified models and the possibility of cycles. *Journal of Economic Theory*, 55(2), 416-427.

[56] Ortoleva, P. (2012). Modeling the change of paradigm: Non-Bayesian reactions to unexpected news. *American Economic Review*, 102(6), 2410-36.

[57] Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369-381.

[58] Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1), 37-82.

[59] Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics*, 6(3), 379-421.

[60] Strotz, R. H. (1955-1956). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165-180.

[61] Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755-769.

[62] Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological bulletin*, 103(2), 193.

A Proofs

A.1 Some Auxiliary Lemmas

I first define several concepts, some of which are also seen in Berk (1966)'s proof. Let v be an finite measure on $(\Theta, \mathfrak{B}_\Theta)$ that admits a density function $f_v \equiv \frac{dv}{dm}$ which is continuous on its support, $\text{supp}(v)$.[26] The support of v refers to the closure of points f_v taking strictly positive values on. For all $\mathcal{U} \subset \Theta$, all continuous function $G(\alpha, \theta) : \mathcal{A} \times \Theta \to \mathbb{R}$, and all $t \in \{1, 2, ...\} \cup \{\infty\}$, I define

$$\mathcal{U}||G(\alpha,\theta)||_t^{\theta,v} \equiv \left[\int_{\mathcal{U}} |G(\alpha,\theta)|^t dv(\theta)\right]^{1/t} \text{ and } \mathcal{U}||G(\alpha,\theta)||_\infty^{\theta,v} \equiv \sup_{\theta \in \mathcal{U} \cap \text{supp}(v)} |G(\alpha,\theta)|.$$

Further define

$$H_t(\alpha,\theta) \equiv \frac{1}{t}\sum_{i=1}^{t} \log f(s_i|\alpha,\theta) \text{ and } H(\alpha,\theta) \equiv \mathbb{E}^* \log f(s_i|\alpha,\theta).$$

We have the following lemmas.

Lemma 4. [Continuity] *(1) $H(\alpha,\theta) : \mathcal{A} \times \Theta \to \mathbb{R}$ is continuous (2) $\mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v} : \mathcal{A} \to \mathbb{R}_{++}$ is continuous for all compact $\mathcal{U} \subset \Theta$.*

Proof. (1) is a direct result from Assumption 1 and 2 combined with the dominated convergence theorem. The continuity of $\mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v}$ follows from (1) and Berge's maximum theorem. □

Lemma 5. [Uniform Convergence] *For all compact $\mathcal{U} \subset \Theta$,*

$$\max_{\alpha\in\mathcal{A}} \left|\mathcal{U}||\exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right| \to 0$$

$\mathbb{P}^*$-almost surely.

Proof. On one hand,

$$\begin{aligned}\max_{\alpha\in\mathcal{A}}\left(\mathcal{U}||\exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right) &\le \max_{\alpha\in\mathcal{A}}\left(\mathcal{U}||\exp H_t(\alpha,\theta)||_\infty^{\theta,v} - \mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right)\\ &\le \max_{\alpha\in\mathcal{A}}\left|\mathcal{U}||\exp H_t(\alpha,\theta)||_\infty^{\theta,v} - \mathcal{U}||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right|\\ &\le \max_{\alpha\in\mathcal{A}} \mathcal{U}||\exp H_t(\alpha,\theta) - \exp H(\alpha,\theta)||_\infty^{\theta,v}\end{aligned}$$

where the first inequality is implied by Holder's inequality, and the last inequality is implied by the triangle inequality. From the uniform law of large numbers (ULLN), we have:

$$\max_{(\alpha,\theta)\in\mathcal{A}\times\mathcal{U}\cap\text{supp}(v)} |H_t(\alpha,\theta) - H(\alpha,\theta)| \to 0 \quad \mathbb{P}^* - a.s.,$$

[26]In the main text, I also use v to denote the benchmark measure on the signal space $(S, \mathcal{B}_S)$, but it is easy to infer from the context which one I am referring to.

which further implies that

$$\max_{\alpha\in\mathcal{A}} \mathcal{U}|| \exp H_t(\alpha,\theta) - \exp H(\alpha,\theta)||_\infty^{\theta,v} = \max_{(\alpha,\theta)\in\mathcal{A}\times\mathcal{U}\cap\text{supp}(v)} |\exp H_t(\alpha,\theta) - \exp H(\alpha,\theta)| \to 0 \quad \mathbb{P}^*\text{−}a.s., \tag{3}$$

so we have

$$\limsup \max_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \le 0 \quad \mathbb{P}^* - a.s.. \tag{4}$$

On the other hand, we have

$$\begin{aligned}
&\min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \\
\ge &\min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} \right) + \min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \\
\ge &- \max_{\alpha\in\mathcal{A}} \left| \mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} \right| + \min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \\
\ge &- \max_{\alpha\in\mathcal{A}} \mathcal{U}|| \exp H_t(\alpha,\theta) - \exp H(\alpha,\theta)||_t^{\theta,v} + \min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \\
\ge &- \underbrace{\max_{\alpha\in\mathcal{A}} \mathcal{U}|| \exp H_t(\alpha,\theta) - \exp H(\alpha,\theta)||_\infty^{\theta,v}}_{(a)} + \underbrace{\min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right)}_{(b)}.
\end{aligned}$$

I next show that both (a) and (b) converge to 0 almost surely. The convergence of (a) follows from (3). Notice that

$$\begin{aligned}
\mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v} &= \left[\int_{\mathcal{U}\cap\text{supp}(v)} |\exp H(\alpha,\theta)|^t dv(\theta) \right]^{1/t} \\
&\to \sup_{\theta\in\mathcal{U}\cap\text{supp}(v)} |\exp H(\alpha,\theta)| = \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v}.
\end{aligned}$$

Since $\exp H(\alpha,\theta)$ is continuous, $h_t(\alpha) \equiv \mathcal{U}|| \exp H(\alpha,\theta)||_t^{\theta,v}$ is also continuous in α. We know that $\{h_t(\alpha)\}$ converges uniformly to $h_\infty(\alpha)$ from Dini's theorem (since $\{h_t(\alpha)\}$ is an increasing function sequence), so (b) also converges to 0. Therefore,

$$\liminf \min_{\alpha\in\mathcal{A}} \left(\mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right) \ge 0 \quad \mathbb{P}^* - a.s.. \tag{5}$$

Combining (4) and (5), we have

$$\max_{\alpha\in\mathcal{A}} \left| \mathcal{U}|| \exp H_t(\alpha,\theta)||_t^{\theta,v} - \mathcal{U}|| \exp H(\alpha,\theta)||_\infty^{\theta,v} \right| \to 0 \quad \mathbb{P}^* - a.s..$$

□

Corollary 3. [Composite Uniform Convergence] *For all continuous function* $\omega : \mathbb{R}_{++} \to \mathbb{R}$ *and all* *compact* $\mathcal{U} \subset \Theta$*, we must have*

$$\max_{\alpha \in \mathcal{A}} \left| \omega \left(\mathcal{U} || \exp H_t (\alpha, \theta) ||_t^{\theta, v} \right) - \omega \left(\mathcal{U} || \exp H (\alpha, \theta) ||_\infty^{\theta, v} \right) \right| \to 0 \quad \mathbb{P}^* - a.s.$$

Proof. Denote set $C \subset \mathbb{R}$ as the range of function $\mathcal{U} || \exp H (\alpha, \theta) ||_\infty^{\theta, v}$ (as a function of α defined on $\mathcal{A}$). Due to the fact that $\mathcal{U} || \exp H (\alpha, \theta) ||_\infty^{\theta, v}$ is continuous and $\mathcal{A}$ is compact, C must be a compact set as well. Further denote $C^\varepsilon = \{x \in \mathbb{R} : \min_{y \in C} |x - y| \leq \varepsilon\}$, which is the set of points within distance ε to set C, so C^ε is also compact. Since ω is continuous and C^ε is compact, ω is a uniformly continuous function on C^ε. As a result, for all $\xi > 0$, there exists some $\delta > 0$ such that $|\omega (x) - \omega (y) | < \xi$ whenever $|x - y| < \delta$ and $x, y \in C^\varepsilon$. From Lemma 5, for almost all signal paths and for all α, we make the distance between $\mathcal{U} || \exp H_t (\alpha, \theta) ||_t^{\theta, v}$ and $\mathcal{U} || \exp H (\alpha, \theta) ||_\infty^{\theta, v}$ sufficiently small by making t sufficiently large. As such, for all $\xi > 0$, the difference between $\omega \left(\mathcal{U} || \exp H_t (\alpha, \theta) ||_t^{\theta, v} \right)$ and $\omega \left(\mathcal{U} || \exp H_t (\alpha, \theta) ||_\infty^{\theta, v} \right)$ is a.s. uniformly bounded by ξ for sufficiently large t, which establishes the corollary. □

A.2 Proof of Lemma 1

The idea of the proof resembles Berk (1966), but the main difference is to prove that beliefs converge *uniformly* for all models, which can be established from the previous lemmas.

Proof. For all $x > 0$, I define a sequence of sets $\{\mathcal{U}_\mathcal{A}^x\}$ where:

$$\mathcal{U}_\mathcal{A}^x \equiv \left\{ \theta \in \Theta : \min_{\alpha \in \mathcal{A}} r (\alpha, \theta) \leq \frac{1}{x} \right\}. \tag{6}$$

It is easy to verify that $\{\mathcal{U}_\mathcal{A}^x\}$ is decreasing in x (with respect to set inclusion) with the limit being $\mathcal{U}_\mathcal{A}$. It is easy to verify that $\mathcal{U}_\mathcal{A}^x$ is a closed subset of Θ, hence compact. For all open set U containing $\mathcal{U}_\mathcal{A}$, there must exist some $x > 0$ such that $\mathcal{U}_\mathcal{A}^x \subset U$.[27] If $\mathcal{U}_\mathcal{A}^x = \Theta$, the claim is trivially correct, so the rest of the proof focuses on the case where $\mathcal{U}_\mathcal{A}^x \subsetneqq \Theta$. I also define a set $\mathcal{U}_\mathcal{A}^{-x}$ by flipping the direction of inequality in (6), so $\mathcal{U}_\mathcal{A}^{-x}$ is also a compact set.

With some abuse of notation, I use μ to denote the measure corresponding to the density

[27] From Berk (1965), suppose not, $\mathcal{U}_\mathcal{A}^x \cap U^c$ is a nested system of closed non-empty set, which must have non-empty intersection, i.e., $\cap \mathcal{U}_\mathcal{A}^x \cap U^c = \mathcal{U}_\mathcal{A} \cap U^c \neq \emptyset$, which contradicts $U \supset \mathcal{U}_\mathcal{A}$.

function μ. For all model $\alpha \in \mathcal{A}$, we have:

$$
\begin{aligned}
\min_{\alpha\in\mathcal{A}} \frac{1}{t}\log\left[\frac{\mu_t\left(U|\alpha\right)}{\mu_t\left(U^c|\alpha\right)}\right] &\geq \min_{\alpha\in\mathcal{A}} \frac{1}{t}\log\left[\frac{\mu_t\left(\mathcal{U}_{\mathcal{A}}^x|\alpha\right)}{\mu_t\left(\mathcal{U}_{\mathcal{A}}^{-x}|\alpha\right)}\right] \\
&= \min_{\alpha\in\mathcal{A}} \frac{1}{t}\log\left[\frac{\int_{\mathcal{U}_{\mathcal{A}}^x}\prod_{i=1}^t f\left(s_i|\alpha,\theta\right)\mu\left(\theta\right)dm\left(\theta\right)}{\int_{\mathcal{U}_{\mathcal{A}}^{-x}}\prod_{i=1}^t f\left(s_i|\alpha,\theta\right)\mu\left(\theta\right)dm\left(\theta\right)}\right] \\
&= \min_{\alpha\in\mathcal{A}} \log\frac{\left(\int_{\mathcal{U}_{\mathcal{A}}^x}\left(\exp\left(\frac{1}{t}\sum_{i=1}^t\log f\left(s_i|\alpha,\theta\right)\right)\right)^t\mu\left(\theta\right)dm\left(\theta\right)\right)^{1/t}}{\left(\int_{\mathcal{U}_{\mathcal{A}}^{-x}}\left(\exp\left(\frac{1}{t}\sum_{i=1}^t\log f\left(s_i|\alpha,\theta\right)\right)\right)^t\mu\left(\theta\right)dm\left(\theta\right)\right)^{1/t}} \\
&= \min_{\alpha\in\mathcal{A}} \log\left[\frac{\mathcal{U}_{\mathcal{A}}^x||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}}{\mathcal{U}_{\mathcal{A}}^{-x}||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}}\right] \\
&\geq \underbrace{\min_{\alpha\in\mathcal{A}}\left(\log\left[\frac{\mathcal{U}_{\mathcal{A}}^x||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}}{\mathcal{U}_{\mathcal{A}}^{-x}||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}}\right] - \log\left[\frac{\mathcal{U}_{\mathcal{A}}^x||\exp H\left(\alpha,\theta\right)||_\infty^{\theta,\mu}}{\mathcal{U}_{\mathcal{A}}^{-x}||\exp H\left(\alpha,\theta\right)||_\infty^{\theta,\mu}}\right]\right)}_{(a)} \\
&+ \min_{\alpha\in\mathcal{A}} \log\left[\frac{\mathcal{U}_{\mathcal{A}}^x||\exp H\left(\alpha,\theta\right)||_\infty^{\theta,\mu}}{\mathcal{U}_{\mathcal{A}}^{-x}||\exp H\left(\alpha,\theta\right)||_\infty^{\theta,\mu}}\right]
\end{aligned}
$$

Rearranging the terms of (a), we get

$$
\begin{aligned}
(a) \geq &\min_{\alpha\in\mathcal{A}}\left[\log\left(\mathcal{U}_{\mathcal{A}}^x||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}\right) - \log\left(\mathcal{U}_{\mathcal{A}}^x||\exp H\left(\alpha,\theta\right)||_\infty^{\theta,\mu}\right)\right] \\
&+ \min_{\alpha\in\mathcal{A}}\left[\log\left(\mathcal{U}_{\mathcal{A}}^{-x}||\exp H_t\left(\alpha,\theta\right)||_\infty^{\theta,\mu}\right) - \log\left(\mathcal{U}_{\mathcal{A}}^{-x}||\exp H\left(\alpha,\theta\right)||_t^{\theta,\mu}\right)\right]
\end{aligned} \tag{7}
$$

Corollary 3 implies that

$$
\min_{\alpha\in\mathcal{A}}\left[\log\left(\mathcal{U}||\exp H_t\left(\alpha,\theta\right)||_t^{\theta,\mu}\right) - \log\left(\mathcal{U}||\exp H_t\left(\alpha,\theta\right)||_\infty^{\theta,\mu}\right)\right] \to 0 \quad \mathbb{P}^* - a.s.,
$$

for all compact set $\mathcal{U} \subset \Theta$. Recall that $\mathcal{U}_{\mathcal{A}}^x$ and $\mathcal{U}_{\mathcal{A}}^{-x}$ are both compact, so the RHS of (7) must converge to 0 almost surely, which implies that the lim inf of (a) is non-negative, so we have

$$
\begin{aligned}
\liminf\left(\min_{\alpha\in\mathcal{A}}\frac{1}{t}\log\left[\frac{\mu_t\left(U|\alpha\right)}{\mu_t\left(U^c|\alpha\right)}\right]\right) &\geq \min_{\alpha\in\mathcal{A}}\log\left[\frac{\mathcal{U}^x_{\mathcal{A}}||\exp H\left(\alpha,\theta\right)||^{\theta,\mu}_{\infty}}{\mathcal{U}^{-x}_{\mathcal{A}}||\exp H\left(\alpha,\theta\right)||^{\theta,\mu}_{\infty}}\right] \quad \mathbb{P}^*-a.s. \\
&= \min_{\alpha\in\mathcal{A}}\left(\log\left(\mathcal{U}^x_{\mathcal{A}}||\exp H\left(\alpha,\theta\right)||^{\theta,\mu}_{\infty}\right)-\log\left(\mathcal{U}^{-x}_{\mathcal{A}}||\exp H\left(\alpha,\theta\right)||^{\theta,\mu}_{\infty}\right)\right) \\
&= \min_{\alpha\in\mathcal{A}}\left(\log\left(\max_{\theta\in\mathcal{U}^x_{\mathcal{A}}}\exp\mathbb{E}^*\log f\left(s|\alpha,\theta\right)\right)-\log\left(\max_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}}\exp\mathbb{E}^*\log f\left(s|\alpha,\theta\right)\right)\right) \\
&= \min_{\alpha\in\mathcal{A}}\left(\max_{\theta\in\mathcal{U}^x_{\mathcal{A}}}\mathbb{E}^*\log f\left(s|\alpha,\theta\right)-\max_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}}\mathbb{E}^*\log f\left(s|\alpha,\theta\right)\right) \\
&= \min_{\alpha\in\mathcal{A}}\left(\min_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}}\mathbb{E}^*\log\frac{f\left(s|\alpha^*,\theta^*\right)}{f\left(s|\alpha,\theta\right)}-\max_{\theta\in\mathcal{U}^x_{\mathcal{A}}}\mathbb{E}^*\log\frac{f\left(s|\alpha^*,\theta^*\right)}{f\left(s|\alpha,\theta\right)}\right) \\
&= \min_{\alpha\in\mathcal{A}}\left(\min_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}}\mathcal{R}\left(\alpha,\theta\right)-\min_{\theta\in\mathcal{U}^x_{\mathcal{A}}}\mathcal{R}\left(\alpha,\theta\right)\right) \\
&= \min_{\alpha\in\mathcal{A}}\left(\min_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}}r\left(\alpha,\theta\right)-\min_{\theta\in\mathcal{U}^x_{\mathcal{A}}}r\left(\alpha,\theta\right)\right)\geq\frac{1}{x}>0,
\end{aligned}
$$

where the last weak inequality comes from the facts that: (i) $\min_{\theta\in\mathcal{U}^{-x}_{\mathcal{A}}} r\left(\alpha,\theta\right) \geq 1/x$ from the definition of $\mathcal{U}^{-x}_{\mathcal{A}}$, and (ii) $r\left(\alpha,\theta\right)\geq 0$ from the definition of r. As a result, $\min_{\alpha\in\mathcal{A}}\mu_t\left(U|\alpha\right)\to 1$ $\mathbb{P}^*$-almost surely. For all possible bias τ, we must have $\mu^\tau_t\left(U|\alpha\right)\geq\min_{\alpha\in\mathcal{A}}\mu_t\left(U|\alpha\right)$, so the result is proved. □

A.3 Proof of Lemma 2 and Theorem 1

Proof. Denote by $U_t\left(\theta\right)$ the time-t utility of an individual with bias θ. We have

$$
U_t\left(\theta\right)=\max_{\alpha\in\mathcal{A}}\mu_t\left(\theta|\alpha\right)=\max_{\alpha\in\mathcal{A}}\frac{\prod_{i=1}^t f\left(s_i|\alpha,\theta\right)\mu\left(\theta\right)}{\int\prod_{i=1}^t f\left(s_i|\alpha,\theta'\right)\mu\left(\theta'\right)dm\left(\theta'\right)},
$$

or equivalently,

$$
\begin{aligned}
\frac{1}{t}\log U_t\left(\theta\right) &= \max_{\alpha\in\mathcal{A}}\left(\frac{1}{t}\log\left[\prod_{i=1}^t f\left(s_i|\alpha,\theta\right)\mu\left(\theta\right)\right]-\frac{1}{t}\log\left[\int_\Theta\prod_{i=1}^t f\left(s_i|\alpha,\theta'\right)\mu\left(\theta'\right)dm\left(\theta'\right)\right]\right) \\
&= \max_{\alpha\in\mathcal{A}}\left[H_t\left(\alpha,\theta\right)-\log\left(\Theta||\exp H_t\left(\alpha,\theta\right)||^{\theta,\mu}_t\right)\right]+\frac{1}{t}\log\mu\left(\theta\right) \qquad (8)
\end{aligned}
$$

Uniform law of large number (ULLN) implies that

$$
\max_{\alpha\in\mathcal{A}}\left|H_t\left(\alpha,\theta\right)-H\left(\alpha,\theta\right)\right|\to 0 \quad \mathbb{P}^*-a.s.. \qquad (9)
$$

Corollary 3 implies that

$$\max_{\alpha\in\mathcal{A}} \left| \log\left(\Theta || \exp H_t(\alpha,\theta) ||_t^{\theta,\mu}\right) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right) \right| \to 0 \quad \mathbb{P}^* - a.s.. \tag{10}$$

From (9) and (10), we can find a set E with $\mathbb{P}^*(E) = 1$ such that for all signal path $s^\infty \in E$, for all $\varepsilon > 0$, there exists some T such that for all $t \geq T$, we have

$$\begin{aligned} \forall \alpha \in \mathcal{A}: \quad & |H_t(\alpha,\theta) - H(\alpha,\theta)| \leq \varepsilon/2 \\ & \left|\log\left(\Theta || \exp H_t(\alpha,\theta) ||_t^{\theta,\mu}\right) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right)\right| \leq \varepsilon/2, \end{aligned}$$

which implies that for all $t \geq T$,

$$\begin{aligned} & \max_{\alpha\in\mathcal{A}} \left[H_t(\alpha,\theta) - \log\left(\Theta || \exp H_t(\alpha,\theta) ||_t^{\theta,\mu}\right)\right] \\ = & \max_{\alpha\in\mathcal{A}} \left[H(\alpha,\theta) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right) + H_t(\alpha,\theta) - H(\alpha,\theta) + \log\left(\Theta || \exp H_t(\alpha,\theta) ||_\infty^{\theta,\mu}\right) - \log\left(\Theta || \exp H_t(\alpha,\theta) ||_t^{\theta,\mu}\right)\right] \\ \in & \left[\max_{\alpha\in\mathcal{A}} \left[H(\alpha,\theta) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right)\right] - \varepsilon, \max_{\alpha\in\mathcal{A}} \left[H(\alpha,\theta) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right)\right] + \varepsilon\right]. \end{aligned}$$

Therefore,

$$\max_{\alpha\in\mathcal{A}} \left[H_t(\alpha,\theta) - \log\left(\Theta || \exp H_t(\alpha,\theta) ||_t^{\theta,\mu}\right)\right] \to \max_{\alpha\in\mathcal{A}} \left(H(\alpha,\theta) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right)\right) \quad \mathbb{P}^* - a.s..$$

Consequently, we get

$$\begin{aligned} \frac{1}{t} \log U_t(\theta) \to & \max_{\alpha\in\mathcal{A}} \left(H(\alpha,\theta) - \log\left(\Theta || \exp H(\alpha,\theta) ||_\infty^{\theta,\mu}\right)\right) \quad \mathbb{P}^* - a.s. \\ & = \max_{\alpha\in\mathcal{A}} \left(\mathbb{E}^* \log f(s|\alpha,\theta) - \log\left(\max_{\theta\in\Theta} \exp \mathbb{E}^* \log f(s|\alpha,\theta)\right)\right) \\ & = \max_{\alpha\in\mathcal{A}} \left(\mathbb{E}^* \log f(s|\alpha,\theta) - \max_{\theta\in\Theta} \mathbb{E}^* \log f(s|\alpha,\theta)\right) \\ & = \max_{\alpha\in\mathcal{A}} \left(-\mathbb{E}^* \log \frac{f(s|\alpha^*,\theta^*)}{f(s|\alpha,\theta)} + \min_{\theta\in\Theta} \mathbb{E}^* \log \frac{f(s|\alpha^*,\theta^*)}{f(s|\alpha,\theta)}\right) \\ & = -\min_{\alpha\in\mathcal{A}} r(\alpha,\theta) \end{aligned}$$

Since α_t^θ maximizes $U_t(\theta)$ for each t (hence maximizes $\frac{1}{t}\log U_t(\theta)$), it follows immediately that

$$\mathcal{A}_\infty^\theta \subset \min_{\alpha\in\mathcal{A}} r(\alpha,\theta) \coloneqq \mathcal{A}^\theta \quad \mathbb{P}^* - a.s.,$$

so individuals will only adopt models that minimize the information potential of state θ. Since only models in $\mathcal{A}^\theta$ will be adopted in the limit, the learning problem is essentially the same as if bias-θ individuals only perceive the model set $\mathcal{A}^\theta$. Using Lemma 1, it is easy to verify that limit beliefs of bias-θ individuals will settle on zero-potential states under $\mathcal{A}^\theta$, which gives Theorem 1. □

A.4 Proof of Proposition 1

Proof. Define $\mathcal{A}(\delta) \equiv \{\alpha \in \mathbb{A} : ||\alpha - \alpha^*|| \leq \delta\}$, where $||.||$ denotes the relevant metric in $\mathbb{A}$. Further define

$$\mathcal{U}(\delta) \equiv \left\{\theta \in \Theta : \min_{\alpha \in \mathcal{A}(\delta)} r(\alpha, \theta) = 0\right\}$$

which is the set of zero-potential states when the model set is $\mathcal{A}(\delta)$. From Berge's maximum theorem, we know that $V(\theta, \delta) = \min_{\alpha \in \mathcal{A}(\delta)} r(\alpha, \theta)$ is a continuous function, so $\mathcal{U}(\delta)$ is a compact-valued upper semi-continuous correspondence. Further denote

$$D(\delta) = \max_{\theta \in \mathcal{U}(\delta)} ||\theta - \theta^*||,$$

which describes the "size" of $\mathcal{U}(\delta)$. Applying Berge's theorem again, we know that $D(\delta)$ is also upper semi-continuous functions of δ. Notice that when $\delta = 0$, we have $\mathcal{U}(\delta) = \{\theta^*\}$, so $D(0) = 0$. Since $D(\delta) \geq 0$, the upper semi-continuity of D implies that $D(\delta) \to 0$ as $\delta \to 0$. Note that for all compact set $\mathcal{A}$, we can always bound it using some $\mathcal{A}(\delta)$. Therefore, we must have:

$$\max_{\theta \in \mathcal{U}_{\mathcal{A}}} ||\theta - \theta^*|| \to 0 \quad \text{as } d(\mathcal{A}) \to 0 \tag{11}$$

Further denote $\mathcal{U}^{\varepsilon}_{\mathcal{A}} = \{\theta \in \Theta : \min_{\theta' \in \mathcal{U}_{\mathcal{A}}} ||\theta - \theta'|| \leq \varepsilon\}$, which is the set of states that are within distance ε to $\mathcal{U}_{\mathcal{A}}$. For all bounded and continuous function $h : \Theta \to \mathbb{R}$ and for all $\varepsilon > 0$, we have:

$$\begin{aligned}\lim_{t\to\infty} \left|\int_{\Theta} h \times \mu^{\theta}_t dm - \int_{\Theta} h \times \delta_{\theta^*} dm\right| &= \lim_{t\to\infty} \left|\int_{\mathcal{U}^{\varepsilon}_{\mathcal{A}}} h \times \mu^{\theta}_t dm - \int_{\mathcal{U}^{\varepsilon}_{\mathcal{A}}} h \times \delta_{\theta^*} dm\right| \quad \mathbb{P}^* - a.s. \\ &\leq \max_{\theta \in \mathcal{U}^{\varepsilon}_{\mathcal{A}}} |h(\theta) - h(\theta^*)|\end{aligned}$$

where the first equality comes from the fact that $\lim \mu^{\theta}_t (\mathcal{U}^{\varepsilon}_{\mathcal{A}}) = 1$ for all $\varepsilon > 0$ (implied by Lemma 1). From (11), as $d(\mathcal{A}) \to 0$, we have $\mathcal{U}^{\varepsilon}_{\mathcal{A}} \to B_{\varepsilon}(\theta^*)$ under the Hausdorff metric, where $B_{\varepsilon}(\theta^*)$ denotes the ε-closed neighborhood of θ^*. Therefore, for all $\varepsilon > 0$, we have:

$$\lim_{d(\mathcal{A})\to 0} \lim_{t\to\infty} \left|\int_{\Theta} h \times \mu^{\theta}_t dm - \int_{\Theta} h \times \delta_{\theta^*} dm\right| \leq \max_{\theta \in B_{\varepsilon}(\theta^*)} |h(\theta) - h(\theta^*)| \quad \mathbb{P}^* - a.s. \tag{12}$$

Letting $\varepsilon \to 0$, the RHS of (12) converges to 0 from the continuity of h, so the claim is proved. □

A.5 Proof of Proposition 2 and Corollary 1

Proof. Proposition 2 follows immediately from Lemma 1. Let's then prove Corollary 1. For all $\theta \neq \theta^*$, we have

$$\mathcal{R}(\alpha^*, \theta) = \mathbb{E}^* \log \frac{f(s|\theta^*, \alpha^*)}{f(s|\theta, \alpha)} > -\log \mathbb{E}^* \frac{f(s|\theta, \alpha^*)}{f(s|\theta^*, \alpha^*)} = 0,$$

where the strict inequality comes from Jensen's inequality and Assumption 3. Since Θ is finite, there exists some $\delta > 0$ such that $\mathcal{R}(\alpha^*, \theta) > \delta$ for all $\theta \neq \theta^*$ (note that $\mathcal{R}(\alpha^*, \theta^*) = 0$ by definition). From the continuity of $\mathcal{R}(\alpha,\theta)$ (from Lemma 4), when $||\alpha-\alpha^*||$ is sufficiently small, we have $\mathcal{R}(\alpha,\theta)>\delta/2$ and $\mathcal{R}(\alpha,\theta^*)<\delta/2$ for all $\theta\neq\theta^*$. Therefore, when $\mathcal{A}$ is sufficiently small (around the true model), θ^* is the unique zero-potential state for all models in $\mathcal{A}$, which implies that beliefs will converge to δ_{θ^*} for all possible bias. □

A.6 Proof of Proposition 3

Proof. Since θ^* is not locally dominant, for all $\mathcal{A}$ such that $d(\mathcal{A})>0$, there exists some model $\alpha_0\in\mathcal{A}$ such that $r(\alpha_0,\theta_0)=0$ for some $\theta_0\neq\theta^*$. Therefore, we have $\min_{\alpha\in\mathcal{A}} r(\alpha,\theta_0)=0$. From Theorem 1, the limit belief carrier for individuals with bias θ_0 is:

$$\mathcal{U}_{\mathcal{A}}^{\theta_0}=\left\{\theta'\in\Theta : r\left(\alpha',\theta'\right)=0 \text{ where } r\left(\alpha',\theta_0\right)=0\right\}$$

Since θ^* is singular, we have: (i) $\theta^*\notin\mathcal{U}_{\mathcal{A}}^{\theta_0}$, and (ii) $\mathcal{U}_{\mathcal{A}}^{\theta^*}=\{\theta^*\}$. Besides, it is easy to verify that $\mathcal{U}_{\mathcal{A}}^{\theta_0}$ is a closed set from the continuity of $r(\alpha,\theta)$. From the property of metrizable space, we can find two disjoint open sets U_1 and U_2 to separate $\mathcal{U}_{\mathcal{A}}^{\theta_0}$ and $\mathcal{U}_{\mathcal{A}}^{\theta^*}$ in the sense that $U_1\supset\mathcal{U}_{\mathcal{A}}^{\theta_0}$ and $U_2\supset\mathcal{U}_{\mathcal{A}}^{\theta^*}$. From Theorem 1, $\mu_t^{\theta_0}(U_1)\to 1$ and $\mu_t^{\theta^*}(U_1)\to 0$ almost surely, so we have:

$$\lim_{t\to\infty}|\mu_t^{\theta_0}(U_1)-\mu_t^{\theta^*}(U_1)|=1 \quad \mathbb{P}^*-a.s.$$

, which proves the claim. □

A.7 Proof of Lemma 3 and Theorem 2

We prove the benchmark case and the general case separately for better clarity. This is because when Θ is continuous, the benchmark case requires the bias τ to be a Dirac delta function δ_θ, which cannot be directly accommodated in the general case (i.e., we assume that the every bias τ is a well-defined function, but the Dirac delta function is not a function).

Proof. The proof resembles the proof of the single-biased state. For all possible bias $\tau\in\mathcal{T}$,we have

$$U_t(\tau)=\max_{\alpha\in\mathcal{A}} U_t(\tau|\alpha)=\max_{\alpha\in\mathcal{A}}\frac{\int\prod_{i=1}^t f(s_i|\alpha,\theta)\,\mu(\theta)\,\tau(\theta)\,dm(\theta)}{\int\prod_{i=1}^t f(s_i|\alpha,\theta)\,\mu(\theta)\,dm(\theta)}.$$

I define a measure v such that $v(E)=\int_E\mu(\theta)\,\tau(\theta)\,dm(\theta)$ for all measurable set E. In other words, $\mu(\theta)\,\tau(\theta)$ is the density function of measure v with respect to m. Denote by Θ_τ the support of τ, which are the set of biased states. As such,

$$\begin{aligned}\frac{1}{t}\log U_t(\tau)&=\max_{\alpha\in\mathcal{A}}\left(\frac{1}{t}\log\left[\int_{\Theta_\tau}\prod_{i=1}^{t} f(s_i|\alpha,\theta)\,dv(\theta)\right]-\frac{1}{t}\log\left[\int_{\Theta}\prod_{i=1}^{t} f(s_i|\alpha,\theta)\,\mu(\theta)\,dm(\theta)\right]\right)\\&=\max_{\alpha\in\mathcal{A}}\left[\log\left(\Theta_\tau||\exp H_t(\alpha,\theta)\,||_t^{\theta,v}\right)-\log\left(\Theta||\exp H_t(\alpha,\theta)\,||_t^{\theta,\mu}\right)\right].\end{aligned}\tag{13}$$

Corollary 3 implies that

$$\max_{\alpha\in\mathcal{A}} |\log\left(\Theta|| \exp H_t(\alpha,\theta)||_t^{\theta,\mu}\right) - \log\left(\Theta||\exp H(\alpha,\theta)||_\infty^{\theta,\mu}\right)| \to 0 \quad \mathbb{P}^* - a.s., \tag{14}$$

and

$$\max_{\alpha\in\mathcal{A}} |\log\left(\Theta_\tau|| \exp H_t(\alpha,\theta)||_t^{\theta,v}\right) - \log\left(\Theta_\tau||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right)| \to 0 \quad \mathbb{P}^* - a.s.. \tag{15}$$

Following the exact same argument as in the proof of Lemma 2 and Theorem 1, we have

$$\begin{aligned}
\frac{1}{t}\log U_t(\tau) &\to \max_{\alpha\in\mathcal{A}}\left[\log\left(\Theta_\tau||\exp H(\alpha,\theta)||_\infty^{\theta,v}\right) - \log\left(\Theta||\exp H(\alpha,\theta)||_\infty^{\theta,\mu}\right)\right] \quad \mathbb{P}^* - a.s.\\
&= \max_{\alpha\in\mathcal{A}}\left(\log\left(\max_{\theta\in\Theta_\tau}\exp\mathbb{E}^*\log f(s|\alpha,\theta)\right) - \log\left(\max_{\theta\in\Theta}\exp\mathbb{E}^*\log f(s|\alpha,\theta)\right)\right)\\
&= \max_{\alpha\in\mathcal{A}}\left(-\min_{\theta\in\Theta_\tau}\mathbb{E}^*\log\frac{f(s|\alpha^*,\theta^*)}{f(s|\alpha,\theta)} + \min_{\theta\in\Theta}\mathbb{E}^*\log\frac{f(s|\alpha^*,\theta^*)}{f(s|\alpha,\theta)}\right)\\
&= -\min_{\alpha\in\mathcal{A}}\left[\min_{\theta\in\Theta_\tau} r(\alpha,\theta)\right]
\end{aligned}$$

Since α_t^τ maximizes $U_t(\tau)$ for each t (hence maximizes $\frac{1}{t}\log U_t(\tau)$), it follows immediately that

$$\mathcal{A}_\infty^\theta \subset \min_{\alpha\in\mathcal{A}}\left[\min_{\theta\in\text{supp}(\tau)} r(\alpha,\theta)\right] \quad \mathbb{P}^* - a.s..$$

Similarly, limit beliefs of bias-τ individuals will settle on zero-potential states under models on the RHS. □

A.8 Proof of Corollary 2

Proof. If $\theta_1 \in \mathcal{U}_\mathcal{A}^\tau \backslash \mathcal{V}_\mathcal{A}^\tau$, there must exist some θ_2 such that $\tau(\theta_2) > \tau(\theta_1)$. Denote by $\mathcal{A}(\theta) = \{\alpha\in\mathcal{A} : r(\alpha,\theta) = 0\}$, the set of models under which state θ is a zero-potential state. For all $\alpha_1 \in \mathcal{A}(\theta_1)$ and all $\alpha_2 \in \mathcal{A}(\theta_2)$, we have:

$$\lim U_t(\tau|\alpha_1) \to \tau(\theta_1) < \tau(\theta_2) = \lim U_t(\tau|\alpha_2) \quad \mathbb{P}^* - a.s.$$

where the convergence comes from the fact that every model has a unique zero-potential state, so $\mu(\theta|\alpha)$ converges to the Dirac measure on the zero-potential state. Since all models in $\mathcal{A}(\theta_1)$ are strictly dominated in terms of utilities, so $\mathcal{A}(\theta_1)$ will not be chosen in the limit, which implies that θ_1 is not asymptotically carried on. □

A.9 Verification of $\mathbb{E}a_t \to 0$ in Example 12

I am going to verify that if individuals update according to model α_2, the expected report approaches 0 in the limit.

Proof. I assume a tie-breaking rule that when indifferent between two actions, individuals choose the action with a larger index[28]. Denote by $E_t = \left\{\frac{\mu_t(1|\alpha_2)}{\mu_t(-1|\alpha_2)} \geq 1\right\}$ and $A_t = \left\{\frac{\mu_t(1|\alpha_2)}{\mu_t(0|\alpha_2)} \geq 1\right\}$ and $B_t = \left\{\frac{\mu_t(-1|\alpha_2)}{\mu_t(0|\alpha_2)} > 1\right\}$. By definition, we have:

$$\begin{aligned}\mathbb{E}^* r_t &= \mathbb{P}^*\left(r_t = 1\right) \times 1 + \mathbb{P}^*\left(r_t = -1\right) \times (-1) \\ &= \mathbb{P}^*\left(E_t \cap A_t\right) \times 1 + \mathbb{P}^*\left(E_t^c \cap B_t\right) \times (-1) \quad (16) \\ &= \left[\mathbb{P}^*\left(E_t\right) - \mathbb{P}^*\left(E_t \cap A_t^c\right)\right] - \left[\mathbb{P}^*\left(E_t^c\right) - \mathbb{P}^*\left(E_t^c \cap B_t^c\right)\right]. \quad (17)\end{aligned}$$

Denoting by $S_t = \log\left[\frac{\mu_t(1|\alpha_2)}{\mu_t(-1|\alpha_2)}\right]$ and applying Bayes rule, we get

$$S_{t+1} = S_t + \log\left[\frac{f\left(s_t|1, \alpha_2\right)}{f\left(s_t|-1, \alpha_2\right)}\right] := S_t + X_t,$$

where $X_t = \log\left[\frac{f(s_t|1,\alpha_2)}{f(s_t|-1,\alpha_2)}\right]$. It is easy to verify that state 1 and -1 have the same relative entropy under model α_2, so $\mathbb{E}^* X_t = 0$. Denoting by $\sigma^2 = \mathbb{E}^* X_t^2$, the Central Limit Theorem (CLT) implies that:

$$\frac{S_t}{\sigma\sqrt{t}} \Rightarrow \mathcal{N}\left(0, 1\right),$$

where "$\Rightarrow$" means convergence in distribution. As a result,

$$\lim_{t\to\infty} \mathbb{P}^*\left(E_t\right) = \lim_{t\to\infty} \mathbb{P}^*\left(S_t \geq 0\right) = \lim_{t\to\infty} \mathbb{P}^*\left(\frac{S_t}{\sigma\sqrt{t}} \geq 0\right) = \frac{1}{2}.$$

Similarly, we have $\lim_{t\to\infty} \mathbb{P}^*\left(E_t^c\right) = \frac{1}{2}$. Since the relative entropy of state 1 and -1 under α_2 is strictly than the relative entropy of state 0 under α_2, we have:

$$\lim_{t\to\infty} \mathbb{P}^*\left(A_t\right) = \lim_{t\to\infty} \mathbb{P}^*\left(B_t\right) = 1,$$

which directly implies that

$$\lim_{t\to\infty} \mathbb{P}^*\left(E_t \cap A_t^c\right) = \lim_{t\to\infty} \mathbb{P}^*\left(E_t^c \cap B_t^c\right) = 0.$$

Taking limits on both sides of (17), we get $\lim_{t\to\infty} \mathbb{E}^* r_t = 0$, so the expected action under model α_2 approaches 0. □

A.10 Verification of the Claims in Example 13

(I) Suppose that the agent updates via the *full Bayesian rule*. Given our payoff structures, the agent's max-min expected utility of choosing action l is $V^{FB}(l) = \underline{\mu}(L)$, where $\underline{\mu}(L)$ denotes the lowest probability on state L. Similarly, we have $V^{FB}(r) = \underline{\mu}(R)$, and $V^{FB}(m) = 1/2$. From the previous discussion, the lowest probability on state L arises when the agent updates according to model α_2, where signal b (i.e., the majority signal) is interpreted as "bad news" for state L.

[28] This is only for convenience purpose. The result also holds for other tie-breaking rules.

Symmetrically, the lowest probability on state R arises when he updates according to model α_1. Simple calculations show that

$$\underline{\mu}(L) = \mu(L|\alpha_2) = \frac{\left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^6}{\left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^6 + \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6 + \left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{4}\right)^6} \approx 0.009$$

$$\underline{\mu}(R) = \mu(R|\alpha_1) = \frac{\left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^6}{\left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^6 + \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6 + \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^6} \approx 0.04,$$

which are less than $V^{FB}(m) = 1/2$, so the optimal choice is $a^{FB} = m$.

(II) Suppose that the agent updates via the *maximum likelihood rule*. We can calculate the probabilities of the observed signals under each model,

$$\mathbb{P}(\boldsymbol{s}|\alpha_1) = \frac{1}{3} \times \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^6 + \frac{1}{3} \times \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6 + \frac{1}{3} \times \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^6 \approx 0.0048$$

$$\mathbb{P}(\boldsymbol{s}|\alpha_2) = \frac{1}{3} \times \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^6 + \frac{1}{3} \times \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6 + \frac{1}{3} \times \left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{3}\right)^6 \approx 0.02,$$

so model α_2 is more likely to generate the observed signals. The agent then updates according to α_2, and his posterior on state L is

$$\mu^{ML}(R) = \mu(R|\alpha_2) = \frac{\left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{4}\right)^4}{\left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^4 + \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^4 + \left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{4}\right)^4} \approx 0.53 > \frac{1}{2}.$$

The expected utility from action r is $V^{ML}(r) = \mu^{ML}(R) > \frac{1}{2} \geq \max\left\{V^{ML}(m), V^{ML}(l)\right\}$, so the optimal choice is $a^{ML} = r$.

(III) Suppose that the agent updates via the *biased updating rule* and is biased toward state M. Simple application of Bayes rule shows that

$$\mu(M|\alpha_1) = \frac{\frac{1}{3} \times \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6}{\mathbb{P}(\boldsymbol{s}|\alpha_1)} > \frac{\frac{1}{3} \times \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6}{\mathbb{P}(\boldsymbol{s}|\alpha_2)} = \mu(M|\alpha_2),$$

where the inequality comes from the fact that $\mathbb{P}(\boldsymbol{s}|\alpha_1) < \mathbb{P}(\boldsymbol{s}|\alpha_2)$. Therefore, the biased agent will update according to model α_1. Under model α_1, the probability of state L is

$$\mu^{BS}(L) = \mu(L|\alpha_1) = \frac{\left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^6}{\left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^6 + \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^6 + \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)^6} \approx 0.68 > \frac{1}{2}.$$

The expected utility from action l is $V^{BS}(l) = \mu^{BS}(L) > \frac{1}{2} \geq \max\left\{V^{BS}(m), V^{BS}(r)\right\}$, so the optimal choice should be $a^{BS} = l$.

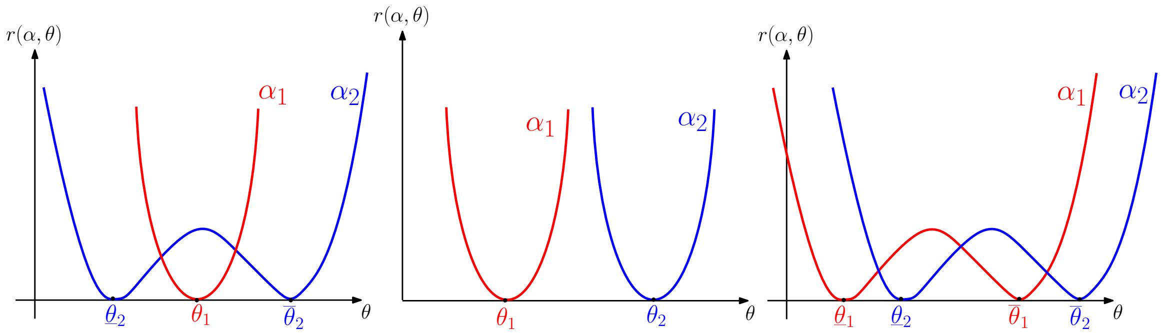

Figure 4: Illustration of τ-complementary

B Supplementary Materials I: Overconfidence under Ambiguity

This subsection expands on the intuition developed in Example 12 and provides conditions under which the overconfidence exists in the limit. For all model $\alpha \in \mathbb{A}$, denote by Θ_α the set of zero-potential states under model α. Let $\overline{\tau}(\alpha)$ and $\underline{\tau}(\alpha)$ denote the highest and the lowest $\tau(\theta)$ for all $\theta \in \Theta_\alpha$ respectively.

Definition 3. Model α τ-*dominates* α', if $\underline{\tau}(\alpha) > \overline{\tau}(\alpha')$. The τ-*core of* $\mathcal{A}$, denoted by $C^\tau(\mathcal{A})$, consists of all models in $\mathcal{A}$ which are not τ-dominated by any other model in $\mathcal{A}$.

The τ-core of $\mathcal{A}$ consists of all models whose zero-potential states are not strictly dominated by any other model in terms of τ. Following the same logic as in Corollary 2, in the limit, only models in the τ-core of $\mathcal{A}$ matter for an individual with bias τ. As a result, when analyzing individuals with bias τ, we can restrict our attention to the τ-core models without loss of generality. The following concept is essential in characterizing the overconfidence effect.

Definition 4. For all $\alpha, \alpha' \in \mathbb{A}$, models α and α' are ***τ-complementary*** if

$$\left[\overline{\tau}(\alpha) - \overline{\tau}\left(\alpha'\right)\right] \cdot \left[\underline{\tau}(\alpha) - \underline{\tau}\left(\alpha'\right)\right] < 0.$$

Model set $\mathcal{A}$ is called ***τ-complementary*** if for all $\alpha \in C^\tau(\mathcal{A})$, there exists some $\alpha' \in \mathcal{A}$ such that α and α' are τ-complementary.

The notion of τ-complementarity describes some "crossing" feature of utilities generated by zero-potential states under α and α'. Roughly speaking, if the highest utility under α is higher than the highest utility under α', then the lowest utility under α must be lower than the lowest utility under α'. Figure 4 illustrates the concept, where the vertical axis denotes the information potential $r(\alpha, \theta)$. Assuming that τ is strictly monotone, the model set $\{\alpha_1, \alpha_2\}$ is τ-complementary in the left graph, but is not τ-complementary in the middle and right graph.

Theorem 3. [The Overconfidence Effect] *Assume that* $|\Theta| < \infty$. *For all bias* $\tau \in \mathcal{T}$, *if* $\mathcal{A}$ *is*

τ-complementary, then for all $\alpha \in \mathcal{A}$, there is some $\varepsilon > 0$ such that:

$$\liminf_{t\to\infty} \mathbb{P}^* \left[U_t(\tau) - U_t(\tau|\alpha) \geq \varepsilon\right] > 0$$

Recall that $U_t(\tau)$ denotes the time-t belief utility of an individual with bias τ, and $U_t(\tau|\alpha)$ denotes the time-t belief utility if this individual updates according to model α. By definition, $U_t(\tau) \geq U_t(\tau|\alpha)$ for all $\alpha \in \mathcal{A}$, so Theorem 3 implies that the expected belief utility of a biased individual is *strictly* higher than that of a Bayesian individual with any feasible model perception. This result is interesting since it shows that the effect of ambiguity cannot be replaced by any specific model, so biased individuals can exhibit behavioral patterns that are absent in the Bayesian framework.

The intuition has been discussed in Example 12. To illustrative it more directly, let us look at the left graph in Figure 4. Suppose that τ is strictly increasing and $\mathcal{A} = \{\alpha_1, \alpha_2\}$. Following the arguments in Example 12, in the limit, the τ-individual will adopt model α_2 when the beliefs under α_2 accumulate around the higher zero-potential state $\overline{\theta}_2$, and adopt model α_1 when the beliefs under α_2 accumulate around the lower zero-potential state $\underline{\theta}_2$. We can think of α_2 as a "riskier" model than α_1 in the sense that it can generate a higher utility gain but can also incur a larger utility loss in the limit (graphically, its zero-potential states are more "spread-out" than model α_1). The presence of the "safe model" α_1 allows individuals to exploit the utility gain as well as hedge against the utility loss from α_2. As a result, individuals are strictly better-off with the presence of model ambiguity.

B.1 Proof of Theorem 3

Proof. Consider any model α and its τ-complementary model α'. Without loss of generality, I assume that $\overline{\tau}(\alpha') > \overline{\tau}(\alpha)$ (the case $\overline{\tau}(\alpha') < \overline{\tau}(\alpha)$ follows an analogous argument). By definition, $\overline{\tau}(\alpha') = \tau(\theta_1)$ for some $\theta_1 \in \Theta_{\alpha'}$, where $\Theta_{\alpha'} = \{\theta_1, \theta_2, ..., \theta_{K+1}\}$. Notice that we must have $K \geq 1$ since $K = 0$ (i.e., $\Theta_{\alpha'} = \{\theta_1\}$) implies that $\overline{\tau}(\alpha') = \underline{\tau}(\alpha') > \overline{\tau}(\alpha) \geq \underline{\tau}(\alpha)$, which contradicts the fact that α and α' are τ-complementary. Therefore, it is meaningful to define two K-dimensional vectors:

$$\boldsymbol{S}_t = \left(\log\frac{\mu_t(\theta_2|\alpha')}{\mu_t(\theta_1|\alpha')}, ..., \log\frac{\mu_t(\theta_{K+1}|\alpha')}{\mu_t(\theta_1|\alpha')}\right) \quad \text{and} \quad \boldsymbol{X}_t = \left(\log\frac{f(s_t|\theta_2, \alpha')}{f(s_t|\theta_1, \alpha')}, ..., \log\frac{f(s_t|\theta_{K+1}, \alpha')}{f(s_t|\theta_1, \alpha')}\right).$$

By Bayes rule, for all $k \in \{1, ..., K\}$, we have

$$\frac{\mu_{t+1}(\theta_{k+1}|\alpha')}{\mu_{t+1}(\theta_1|\alpha')} = \frac{\mu_t(\theta_{k+1}|\alpha')}{\mu_t(\theta_1|\alpha')} \times \frac{f(s_t|\theta_{k+1}, \alpha')}{f(s_t|\theta_1, \alpha')},$$

which implies that

$$\boldsymbol{S}_{t+1} = \boldsymbol{S}_t + \boldsymbol{X}_t \text{ or } \boldsymbol{S}_t = \sum_{i=1}^{t} \boldsymbol{X}_i. \tag{18}$$

For all $k \in \{1, ..., K\}$, we have $\theta_{k+1} \in \Theta_{\alpha'}$, so the relative entropy of state $k+1$ and state 1 must be equal, which implies $\mathbb{E}^* \boldsymbol{X}_t = 0$. Applying the multidimensional Central Limit Theorem (CLT) , we get

$$\frac{\boldsymbol{S}_t}{\sqrt{t}} \Rightarrow \mathcal{N}_K(0, \Sigma) \coloneqq Z \text{ with } \Sigma = \begin{pmatrix} \mathbb{E}^* \boldsymbol{X}_{t1}^2 & \mathbb{E}^* \boldsymbol{X}_{t1} \boldsymbol{X}_{t2} & \cdots & \mathbb{E}^* \boldsymbol{X}_{t1} \boldsymbol{X}_{tK} \\ \mathbb{E}^* \boldsymbol{X}_{t2} \boldsymbol{X}_{t1} & \mathbb{E}^* \boldsymbol{X}_{t2}^2 & & \mathbb{E}^* \boldsymbol{X}_{t2} \boldsymbol{X}_{tK} \\ \vdots & \vdots & & \vdots \\ \mathbb{E}^* \boldsymbol{X}_{tK} \boldsymbol{X}_{t1} & \mathbb{E}^* \boldsymbol{X}_{tK} \boldsymbol{X}_{t2} & \cdots & \mathbb{E}^* \boldsymbol{X}_{t1} \boldsymbol{X}_{tK} \end{pmatrix}, \tag{19}$$

where X_{tk} denotes the k-th component of X_t, and $\mathcal{N}_K(0, \Sigma)$ denotes the K-dimensional multivariate normal distribution with covariance matrix Σ. Notice that it can be verified that Σ is a well-defined since every entry of Σ is finite from Assumption 2. Based on Assumption 3, we can further show that Σ is positive definite, so Z admits a strictly positive density function on $\mathbb{R}^K$.[29] As a result, from (19), for all $M \in \mathbb{R}_{++}$, we have

$$\lim_{t\to\infty} \mathbb{P}^* (\boldsymbol{S}_t \leq -\log M \cdot \boldsymbol{e}) = \lim_{t\to\infty} \mathbb{P}^* \left(\frac{\boldsymbol{S}_t}{\sqrt{t}} \leq -\frac{\log M}{\sqrt{t}} \boldsymbol{e} \right) = \mathbb{P}(Z \leq 0) \coloneqq \xi > 0, \tag{20}$$

where $\boldsymbol{e} = (1, 1, ..., 1) \in \mathbb{R}^K$. Defining $E_t = \{\boldsymbol{S}_t \leq -\log M \cdot \boldsymbol{e}\}$, for all signal paths $s^\infty \in E_t$, we have

$$\mu_t(\theta_{k+1}|\alpha') \leq \frac{1}{M} \cdot \mu_t(\theta_1|\alpha') \quad \text{for all } k \in \{1, ..., K\}.$$

Summing up all $k \in \{1, ..., K\}$, we get

$$\mu_t(\Theta_{\alpha'}|\alpha') \leq \left(1 + \frac{K}{M}\right) \times \mu_t(\theta_1|\alpha'),$$

so (20) implies that

$$\lim_{t\to\infty} \inf \mathbb{P}^* \left(\mu_t(\theta_1|\alpha') \geq \frac{M}{M+K} \cdot \mu_t(\Theta_{\alpha'}|\alpha') \right) \geq \xi,$$

or equivalently,

$$\begin{aligned} \lim_{t\to\infty} \inf \mathbb{P}^* & \left(\left\{ \mu_t(\theta_1|\alpha') \geq \frac{M}{M+K} \cdot \mu_t(\Theta_{\alpha'}|\alpha') \right\} \cap \left\{ \mu_t(\Theta_{\alpha'}|\alpha') \geq \frac{M+K}{M+K+1} \right\} \right) \\ + \mathbb{P}^* & \left(\left\{ \mu_t(\theta_1|\alpha') \geq \frac{M}{M+K} \cdot \mu_t(\Theta_{\alpha'}|\alpha') \right\} \cap \left\{ \mu_t(\Theta_{\alpha'}|\alpha') < \frac{M+K}{M+K+1} \right\} \right) \geq \xi > 0 \end{aligned} \tag{21}$$

Applying Berk (1966)'s result, we have $\lim_{t\to\infty} \mathbb{P}^* \left(\mu_t(\Theta_{\alpha'}|\alpha') < \frac{M+K}{M+K+1} \right) = 0$, so (21) implies that

$$\lim_{t\to\infty} \inf \mathbb{P}^* \left(\left\{ \mu_t(\theta_1|\alpha') \geq \frac{M}{M+K} \cdot \mu_t(\Theta_{\alpha'}|\alpha') \right\} \cap \left\{ \mu_t(\Theta_{\alpha'}|\alpha') \geq \frac{M+K}{M+K+1} \right\} \right) \geq \xi,$$

[29] see Fudenberg, Lanzani and Strack (2020) for a detailed argument on the positive definiteness.

which directly implies that

$$\lim_{t\to\infty}\inf \mathbb{P}^*\left(\mu_t\left(\theta_1|\alpha'\right) \geq \frac{M}{M+K+1}\right) \geq \xi > 0 \tag{22}$$

for all possible $M \in \mathbb{R}_{++}$.

For simplicity in notation, I normalize the τ such that the highest value is bounded by 1 (e.g., by multiplying some constant. It is easy to see that this transformation will not change the problem). Define $\varepsilon = \frac{1}{3}\left[\overline{\tau}\left(\alpha'\right) - \overline{\tau}\left(\alpha\right)\right]$, which is a strictly positive number by definition. From the definition, we know that $U_t\left(\tau\right) = \max_\alpha U_t\left(\tau|\alpha\right) \geq U_t\left(\tau|\alpha'\right)$, so

$$\liminf_{t\to\infty}\mathbb{P}^*\left[U_t\left(\tau\right) - U_t\left(\tau|\alpha\right) \geq \varepsilon\right] \geq \liminf_{t\to\infty}\mathbb{P}^*\left[U_t\left(\tau|\alpha'\right) - U_t\left(\tau|\alpha\right) \geq \varepsilon\right].$$

Define $E_t = \left\{U_t\left(\tau|\alpha'\right) - U_t\left(\tau|\alpha\right) \geq \varepsilon\right\}$ and $E_t^1 = \left\{\mu_t\left(\theta_1|\alpha'\right) \geq \frac{\varepsilon+1}{3\varepsilon+1}\right\}$ and $E_t^2 = \left\{\mu_t\left(\Theta_\alpha|\alpha\right) \geq \frac{\varepsilon+1}{3\varepsilon+1}\right\}$. For all $s^\infty \in E_t^1 \cap E_t^2$, we have:

$$\begin{aligned} U_t\left(\tau|\alpha'\right) - U_t\left(\tau|\alpha\right) &\geq \mu_t\left(\theta_1|\alpha'\right)\cdot\overline{\tau}\left(\alpha'\right) - \left(\mu_t\left(\Theta_\alpha|\alpha\right)\cdot\overline{\tau}\left(\alpha\right) + \left(1-\mu_t\left(\Theta_\alpha|\alpha\right)\right)\cdot 1\right) \\ &\geq \frac{\varepsilon+1}{3\varepsilon+1}\left[\overline{\tau}\left(\alpha'\right) - \overline{\tau}\left(\alpha\right)\right] - \frac{2\varepsilon}{3\varepsilon+1} = \varepsilon, \end{aligned}$$

so $s^\infty \in E_t$, which implies that $E_t^1 \cap E_t^2 \subset E_t$. As a result,

$$\begin{aligned} \liminf_{t\to\infty}\mathbb{P}^*\left[E_t\right] &\geq \liminf \mathbb{P}^*\left[E_t^1 \cap E_t^2\right] \\ &\geq \liminf\left(\mathbb{P}^*\left[E_t^1\right] - \mathbb{P}^*\left[E_t^1 \cap \left(E_t^2\right)^c\right]\right) \\ &\geq \liminf\left(\mathbb{P}^*\left[E_t^1\right] - \mathbb{P}^*\left[\left(E_t^2\right)^c\right]\right) \\ &\geq \liminf \mathbb{P}^*\left[E_t^1\right] + \liminf \mathbb{P}^*\left[E_t^2\right] - 1 \end{aligned} \tag{23}$$

From Berk's result, we know that $\lim \mathbb{P}^*\left(E_t^2\right) = 1$. From (22), we can set $M = \frac{\varepsilon+1}{2\varepsilon}\left(K+1\right)$, so $\liminf \mathbb{P}^*\left(E_t^1\right) \geq \xi > 0$. Consequently,

$$\liminf_{t\to\infty}\mathbb{P}^*\left[U_t\left(\tau|\alpha'\right) - U_t\left(\tau|\alpha\right) \geq \varepsilon\right] = \liminf_{t\to\infty}\mathbb{P}^*\left[E_t\right] \geq \xi > 0,$$

so the claim is proved. □

C Supplementary Materials II: Belief-Dependent Bias

In this section, I present a more general version of ***belief-dependent biased updating rule*** than Examples 15 and characterize limit beliefs under this rule. Suppose that the state space is finite, that is, $\Theta = \{\theta_1, ..., \theta_N\}$ for some finite N, and the signal space S is also finite. The belief simplex $\Delta\left(\Theta\right)$ is divided into N regions $\Delta_{\theta_1}, ..., \Delta_{\theta_N}$, which can possibly intersect. For all $\theta \in \Theta$, the Dirac belief δ_θ is uniquely contained in the interior of region Δ_θ and not in any other region. At time $t = 0$, every individual is endowed with some initial biased state $\theta \in \Theta$ and a full-support prior $\mu_0^\theta \in \Delta_\theta$. The biased state can change overtime, where b_t^θ denotes the time-t biased state of the individual with initial bias θ. The bias process $\left\{b_t^\theta\right\}$ evolves with beliefs according to the following rule:

$$\forall \theta, \theta_{i_1}, .., \theta_{i_k} \in \Theta : \quad b_t^\theta \in \{\theta_{i_1}, \theta_{i_2}, ..\theta_{i_k}\} \text{ if } \mu_t^\theta \in \Delta_{\theta_{i_1}} \cap \Delta_{\theta_{i_2}} .. \cap \Delta_{\theta_{i_k}} \tag{24}$$

In other words, if the time-t belief μ_t^θ enters some region, the time-t bias b_t^θ changes to the state that corresponds to that region. For the case where μ_t^θ belongs to the intersection of multiple regions, b_t^θ is determined by some tie-breaking rule. For realistic purposes, I assume that the tie-breaking rule features inertia on the initial bias. More specifically, if $b_{t-1}^\theta = \theta$ and $\mu_t^\theta \in \Delta_\theta$, then we must have $b_t^\theta = \theta$. In other words, if the last-period bias is the initial bias, and if the current belief stays in the region favoring the initial bias, then the individual will stick to his initial bias. Given the last period bias b_{t-1}^θ, the time-t belief μ_t^θ is updated according to the model that can best support b_{t-1}^θ. More precisely,

$$\forall \theta, \theta' \in \Theta : \quad \mu_t^\theta \left(\theta'\right) = \mu_{t-1} \left(\theta' | \alpha_t^\theta\right) \text{ where } \alpha_t^\theta \in \arg\max_{\alpha \in \mathcal{A}} \mu \left(b_{t-1}^\theta | \alpha\right), \tag{25}$$

where the specific value of α_t^θ is determined by some tie-breaking rule when there are multiple maximizers. Notice that the only difference between (25) and the biased updating rule is that the bias b_t^θ can change overtime. Below is another example.

Example 17. [Belief-Dependent Bias II] Define $\Delta_G = \{\mu : \mu(G) \geq k\}$ and $\Delta_B = \{\mu : \mu(B) \geq k\}$, for some $k \leq 1/2$. Now suppose that the updating rule of $\left\{b_t^\theta\right\}$ evolves as follows

$$b_t^\theta = \begin{cases} G & \text{if } \mu_t(G) > 1 - k \\ B & \text{if } \mu_t(G) < k \\ b_{t-1}^\theta & \text{if } \mu_t(G) \in [k, 1-k] \end{cases}.$$

If the belief uniquely belongs to Δ_G (or Δ_B), the bias is just equal to G (or B). If the belief falls into the intersection of Δ_G and Δ_B, the bias is determined by the last-period bias. This captures the idea that individuals exhibit some inertia to the current bias, that is, they are willing to switch their bias only if they face strong opposing evidence. When $k = 1/2$, it corresponds to confirmation bias model in Example 15; when $k = 0$, it corresponds to the benchmark situation where individuals hold a fixed bias.

C.1 Characterizations of Limit Beliefs

This subsection provides the characterizations of limit beliefs under the belief-dependent biased updating rule. For simplicity of discussion, this section maintains the assumption that every model has a unique zero-potential state.

Assumption 4. *For all $\alpha \in \mathcal{A}$, there exists a unique $\theta \in \Theta$ such that $r(\alpha, \theta) = 0$.*

In other words, under every model, there is a unique state that delivers the "smallest" distance between the induced distribution and the true distribution. This assumption is adopted throughout this section. The first result is that all beliefs will converge almost surely as stated below.

Proposition 4. *For all initial bias $\theta \in \Theta$, we have $\mu_t^\theta \to \mu_\infty^\theta$ $\mathbb{P}^*$-almost surely, where μ_∞^θ is a random variable that satisfies*

$$\sum_{\theta' \in \mathcal{U}_\mathcal{A}} \mathbb{P}^* \left(\mu_\infty^\theta = \delta_{\theta'} \right) = 1,$$

where $\mathcal{U}_\mathcal{A}$ denotes the set of zero-potential states under $\mathcal{A}$.

Proposition 4 says that: (i) beliefs will almost surely converge, and (ii) the limit belief is a Dirac belief on some zero-potential state. One natural question is that: will every zero-potential state be visited with a strictly positive probability? Unfortunately, it is not always true. Below is one example.

Example 18. Consider the case described in Example 15. There are two individuals, one with initial bias G, the other with initial bias B. Both individuals hold the same prior that $\mu_0(G) = \mu_0(B) = 1/2$. The set of models is $\mathcal{A} = \{\alpha_1, \alpha_2\}$ and the data-generating processes are described below

α_1	a	b
G	$3/4$	$1/4$
B	$1/4$	$3/4$

α_2	a	b
G	$1/4$	$3/4$
B	$3/4$	$1/4$

,

where a and b are two possible signals. Suppose that the true state is G and the correct model α_1. It is easy to see that both states are zero-potential states: if individuals update according to model α_1, they will believe in state G; if they update according to model α_2, they will believe in state B.

Consider an individual with initial bias G, if he received more signal a, he would adopt model α_1; if he received more signal b, he would adopt model α_2. In both cases, the belief on state G is higher than $1/2$, which prevents the individual from modifying his initial bias. As a result, we must have $\mu_t^G(G) \geq 1/2$ for all t. Proposition 4 further implies that $\mu_t^G \to \delta_G$ almost surely, which means that every individual will perfectly confirm his bias with probability 1, so not every zero-potential state will be visited with a positive probability.

In this example, all signals are "controversial" in the sense that they can be interpreted as both good news and bad news for every state. Therefore, individuals can always confirm their initial biases by misinterpreting any negative news as positive. The following assumption restricts this kind of misinterpretation.

Assumption 5. [Existence of non-controversial signals] *For all $\theta \in \Theta$, there exists some signal $s_\theta \in S$ such that:*

$$\min_{\theta' \neq \theta} \frac{f(s|\alpha, \theta)}{f(s|\alpha, \theta')} > 1 \text{ for all } \alpha \in \mathcal{A}$$

Assumption 5 says that for every state θ, there a signal s_θ that is unanimously regarded as the good news for state θ by all models. Under this assumption, it is impossible to interpret any signal as good news to an arbitrary state as in Example 18. Notice that Assumption 5 can be weak since it only requires the existence of one such s_θ without restricting how large the probability is. It is possible that all models only agree on a tiny fraction of signals but disagree on signals that occur with a large probability. Under this assumption, we have the following result.

Proposition 5. *If Assumption 5 holds, then for all initial bias $\theta \in \Theta$ and for all state $\theta' \in \mathcal{U}_{\mathcal{A}}$, we have $\mathbb{P}^* \left(\mu_\infty^\theta = \delta_{\theta'}\right) > 0$.*

Proposition 4 shows a strong result that for *all* possible initial bias, *every* zero-potential state will be visited with a strictly positive probability. One implication is that **complete learning may not occur for all individuals**. Unless the true state θ^* minimizes relative entropy for all possible models (i.e., $\mathcal{U}_{\mathcal{A}} = \{\theta^*\}$), incorrect learning arises with a strictly positive probability for every individual. Even if an individual is initially biased toward the true state, he may arrive at some incorrect state with a positive probability. Recall that in the case with fixed bias, when the model set is correctly specified, individuals with correct bias can almost surely learn the true state. As such, Proposition 4 suggests that allowing bias to change can even make it harder to achieve correct learning in some sense.

Previous two propositions hold for individuals with all possible initial biases. However, it is conceivable that individuals starting from different regions (i.e., with different initial biases) may visit each state with different probabilities. Therefore, it becomes natural to ask: how does the distribution of μ_∞^θ vary with the initial bias θ? This question is generally very hard to answer since it is a challenging task to solve for the exact distribution of μ_∞^θ. However, it is possible to characterize the distribution in a limit case where the "inertia" of the initial bias is sufficiently large.

Definition 5. For every θ, define $R_\theta = \sup \left\{ r \geq 0 : B_{\frac{r}{r+1}} \left(\delta_\theta\right) \subset \Delta_\theta \right\}$, where $R_\theta \in [0, \infty]$.

Intuitively, the magnitude of R_θ measures the size of Δ_θ, which captures the degree of reluctance of switching the bias. If $R_\theta \to 0$, the region Δ_θ approaches the Dirac belief δ_θ, meaning that individuals are only willing to hold the bias θ within a small neighborhood of δ_θ. If $R_\theta \to \infty$, the region Δ_θ approaches the whole space, and individuals with initial bias θ will stick to it for a large set of beliefs. We have the following proposition.

Proposition 6. *For all initial bias $\theta \in \Theta$, as $R_\theta \to \infty$, we have*

$$\sum_{\theta' \in \mathcal{U}_{\mathcal{A}}^\theta} \mathbb{P}^* \left(\mu_\infty^\theta = \delta_{\theta'}\right) \to 1,$$

where $\mathcal{U}_{\mathcal{A}}^\theta$ denotes the set of zero-potential states under models that minimize the potential of state θ.

When the inertia is sufficiently large, with a large probability, individuals will land on zero-potential states under models that minimize the potential of their initial biases, which correspond to the limit belief carrier in the benchmark case (see Theorem 1). This proposition builds the connection between the model with fixed bias and the model with belief-dependent bias. It demonstrates that the fixed-bias model approximates the situation where the degree of inertia is sufficiently large.

C.2 Proof of Proposition 4

Let $\theta(\alpha)$ be the zero-potential state under model α. We have

$$\mathbb{P}^*\left(\left\{\omega : \min_{\alpha\in\mathcal{A}} |\mu_t(\theta(\alpha)|\alpha) - 1| \to 0\right\}\right) = 1 \tag{26}$$

where the uniform convergence comes from the same reasoning as in the proof of Lemma 1. Define $B_\varepsilon(\delta_\theta) \equiv \{\mu \in \Delta(\Theta) : \mu(\theta) > 1-\varepsilon\}$, so (26) further implies that for all (small) $\varepsilon \in (0,1)$ and for all (large) $\delta \in (0,1)$, there exists some $T < \infty$ such that:

$$\mathbb{P}^*(E) \coloneqq \mathbb{P}^*\left(\left\{\omega : \mu_t(.|\alpha) \in B_\varepsilon\left(\delta_{\theta(\alpha)}\right) \text{ for all } \alpha \in \mathcal{A} \text{ and } t \geq T\right\}\right) > \delta. \tag{27}$$

In (27), I set ε to be sufficiently small such that $B_\varepsilon\left(\delta_{\theta(\alpha)}\right)$ is uniquely contained in the region containing $\delta_{\theta(\alpha)}$, that is, $B_\varepsilon\left(\delta_{\theta(\alpha)}\right) \subset \Delta_{\theta(\alpha)} \backslash \cup_{\theta\neq\theta(\alpha)} \Delta_\theta$. We then have the following claim

Claim 1. *For all signal path $\omega \in E$, the belief μ_t^θ is trapped in $B_\varepsilon(\delta_{\theta'})$ for some θ' forever.*

Proof. Let α_T be the time-T model adopted by the individual, where T is as defined in (27).

(1) When $t = T$, $\mu_T^\theta = \mu_T(.|\alpha_T)$, by definition, we have $\mu_T^\theta \in B_\varepsilon\left(\delta_{\theta(\alpha_T)}\right)$. For sufficiently small ε, $\mu_T^\theta \in B_\varepsilon\left(\delta_{\theta(\alpha_T)}\right)$ implies that $\mu_T^\theta \in \Delta_{\theta(\alpha)}$ and $\mu_T^\theta \notin \Delta_\theta$ for all $\theta \neq \theta(\alpha)$. Therefore, $b_T = \theta(\alpha_T)$.

(2) When $t = T+1$,

$$\mu_{T+1}^\theta(\theta(\alpha_T)) = \max_{\alpha\in\mathcal{A}} \mu_{T+1}(\theta(\alpha_T)|\alpha) \geq \mu_{T+1}(\theta(\alpha_T)|\alpha_T) > 1-\varepsilon,$$

where the equality comes from $b_T = \theta(\alpha_T)$, and the last weak inequality comes from the definition of E. Consequently, $\mu_{T+1}^\theta \in B_\varepsilon\left(\delta_{\theta(\alpha_T)}\right)$ and $b_{T+1} = \theta(\alpha_T)$.

(3) The same reasoning applies to all $t \geq T+2$, so $\mu_t^\theta \in B_\varepsilon\left(\delta_{\theta(\alpha_T)}\right)$ and $b_t = \theta(\alpha_T)$ for all $t \geq T$. □

Previous claim implies that

$$E \subset \left\{\omega : \mu_t^\theta \in B_\varepsilon(\delta_{\theta'}) \text{ for some } \theta' \in \Theta \text{ for all } t \geq T\right\} \coloneqq E_1.$$

By definition,

$$\mathbb{P}^*\left(\left\{\omega : |\limsup \mu_t^\theta(\theta'') - \liminf \mu_t^\theta(\theta'')| \leq \varepsilon \text{ for all } \theta'' \in \Theta\right\}\right) \geq \mathbb{P}^*(E_1) \geq \delta \tag{28}$$

Note that (28) holds for arbitrary ε and δ between 0 and 1. Letting $\varepsilon \to 0$ and $\delta \to 1$, we have

$$\mathbb{P}^*\left(\left\{\omega : \mu_t^\theta \text{ converges}\right\}\right) = 1,$$

so μ_t^θ will converge to some limit measure μ_∞^θ. Previous discussion shows that μ_t^θ will be trapped in arbitrarily small neighborhood around some Dirac belief $\delta_{\theta(\alpha)}$ with arbitrarily large probability. Therefore, whenever μ_t^θ converges, it must converge to some Dirac belief on a zero-potential state. In other words,

$$\sum_{\theta' \in \mathcal{U}_\mathcal{A}} \mathbb{P}^*\left(\mu_\infty^\theta = \delta_{\theta'}\right) = 1,$$

so the proposition is proved.

C.3 Proof of Proposition 5

For state $\theta' \in \mathcal{U}_\mathcal{A}$, let α' be any model such that $r(\alpha', \theta') = 0$. We first have the following lemma:

Lemma 6. *For all $\varepsilon, \delta \in (0,1)$, there exists some $\varepsilon' \in (0, \varepsilon)$ such that for all $\mu \in B_{\varepsilon'}(\delta_{\theta'})$, we have:*

$$\mathbb{P}_\mu^*\left(\omega : \mu_t\left(|\alpha'\right) \in B_\varepsilon\left(\delta_{\theta'}\right) \text{ for all } t \geq 1\right) > \delta$$

where $\mathbb{P}_\mu^*$ denotes the probability measure if the individual has a prior μ. In other words, the probability that the posterior always stays inside a small neighborhood around $\delta_{\theta'}$ can be made arbitrarily large if the prior is sufficiently close to $\delta_{\theta'}$.

Proof. This is an intermediate result in the proof of Theorem 1 in Frick, Iijima and Ishii (2020). Detailed proofs are omitted for exposition. The idea is to make use of the supermartingale property near the Dirac belief. □

Proof of Proposition 5

Proof. We can pick the ε such that $B_\varepsilon(\delta_{\theta'}) \subset \Delta_{\theta'} \setminus \cup_{\theta \neq \theta'} \Delta_\theta$ and fix any $\delta > 0$. By Assumption 5, there exists some $K > 1$ such that

$$\forall \theta \neq \theta' : \quad \frac{f\left(s_{\theta'}|\alpha, \theta'\right)}{f\left(s_{\theta'}|\alpha, \theta\right)} > K \text{ for all } \alpha \in \mathcal{A}$$

Let $T = \left\lceil \log_K^{\frac{1-\varepsilon'}{\mu_0(\theta')}} \right\rceil + 1$, where $\mu_0(\theta')$ denotes the individual's initial prior on state θ'. If $s_1 = ... = s_T = s_{\theta'}$, we must have $\mu_T(\theta'|\alpha) > 1 - \varepsilon'$ for all model $\alpha \in \mathcal{A}$, so $\mu_T^\theta \in B_{\varepsilon'}(\delta_{\theta'})$ and $b_T = \theta'$. Due to the fact that data-generating process has full support, it occurs with a strictly positive probability that the first T consecutive signals are $s_{\theta'}$. Denote by E the event that $\mu_t(|\alpha')$ is trapped in $B_\varepsilon(\delta_{\theta'})$ for all $t \geq T$. Combining Lemma 6 and the fact that $\{s_1 = ... = s_T = s_{\theta'}\}$ is a positive-probability event, it is easy to see that $\mathbb{P}^*(E) > 0$. The rest of the proof resembles the proof of Proposition 4.

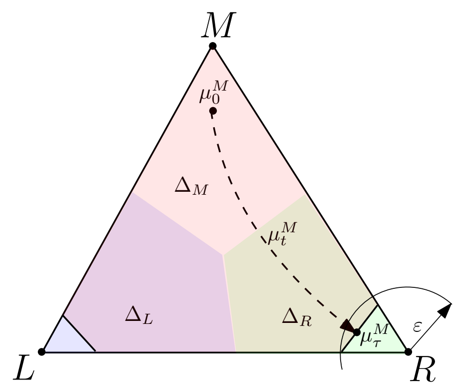

Figure 5: Illustration of Proposition 6

For all signal paths in E, we have $b_T = \theta'$, so

$$\mu^{\theta}_{T+1}\left(\theta'\right) \geq \max_{\alpha \in \mathcal{A}} \mu_{T+1}\left(\theta' | \alpha\right) \geq \mu_{T+1}\left(\theta' | \alpha'\right) > 1 - \varepsilon,$$

where the last inequality comes from the fact that $\mu_{T+1}\left(|\alpha'\right) \in B_{\varepsilon}\left(\delta_{\theta'}\right)$. Therefore, $\mu^{\theta}_{T+1} \in B_{\varepsilon}\left(\delta_{\theta'}\right)$ and $b_{T+1} = \theta'$. The same argument applies for all $t > T$. Consequently, for all signal paths in E, we have $\mu^{\theta}_t \in B_{\varepsilon}\left(\delta_{\theta'}\right)$ for all $t \geq T$. Proposition 4 implies that μ^{θ}_t converges to some Dirac belief almost surely, so the only possible limit is $\delta_{\theta'}$. As a result, for all signal paths in E, we have $\mu^{\theta}_t \to \delta_{\theta'}$ with exception only on null events. Recall that $\mathbb{P}^*(E) > 0$, so $\mathbb{P}^*\left(\mu^{\theta}_t \to \delta_{\theta'}\right) > 0$. Since θ' is an arbitrary zero-potential state, Proposition 5 is thus proved. □

C.4 Proof of Proposition 5

A Sketch of the Proof

Proposition 6 is intuitively straightforward, since we would expect the fixed bias case should be approximated by the situation where inertia is sufficiently large. To establish this limit-preserving property, we still need some technical arguments. The idea of the proof is illustrated in Figure 5. Consider an individual whose initial bias is M and suppose that $\mathcal{U}^M_{\mathcal{A}} = \{R\}$. Since M is not a zero-potential state,[30] the belief will almost surely escape from Δ_M in finite time. Besides, the larger the Δ_M becomes (i.e., the larger the inertia becomes), the longer the belief remains in Δ_M. Consequently, the escaping belief μ^M_τ approaches the limit case where individuals hold a fixed bias M. From Theorem 1, we know that if the bias is fixed, the belief $\mu^M_t \to \delta_R$ almost surely. As a result, if Δ_M is sufficiently large, with a sufficiently large probability, μ^M_τ will land in a small neighborhood of δ_R and stay within that neighborhood forever, in which case we must have $\mu^M_t \to \delta_R$ from Proposition 4.

[30] if it is, M should belong to $\mathcal{U}^M_{\mathcal{A}}$ by definition.

Formal Proof

Defining a stopping time $\tau = \inf\left\{t : \mu_t^\theta \notin \Delta_\theta\right\}$, we have the following lemmas.

Lemma 7. *There exists some $\theta_0 \in \Theta$ and some constant $A > 0$ such that*

$$\frac{\mu_\tau(\theta_0)}{\mu_\tau(\theta)} / \frac{\mu_0(\theta_0)}{\mu_0(\theta)} \geq A \times (R_\theta - 1) \tag{29}$$

The LHS of (29) represents the increment of the likelihood ratio between state θ_0 and θ at the stopping time, τ. Notice that the RHS of (29) is increasing in R_θ. Lemma 7 states a simple fact that when the inertia increases, it becomes harder to escape from Δ_θ in the sense that it requires a larger increment of beliefs for some state.

Proof. From the definition of τ and R_θ, we have $\mu_\tau(\theta) \leq \frac{1}{R_\theta}$, which implies that $\sum_{\tilde{\theta}\neq\theta} \mu_\tau\left(\tilde{\theta}\right) > 1 - \frac{1}{R_\theta}$. Therefore, there must exist some $\theta_0 \in \Theta$ such that $\mu_\tau(\theta_0) \geq \frac{1}{N}\left(1 - \frac{1}{R_\theta}\right)$, which further implies that $\frac{\mu_\tau(\theta_0)}{\mu_\tau(\theta)} \geq \frac{1}{N}(R_\theta - 1)$. □

Lemma 8. *For all $T < \infty$, as $R_\theta \to \infty$, we have $\mathbb{P}^*(\tau < T) \to 0$.*

This lemma states a simple fact that as the inertia increases, it takes longer for individuals to change their biases.

Proof. From Lemma 7, we know that:

$$\begin{aligned}
\mathbb{P}^*(\tau < T) &\leq \sum_{t<T}\sum_{\theta_0\neq\theta} \mathbb{P}^*\left(\frac{\mu_t(\theta_0)}{\mu_t(\theta)} / \frac{\mu_0(\theta_0)}{\mu_0(\theta)} \geq A \times (R_\theta - 1)\right) \\
&= \sum_{t<T}\sum_{\theta_0\neq\theta} \mathbb{P}^*\left(\prod_{k\leq t} \frac{f(s_k|\alpha,\theta_0)}{f(s_k|\alpha,\theta)} \geq A \times (R_\theta - 1)\right) \\
&\leq \sum_{t<T}\sum_{\theta_0\neq\theta} \mathbb{P}^*\left(\sum_{k\leq t} \left|\log \frac{f(s_k|\alpha,\theta_0)}{f(s_k|\alpha,\theta)}\right| \geq \log A(R_\theta - 1)\right)
\end{aligned}$$

Applying Markov's inequality, we get:

$$\begin{aligned}
\mathbb{P}^*\left(\sum_{k\leq t} \left|\log \frac{f(s_k|\alpha,\theta_0)}{f(s_k|\alpha,\theta)}\right| \geq \log A(R_\theta - 1)\right) &\leq \frac{t \cdot \mathbb{E}^* \left|\log \frac{f(s|\alpha,\theta_0)}{f(s|\alpha,\theta)}\right|}{\log A(R_\theta - 1)} \\
&\leq \frac{T\left(\mathbb{E}^*|\log f(s|\alpha,\theta_0)| + \mathbb{E}^* \log|f(s|\alpha,\theta)|\right)}{\log A(R_\theta - 1)} \\
&\leq \frac{Tb}{\log A(R_\theta - 1)},
\end{aligned}$$

for some constant $b \geq 0$, where the last inequality comes from the fact that $\mathbb{E}^* |\log f(s|\alpha, \theta_0)|$ is uniformly bounded (implied by the Assumption 2 and Jensen's inequality). Therefore, we have:

$$\mathbb{P}^* (\tau < T) \leq \frac{T^2 N b}{\log A (R_\theta - 1)} \to 0 \quad \text{as } R_\theta \to \infty$$

□

Based on the lemmas above, Proposition 6 can be proved as follows.

Proof. (i) Consider first the case where $\theta \in \mathcal{U}_{\mathcal{A}}^\theta$. Since every model has a unique zero-potential state, we have $\mathcal{U}_{\mathcal{A}}^\theta = \{\theta\}$. Let α be the model such that $r(\alpha, \theta) = 0$. From the fact that $\mu_t(.|\alpha) \to \delta_\theta$, for all $\varepsilon, \delta \in (0, 1)$, there exists some T such that

$$\mathbb{P}^* (\mu_t(.|\alpha) \in B_\varepsilon(\delta_\theta) \text{ for all } t \geq T) \geq 1 - \delta^2,$$

Similar to the previous proof, I choose the ε to be sufficiently small such that $B_\varepsilon(\delta_\theta)$ is contained in the interior of $\Delta_\theta \backslash \cup_{\theta' \neq \theta} \Delta_{\theta'}$. From Lemma 8, there exists some $R < \infty$ such that when $R_\theta \geq R$, we have:

$$\mathbb{P}^* (\{\mu_t(.|\alpha) \in B_\varepsilon(\delta_\theta) \text{ for all } t \geq T\} \cap \{\tau \geq T + 1\}) \geq 1 - \delta \tag{30}$$

Denote by E_1 the event inside the $\mathbb{P}^*$ operator on the LHS of (30). For all $s^\infty \in E_1$, since $\tau(s^\infty) > T$, we must have $\mu_T^\theta \in \Delta_\theta$, so $b_T^\theta = \theta$. Since $\mu_t(.|\alpha) \in B_\varepsilon(\delta_\theta)$ for all $t \geq T$, we have:

$$\mu_{T+1}^\theta(\theta) \geq \max_{\alpha \in \mathcal{A}} \mu_{T+1}(\theta|\alpha) > 1 - \varepsilon,$$

so $\mu_{T+1}^\theta \in B_\varepsilon(\delta_\theta)$, thus $b_{T+1}^\theta = \theta$. The same reasoning applies for all $t \geq T + 1$, so we have μ_t^θ remains in the small ball $B_\varepsilon(\delta_\theta)$. Since μ_t^θ converges to a Dirac belief almost surely, we must have for almost every signal path s^∞ in E_1, we have $\mu_t^\theta \to \delta_\theta$. Therefore, $\mathbb{P}^* \left(\mu_t^\theta \to \delta_\theta\right) \geq 1 - \delta$ whenever $R_\theta \geq R$. Since δ can be chosen arbitrarily, we must have $\lim_{R_\theta \to \infty} \mathbb{P}^* \left(\mu_t^\theta \to \delta_\theta\right) = 1$.

(ii) Suppose that $\theta \notin \mathcal{U}_{\mathcal{A}}^\theta$. The idea of the proof is similar. Denote by $\hat{\mu}_t^\theta$ the belief of the individual with bias θ as in the benchmark model (i.e., this individual has a fixed bias θ). Theorem 1 implies that for all $\varepsilon, \delta \in (0, 1)$, there exists some T such that

$$\sum_{\theta' \in \mathcal{U}_{\mathcal{A}}^\theta} \mathbb{P}^* \left(\hat{\mu}_t^\theta \in B_\varepsilon(\delta_{\theta'}) \text{ for all } t \geq T\right) \geq 1 - \delta^2 \tag{31}$$

Similarly, I choose the ε to be sufficiently small such that each $B_\varepsilon(\delta_{\theta'})$ is contained in the interior of $\Delta_{\theta'} \backslash \cup_{\theta'' \neq \theta', \theta} \Delta_{\theta''}$. From Lemma 8, there exists some $R < \infty$ such that when $R_\theta \geq R$, we have:

$$\sum_{\theta' \in \mathcal{U}_{\mathcal{A}}^\theta} \mathbb{P}^* \left(\left\{\hat{\mu}_t^\theta \in B_\varepsilon(\delta_{\theta'}) \text{ for all } t \geq T\right\} \cap \{\tau \geq T + 1\}\right) \geq 1 - \delta$$

Since $\theta \notin \mathcal{U}_{\mathcal{A}}^\theta$, we must have $\mathbb{P}^* (\tau < \infty) = 1$. This is because if $\{\tau = \infty\}$ occurs with a positive probability, then we have $\hat{\mu}_t^\theta = \mu_t^\theta$ for all t with a positive probability. If that is the case, we must have $\mu_t^\theta = \hat{\mu}_t^\theta \to \delta_{\theta'}$ for some $\theta' \in \mathcal{U}_\mathcal{A}^\theta$ with a positive probability. However, since $\theta \notin \mathcal{U}_\mathcal{A}^\theta$, $\mu_t^\theta(\theta)$ must escape Δ_θ in finite time, which is a contradiction. For each $\theta' \in \mathcal{U}_\mathcal{A}^\theta$, denote by $E_{\theta'}$ the event inside the $\mathbb{P}^*$ operator above. For all $s^\infty \in E_{\theta'}$ it is a routine to verify that $\mu_{\tau+k}^\theta \in B_\varepsilon(\delta_{\theta'})$ for all $k \geq 1$. From the facts that (i) μ_t^θ almost surely converges to some Dirac measure, and (ii) $\tau < \infty$, we must have $\mu_t^\theta \to \delta_{\theta'}$ for all signal paths in $E_{\theta'}$ (expect for null events), so

$$\sum_{\theta' \in \mathcal{U}_\mathcal{A}^\theta} \mathbb{P}^*\left(\mu_t^\theta \to \delta_{\theta'}\right) \geq 1 - \delta.$$

Since δ is chosen arbitrarily between 0 and 1, we must have $\sum_{\theta' \in \mathcal{U}_\mathcal{A}^\theta} \mathbb{P}^*\left(\mu_t^\theta \to \delta_{\theta'}\right) \to 1$ as $R_\theta \to \infty$. □