Biased Information and Opinion Polarisation*

Richard M. H. Suen†

First Version: 26 May, 2025.

This Version: 30 June, 2026.

Abstract

Why do people form polarised opinions after receiving the same information? Why does disagreement persist even when public information is abundant? This paper presents a tractable Bayesian learning model that can address these questions. Agents in our model interpret the same signals differently because they have different perceptions about the biasedness of the signal source. These perceptions will evolve over time and interfere how the agents learn from the signals. We show that persistent disagreement and opinion polarisation can readily emerge among fully rational agents when they are uncertain and disagree about the bias in the commonly observed signals. Exposure to these signals can in some cases amplify the initial disagreement and lead the agents further away from the truth.

Keywords: Bayesian Learning, Biased Signals, Disagreement, Opinion Polarisation.

JEL Classification: C11, D83

---

[*] We thank two anonymous referees and the Editor for helpful comments and suggestions.

[†] School of Economics, University of Leicester, Leicester LE1 7RH, United Kingdom. Email: *mhs15@le.ac.uk*

1 Introduction

Polarisation of opinion is ubiquitous in public discourse. On a wide range of issues from vaccination, immigration to climate change, people form divergent views even when public information is abundant and readily available.[1] When predicting economic conditions, business executives and professional forecasters also tend to disagree despite access to the same economic data (Kandel and Zilberfarb, 1999; Lahiri and Sheng, 2008).

These observations pose a challenge to the canonical Bayesian learning model which contends that persistent disagreement and opinion polarisation will not occur when rational agents observe the same information.[2] The current study proposes a tractable variant of the Bayesian model that can reconcile this conflict between theory and observations. Agents in our model interpret the same signals differently because they have different perceptions about the biasedness of the signal source. These perceptions will evolve over time and in turn affect how they learn from the signals. This modelling approach is motivated by the documented pervasiveness of biased reporting in mass media and a dwindling trust in them from the public.[3] Our main finding is that persistent disagreement and opinion polarisation can readily emerge among fully rational agents when they are uncertain and disagree about the bias in the commonly observed signals. The model developed here can be applied to economic and political contexts that involve opinion aggregation and collective decision-making.[4]

Our analysis is based upon a prototypical Bayesian learning model in which fully rational economic agents seek to infer an unknown state of the world (a payoff-relevant parameter) from a stream of noisy signals. We depart from this model by assuming that the signals are not only fraught with errors, they are also *potentially* distorted by an unknown factor which we refer to as *bias*. We use the term "bias" in a broad sense to include any systemic distortions that are

---

[1] Recent Gallup poll shows that public opinions on the importance of childhood vaccinations remain deeply divided in the United States (Jones, 2024), despite years of research and campaigns advocating their safety and efficiency. Similarly, public perception on migrant criminality remains polarised, even though evidence shows that immigrants commit fewer crimes than the domestically-born population (Abramitzky *et al.*, 2024; Marie and Pinotti, 2024). The cause of climate change has long been a contentious issue. According to a recent survey (Leiserowitz *et al.* 2026), 58% of Americans think that global warming is mainly driven by human activities, while 29% think that it is due to natural causes.

[2] Throughout this paper, we use the terms "opinion polarisation" and "opinion divergence" interchangeably to refer to situations in which two agents receive the same information but form opposite views, leading to greater disagreement.

[3] Puglisi and Snyder (2015) provide a comprehensive survery on the empirical evidence of biased reporting in traditional news media (newspapers and cable news). See Garz *et al.* (2020) for a more recent study. Empirical evidence showing a receding confidence in mass media can be found in Brenan (2024) based on U.S. sample and Newman *et al.* (2024) based on a large group of countries.

[4] One example is voting in political elections. Two recent studies, Aytimur and Suen (2025) and Little *et al.* (2025), examine electoral competition between two political parties when voters receive potentially biased information.

embedded in the signal-generating process, but cannot be directly observed by the agents. The word "potentially" underscores the agents' uncertainty about the extent of the distortion. It is this uncertainty, rather than the actual bias *per se*, that drives our main results. The ideas of potentially biased signals are best illustrated by the example of news reporting in mass media. The content of news reports often contains factual errors or other forms of unintended mistakes that are very costly (if not impossible) to eradicate. In terms of modelling, these errors are typically formulated as an independent and identically distributed "noise" process, which can be averaged out in a sufficiently large sample. But what makes into the news in the first place is determined by a multitude of factors (editor's political views, self-censorship, viewership concerns, pressures from sponsors and owner etc.) that permeate the news creation process and affect its content. Unlike the noises, this type of systemic distortions cannot be easily filtered out through averaging. Readers and viewers are conscious of these biases, though they may not know or agree on the actual extent. Crucially, their perceptions about the media outlet will influence how they interpret the news generated by it.[5] In our model, this perception is captured by the agents' belief about the bias in the signals.

Examples of systemic distortions in information abound in other contexts.[6] The focus of this paper is not on the reasons why news or other forms of information are biased. Our primary goal is to analyse the implications of potentially biased signals on the Bayesian learning process. The premise of our approach is different from that of the cognitive bias literature in behavioural economics [see, for instance, Barberis (2018) and Benjamin (2019) for extensive reviews]. This literature focuses on the systematic learning errors that arise due to the learners' own cognitive limits or biases (overconfidence, correlation neglect, motivated reasoning etc.). Benjamin (2019, p.70) defines bias as "any deviation from correct reasoning about probabilities or Bayesian updating". In contrast, we maintain the assumption that economic agents are fully rational and use correct Bayesian updating.

We model systemic bias as a separate additive term embedded in each realisation of the signal and make three crucial assumptions. First, the magnitude of the bias term is unknown. Second, agents form an initial subjective belief about the bias, just as they do for the unobserved state

---

[5] Transparency, fairness and neutrality are among the key factors that determine the trustworthiness of a news outlet [Newman *et al.* (2024, p.36)]. People lose confidence in a news outlet when they perceive it as taking sides, withholding information, or being controlled by powerful people to promote their own interests rather than those of the general public [Newman and Fletcher (2017)].

[6] For instance, systemic bias may arise in any kind of performance evaluation due to the evaluator's bias and value judgement, or due to a flawed evaluation process which favours certain group of evaluatees. More generally, any ill-devised sampling or data collection method may lead to systemic distortions in measurements.

of the world. Upon the arrival of new information, agents update their beliefs about the two unknowns jointly using Bayes' rule. As a result, agents' perception about the signal source is also changing over time.[7] Third, there is a lack of consensus among the agents which leads to disagreement in their initial beliefs. Throughout the paper, we confine our attention to Gaussian signal process which is one of the most commonly used specifications in economics and related fields.[8] Using a conjugate bivariate normal distribution as prior, the model admits an exact closed-form solution for all subsequent revised beliefs, including the limit. This tractability allows us to derive sharp results and explain clearly the various mechanisms that lead to opinion polarisation. We explore this model at three different levels: individual, interpersonal and aggregate levels. In Section 2, we first characterise the belief formation process for an individual learner. In Section 3, we compare two learners who observe the same sequence of signals but have different initial beliefs. Here we examine the conditions under which polarised opinions about the hidden state emerge in the long run. Finally, in Section 4, we extend our analysis to a large population of agents with heterogeneous initial beliefs. The primary concern here is how disagreement about the systemic bias will affect long-term disagreement about the hidden state within the population.

Before presenting our main results, we first recall the key implications of unbiased signals as a point of reference. Consider an alternate environment in which all the signals are truly unbiased, and this is *common knowledge* among the agents. This means they all accept that the bias term is identical to zero and there is *no uncertainty* about it.[9] We refer to this as the conventional model. With Gaussian signals, the conventional model has three main predictions that are most relevant to the current study. First, after observing a sufficiently long stream of signals, a Bayesian learner will be able to infer the true value of the hidden state.[10] Second, any initial disagreement across agents will eventually disappear, hence there is no room for persistent disagreement.[11] Third, in the short run, disagreement may remain but polarisation will never occur.[12] More specifically,

---

[7] When applied to the context of news consumption, this is in line with the empirical evidence that people's perception about mass media is evolving over time.

[8] We also assume that there is only one source of public signal. Our model can be readily extended to allow for multiple public signals, each with an unknown bias term. We choose the single-signal version to convey the key messages in the most straightforward manner, without burdening the reader with excessive technical details.

[9] In terms of a bivariate-normal initial belief, this means the marginal distribution of the bias term has zero mean and zero variance (denoted by $\sigma^2_{b,0} = 0$). On the contrary, the signals in our model may be truly unbiased but this is not known with certainty so that $\sigma^2_{b,0} > 0$.

[10] This means the agent's limiting belief is degenerate at the true value of the hidden state. Hence, the agent's long-run estimate is both unbiased and consistent. This result remains valid if the hidden state has a finite number of possible values (i.e., a discrete random variable) but the noises are drawn from a continuous distribution [see DeGroot (1970, Section 10.5)]. In more general settings, the limiting distribution can be non-degenerate. See, for instance, Chamley (2004, Section 2.4).

[11] Blackwell and Dubins (1962) show that this result holds in a general environment under some mild conditions.

[12] See Bullock (2009, Proposition 3) for a formal statement of this result in the Gaussian model. This result holds in general if the likelihood function (i.e., the density function of the signal conditional on the hidden state) satisfies

two agents observing the same signal will never revise their estimates in opposite directions and move further apart.

These results are no longer guaranteed once we introduce an unknown bias. Uncertainty about the bias influences how agents interpret the signals and update their beliefs. This can give rise to some unconventional learning dynamics which are described in the next paragraph. In the long run, each agent's belief converges to a distribution which is non-degenerate and dependent on the initial belief. Except for some knife-edge special cases, agents will not be able to identify the true value of the hidden state. Instead, they can only identify the *sum* of the hidden state and the bias. Introducing an unknown bias in the signals therefore creates an identification problem which results in biased and inconsistent long-run estimates.[13] When applied to interpersonal comparisons, we show that agents with different initial beliefs may form drastically different long-run estimates. In some cases, exposure to the same signals can drive the agents' opinions further away from each other and also further away from the truth. These effects can arise even when the signals are truly unbiased; the key factor is the agents' initial belief about the bias term, which captures their perceptions about the signal-generating process.

An immediate implication of these results is that any initial disagreement among the agents will persist in the long run. Opinion polarisation, on the other hand, is less straightforward. In order to explain how polarisation can emerge, we first describe two other predictions of our model that are not possible in the conventional one. We label these as **defiant learning** and **misguided learning.**[14] Defiant learning happens when an agent revises her estimate for the hidden state in the *opposite* direction as suggested by the signals. This cannot happen in the conventional model because the signals therein are always positively correlated with the hidden state. Therefore, any higher-than-expected signal indicates to the agent that her initial estimate is too low, which motivates an upward revision. As a result, agents in the conventional model will always revise their estimates in the same direction as suggested by the signals. We refer to these as conventional learners. On the contrary, an agent in our model may perceive the potentially biased signals as negatively correlated with the hidden state. This happens when the agent believes that

---

the monotone likelihood ratio property. Baliga *et al.* (2013, Theorem 1) establish this result for the case when the hidden state is a discrete random variable.

[13] Andreoni and Mylovanov (2012) and Acemoglu *et al.* (2016) present other settings in which identification problem can emerge in the Bayesian learning process. We will discuss these papers briefly in the "Related Literature" part of the Introduction. There is also an extensive literature in statistics and econometrics that study the issue of partial identification [see, for instance, Kline and Tamer (2023) for a recent review]. The focus of this literature, however, is different from ours.

[14] Our model also predicts a third type of unconventional learning dynamics which we refer to as opinion reversal. Details of this can be found in Section 2.4 and Section 3. We focus on defiant learning and misguided learning here because they are more directly related to opinion polarisation.

the hidden state and the unknown bias are strongly negatively correlated. To see how this can lead to defiant learning, consider the following example: Suppose an agent reads a news report about the performance of a new vaccine which exceeds her initial expectation. However, due to a lack of confidence on the news source, the agent suspects that the report is biased and the bias is negatively correlated with the vaccine's true performance.[15] Defiant learning happens if the agent infers from the positive news a higher-than-expected bias term. Since the latter tends to happen when the vaccine's true performance is lower than expected, this motivates the agent to revise her expectation downward. We refer to such agent as a defiant learner.

Misguided learning, on the other hand, happens when an agent's long-run estimate for the hidden state is further away from the true value than her initial estimate. In other words, the agent is led further astray from the truth after being exposed to the signals. In general, misguided learning happens when an agent overestimates [resp., underestimates] the hidden state initially and subsequent information leads her to revise her estimates upward [resp., downward] more often than otherwise. This type of learning outcome cannot happen in the conventional model because any initial misjudgment will be corrected subsequently. This self-correcting mechanism is disrupted in our model due to the agents' uncertainty about the bias term, which makes misguided learning possible.

Both defiant learning and misguided learning are relevant to permanent opinion polarisation, i.e., divergence in two agents' long-run estimates for the hidden state. First, if one agent is a conventional learner and the other is a defiant learner, then they will always respond to the same signals in opposite directions. Second, if both agents are conventional learners, or both are defiant learners, then polarisation happens when the misguided learning mechanism works on both of them but in opposite directions. In this case, instead of driving the agents further away from the truth, this mechanism drives their opinions further apart.

Finally, in the aggregate analysis, we consider a large population of agents with heterogeneous initial beliefs. The main question is whether aggregate disagreement about the hidden state will be reduced or exacerbated after the agents are exposed to the signals. Three lessons emerge from this analysis. First, holding other things constant, greater initial disagreement about the bias raises the extent of long-run disagreement. This is true even when there is no initial disagreement in the hidden state. In other words, differences in the perception about the signal-generating process can in itself lead to persistent disagreement. Second, any initial disagreement in the hidden state will

---

[15] For example, the agent may think that the news report (or the study in the news report) is sponsored and manipulated by the drug company in order to promote their commercial interests.

persist in the long run due to the agents' uncertainty about the bias term.[16] This is true even if the signals are truly unbiased. For a population of conventional learners, aggregate disagreement is reduced *ex post*, even though the signals are potentially distorted. But for a population of defiant learners, aggregate disagreement is exacerbated by the signals. Third, differences in the covariance matrix in the agents' initial beliefs will induce different responses to the same signals (e.g., conventional versus defiant learning), leading to aggregate disagreement in the long run.

Related Literature Several existing studies have explored the possibility of permanent disagreement and opinion polarisation within the Bayesian paradigm. Dixit and Weibull (2007) point out that opinion polarisation is not possible in the conventional model if the likelihood function exhibits monotone likelihood ratio (MLR) property. This observation motivates them to devise models where the MLR property does not hold. Short-term polarisation in the mean estimates can emerge under this approach, but not permanent polarisation.[17] Baliga *et al.* (2013) take a different approach: they assume that agents have ambiguity-averse preferences which give rise to a hedging motive when forming their posterior estimates. They show that opinion divergence can emerge between two Bayesian learners with sufficiently extreme and polarised initial beliefs. The current study is close in spirit to Andreoni and Mylovanov (2012) and Acemoglu *et al.* (2016) but differs substantially in details and findings. A common theme that threads through these studies is that disagreement arises because agents interpret the public signals differently. In the theoretical model of Andreoni and Mylovanov (2012), agents receive both public signals and idiosyncratic private signals, and how they interpret the former depends on the latter.[18] Acemoglu *et al.* (2016) show that long-run disagreement can readily emerge when agents are uncertain about the signal-generating process. They consider a general setup without specifying the reason for this uncertainty and they do not explore the possibility of opinion polarisation. Our study complements and extends this work in three substantive ways. First, uncertainty about the bias in our model can be seen as one reason why agents are unsure about the signal-generating process.[19] This provides a specific context under which disagreement will occur and links the the-

---

[16] On the contrary, there is no uncertainty about the bias term in the conventional model and any initial disagreement among the agents will disappear in the long run.

[17] See Dixit and Weibull (2007, p.7353) and the remarks made by Baliga *et al.* (2013, p.3081-3082).

[18] Kondor (2012) presents a financial trading model in which agents receive both public and private information about a fundamental (an unknown parameter). In this setup, agents have incentives to learn the private information of their trading partners. Public information can lead to greater disagreement (and more trading) in Kondor's model by increasing disagreement in higher-order expectations (i.e., expectations about other agents' expectations). This mechanism is not present in our model and the other studies reviewed here.

[19] There is a subtle difference between the two models. In ours, agents update their beliefs about the hidden state and the bias term simultaneously. Hence, their uncertainty about the signal-generating process is evolving

ory to the empirical evidence on biased information. Second, we demonstrate the possibility of both permanent disagreement and opinion polarisation in a tractable manner. Third, we present novel results, such as defiant learning and misguided learning, which can be of interests in other applications.

Other recent studies are similar to ours in certain ways, but their main interests are not about interpersonal disagreement and opinion polarisation. Heidhues *et al.* (2018) present a learning model in which misguided learning can happen if the learner has erroneous or mis-specified belief about the signal-generating process. In their model, the signals not only include an unobserved confounding factor (the agent's ability) but also an endogenous choice variable (action). Liang and Xu (2020) consider a model in which agents receive multiple sources of biased signals. Their main interest is on efficient information aggregation among the agents, rather than opinion polarisation. Bayesian learning models with biased signals have also appeared in the political science literature [see, for instance, Little and Pepinsky (2021), Aytimur and Suen (2025) and Little *et al.* (2025)].

2 Learning from Biased Signals

2.1 The Setup

Consider an agent who cares about an unobserved state $s \in \mathbb{R}$. In each time period $t \in \{1, 2, ...\}$, the agent receives a noisy and potentially biased signal m_t defined as

$$m_t = s + b + \varepsilon_t,$$

where $b \in \mathbb{R}$ is an unknown parameter that captures the inherent bias in the signal-generating process and $\{\varepsilon_t\}_{t=1}^{\infty}$ is a sequence of independent and identically distributed noises. Each ε_t is drawn from a normal distribution with mean zero and variance σ_ε^2. The statistical properties of $\{\varepsilon_t\}_{t=1}^{\infty}$ are known to the agent at the outset. While the agent's primary concern is on the unobserved state s, the distortions caused by the bias term cannot be ignored. Therefore, the agent forms an initial subjective belief on *both* s and b, and revise the joint distribution using Bayes' rule upon the arrival of new information. We maintain the standard assumption that the noise process is statistically independent from the agent's initial belief. The latter takes the form

---

over time. This mechanism is absent in Acemoglu *et al.* (2016).

of a bivariate normal distribution with mean vector $\mathbf{x}_0$ and covariance matrix $\mathbf{\Sigma}_0$ specified as

$$\mathbf{x}_0 = \begin{bmatrix} s_0 \\ b_0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{\Sigma}_0 = \begin{bmatrix} \sigma_{s,0}^2 & \omega_0 \\ \omega_0 & \sigma_{b,0}^2 \end{bmatrix}.$$

The notations $\sigma_{s,0}^2$, $\sigma_{b,0}^2$ and ω_0 denote, respectively, the variance of s, the variance of b and the covariance between the two in the initial belief. A higher value of $\sigma_{b,0}^2$ indicates that the agent is more uncertain about the bias term. The parameter $\sigma_{s,0}^2$ can be interpreted similarly. We also define ρ_0 as the corresponding correlation coefficient, so that $\omega_0 = \rho_0 \sigma_{s,0} \sigma_{b,0}$. Throughout the paper, we maintain the assumptions that $\sigma_{s,0} > 0$, $\sigma_{b,0} > 0$ and $\rho_0 \in (-1, 1)$. These guarantee that $\mathbf{\Sigma}_0$ is invertible and positive definite. A positive value of ρ_0 means that, in the agent's initial belief, the unknown bias tends to exaggerate or complement the effect of the hidden state, whereas a negative value means that it tends to contradict the effect of the hidden state. We note that a nonzero value of ρ_0 is not essential for generating opinion polarisation between agents.

Using the elements of $\mathbf{\Sigma}_0$, we can define two other moments that are crucial for subsequent analysis. First note that conditional on (s, b), the signals are independent and identically distributed over time. Therefore, before observing any signals (i.e., at $t = 0$), the agent's perceived covariances between (s, b) and any m_t, $t \geq 1$, are identical and given by

$$\lambda_0 \equiv Cov\,(s, m_t) = \sigma_{s,0}^2 + \omega_0 \qquad \text{and} \qquad \theta_0 \equiv Cov\,(b, m_t) = \sigma_{b,0}^2 + \omega_0. \tag{1}$$

When taken separately, λ_0 and θ_0 can be either positive or negative. Specifically,

$$\lambda_0 \gtrless 0 \qquad \text{if and only if} \qquad \rho_0 \gtrless -\frac{\sigma_{s,0}}{\sigma_{b,0}},$$

and

$$\theta_0 \gtrless 0 \qquad \text{if and only if} \qquad \rho_0 \gtrless -\frac{\sigma_{b,0}}{\sigma_{s,0}}.$$

Note that a negative value of ρ_0 is a necessary condition for either $\lambda_0 < 0$ or $\theta_0 < 0$. The sum of these two covariances, however, must be non-negative since $\lambda_0 + \theta_0 = var\,(s + b)$ in the agent's initial belief. This rules out the case when λ_0 and θ_0 are *both* negative. Using these notations, we can express the (unconditional) variance of each individual signal as

$$var\,(m_t) = \lambda_0 + \theta_0 + \sigma_\varepsilon^2.$$

Table 1 summarises the key notation introduced so far. We take these as the fundamentals of our model.

Table 1 Parameters in Initial Belief

Symbol	Meaning	Symbol	Meaning
s_0	Estimate of s	ω_0	Covariance between s and b
b_0	Estimate of b	ρ_0	Correlation between s and b
$\sigma^2_{s,0}$	Variance of s	λ_0	Covariance between s and m_t
$\sigma^2_{b,0}$	Variance of b	θ_0	Covariance between b and m_t

2.2 Closed-Form Solution for Updated Belief

After observing the signal in each period, the agent updates her belief about (s, b) using Bayes' rule. Let $\mathbf{x} = [s \quad b]'$ denote the true value of the unobservables and $\mathbf{m}^t = \{m_1, ..., m_t\}$ be a history of signals up to time t. Conditional on $\mathbf{m}^t$, the agent's revised belief takes the form of a bivariate normal distribution with mean vector

$$\widehat{\mathbf{x}}_t = \begin{bmatrix} \widehat{s}_t \\ \widehat{b}_t \end{bmatrix} \equiv E\left[\mathbf{x} \mid \mathbf{m}^t\right],$$

which contains the agent's updated estimate for $\mathbf{x}$, and covariance matrix

$$\widehat{\mathbf{\Sigma}}_t = \begin{bmatrix} \widehat{\sigma}^2_{s,t} & \widehat{\omega}_t \\ \widehat{\omega}_t & \widehat{\sigma}^2_{b,t} \end{bmatrix} \equiv E\left[\left(\mathbf{x} - \widehat{\mathbf{x}}_\mathbf{t}\right)\left(\mathbf{x} - \widehat{\mathbf{x}}_\mathbf{t}\right)' \mid \mathbf{m}^t\right].$$

The model admits a closed-form solution for the elements in $\widehat{\mathbf{x}}_t$ and $\widehat{\mathbf{\Sigma}}_t$ which is presented below.

Proposition 1 *Starting from the initial conditions $\widehat{\mathbf{x}}_0 = \mathbf{x}_0$ and $\widehat{\mathbf{\Sigma}}_0 = \mathbf{\Sigma}_0$, the elements of $\widehat{\mathbf{x}}_t$ and $\widehat{\mathbf{\Sigma}}_t$ at any time $t \geq 0$ are given by*

$$\widehat{\sigma}^2_{s,t} = \frac{\left[\sigma^2_\varepsilon + \left(1 - \rho^2_0\right)\sigma^2_{b,0}t\right]\sigma^2_{s,0}}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}, \tag{2}$$

$$\widehat{\sigma}^2_{b,t} = \frac{\left[\sigma^2_\varepsilon + \left(1 - \rho^2_0\right)\sigma^2_{s,0}t\right]\sigma^2_{b,0}}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}, \tag{3}$$

$$\widehat{\omega}_t = \frac{\omega_0\sigma^2_\varepsilon - \left(1 - \rho^2_0\right)\sigma^2_{s,0}\sigma^2_{b,0}t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}, \tag{4}$$

$$\widehat{s}_t = s_0 + \kappa_t \left(\overline{m}_t - s_0 - b_0\right), \tag{5}$$

$$\widehat{b}_t = b_0 + \eta_t \left(\overline{m}_t - s_0 - b_0\right), \tag{6}$$

where $\overline{m}_t \equiv \sum_{i=1}^{t} m_i/t$ *is the average value of the realised signals up to time* t,

$$\kappa_t \equiv \frac{\lambda_0 t}{\sigma_\varepsilon^2 + (\lambda_0 + \theta_0)\, t} \qquad \textit{and} \qquad \eta_t \equiv \frac{\theta_0 t}{\sigma_\varepsilon^2 + (\lambda_0 + \theta_0)\, t}. \tag{7}$$

Unless otherwise stated, all proofs can be found in the Appendix. Here we focus on the interpretation and intuition of the solution. Equations (2)-(4) form a stand-alone system that completely characterises the dynamics of $\widehat{\mathbf{\Sigma}}_t$. In particular, these equations are independent of the initial estimates $\mathbf{x}_0 = (s_0, b_0)$, the history of realised signals $\mathbf{m}^t$ and the sequence of revised estimates $\left\{\widehat{s}_1, \widehat{b}_1, ..., \widehat{s}_t, \widehat{b}_t\right\}$. This allows us to analyse the dynamics of $\widehat{\mathbf{\Sigma}}_t$ separately. Equations (5)-(6) describe how the agent updates the estimates for the hidden state and the unknown bias. This updating process will be explained fully after we present two other sets of results.

The first one concerns the perceived covariances between (s, b) and the signals in the revised belief. Similar to (1), these covariances are defined as

$$\widehat{\lambda}_t \equiv Cov_t\left(s, m_{t+k}\right) \qquad \text{and} \qquad \widehat{\theta}_t \equiv Cov_t\left(b, m_{t+k}\right),$$

where m_{t+k} is the signal in some future time period $t+k$, with $k \geq 1$. The time subscript in the covariance function indicates that it is derived from the revised belief at time t. These covariances are shown in Proposition 2. The key point to note is that the sign of $\widehat{\lambda}_t$ and $\widehat{\theta}_t$ are determined by that of λ_0 and θ_0, respectively. Therefore, if the perceived covariances between the signals and (s, b) are positive [resp., negative] before the signals are realised, they will remain positive [resp., negative] *ex post*. This highlights the importance of λ_0 and θ_0 in the learning process. As shown below, most of our main results are related to (the sign of) these two parameters.

Proposition 2 *At any time* $t \geq 1$, *after the agent revised her belief, the perceived covariances between* (s, b) *and the signals are given by*

$$\widehat{\lambda}_t = \widehat{\sigma}_{s,t}^2 + \widehat{\omega}_t \gtreqless 0 \qquad \textit{if and only if} \qquad \lambda_0 \gtreqless 0,$$

$$\widehat{\theta}_t = \widehat{\sigma}_{b,t}^2 + \widehat{\omega}_t \gtreqless 0 \qquad \textit{if and only if} \qquad \theta_0 \gtreqless 0.$$

Our next result establishes some of the asymptotic properties of the closed-form solution.

Proposition 3

(a) *Both* $\{\widehat{\sigma}_{s,t}^2\}_{t=0}^{\infty}$ *and* $\{\widehat{\sigma}_{b,t}^2\}_{t=0}^{\infty}$ *are monotonically decreasing sequences that converge to the same, strictly positive limit, i.e.,* $\lim_{t\to\infty}\widehat{\sigma}_{s,t}^2 = \lim_{t\to\infty}\widehat{\sigma}_{b,t}^2 > 0$.

(b) *The combined estimate* $\widehat{s}_t + \widehat{b}_t$ *converges in probability to the true sum* $s + b$.

(c) *Both* $\{\widehat{\omega}_t\}_{t=0}^{\infty}$ *and* $\{\widehat{\rho}_t\}_{t=0}^{\infty}$ *are convergent and the limits satisfy* $\lim_{t\to\infty}\widehat{\omega}_t < 0$ *and* $\lim_{t\to\infty}\widehat{\rho}_t = -1$.

Part (a) of Proposition 3 states that the precision of the agent's estimates (as measured by the reciprocal of $\widehat{\sigma}_{s,t}^2$ and $\widehat{\sigma}_{b,t}^2$) improves after each update. This means learning occurs even though the signals are distorted by the unknown bias. But instead of converging to zero, both $\widehat{\sigma}_{s,t}^2$ and $\widehat{\sigma}_{b,t}^2$ converge to a strictly positive value. This means the agent remains uncertain about the unknown parameters even after observing an infinite stream of signals. This highlights one of the main differences between the current model and the conventional model with known unbiased signals. In the latter, the true value of b is zero and this is known to the agent so that $b_0 = 0$ and $\sigma_{b,0} = 0$. Substituting $b = b_0 = \sigma_{b,0} = 0$ into the closed-form solution yields $\widehat{b}_t = \widehat{\sigma}_{b,t}^2 = \widehat{\omega}_t = 0$ for all t, and

$$\widehat{\sigma}_{s,t}^2 = \frac{\sigma_\varepsilon^2\sigma_{s,0}^2}{\sigma_\varepsilon^2 + \sigma_{s,0}^2 t}, \tag{8}$$

$$\widehat{s}_t = s_0 + \frac{\sigma_{s,0}^2 t}{\sigma_\varepsilon^2 + \sigma_{s,0}^2 t}\left(\overline{m}_t - s_0\right). \tag{9}$$

From these equations we can deduce that, as t approaches infinity, $\widehat{\sigma}_{s,t}^2$ will reduce to zero while $\widehat{s}_t$ converges to the true state.[20] This means after observing a sufficiently long stream of unbiased signals, the agent can filter out the noises and identify the hidden state. Part (a) of Proposition 3 shows that this well-known result from the conventional model is no longer valid once we introduce an unknown bias term.

Parts (b) and (c) of Proposition 3 are closely related. Even though the agent cannot identify the true value of s and b separately, she can still learn the true value of their sum. Specifically, the result in part (b) implies that, after observing a sufficiently long sequence of signals, the chance that $\widehat{s}_t + \widehat{b}_t$ is somewhat different from $s + b$ is arbitrarily small. An immediate implication is that, in the long run, the agent will have a good grasp of the value of $s + b$. Any new information

---

[20] This is the case because, by the law of large number, the sample mean $\overline{m}_t$ converges in probability to the true state s and the fraction $\sigma_{s,0}^2 t / \left(\sigma_\varepsilon^2 + \sigma_{s,0}^2 t\right)$ converges to unity.

that induces her to adjust $\widehat{s}_t$ upward will necessarily lead to a downward adjustment in $\widehat{b}_t$. Hence, the two must be *perfectly* negatively correlated in the long run which is established in part (c).

The rest of Section 2 concerns the dynamic and asymptotic properties of $\widehat{s}_t$ and $\widehat{b}_t$. Unlike equations (2)-(4), the dynamic system for these variables are contingent on the history of realised signals through the sufficient statistic $\overline{m}_t$. Note that the expected value of $\overline{m}_t$ under the agent's initial belief is $s_0 + b_0$. Hence, according to (5) and (6), the agent will revise her estimates for (s, b), either above or below the initial values (s_0, b_0), based on the unexpected component of the signals, which is $(\overline{m}_t - s_0 - b_0)$. The direction of the adjustment is governed by the sign of κ_t and η_t, which are in turn determined by the sign of λ_0 and θ_0, respectively.

Suppose at $t = 0$, the agent perceives the signals as positively correlated with the hidden state, i.e., $\lambda_0 > 0$. Then the coefficient κ_t is strictly positive for all $t \geq 1$. According to (5), this means in all subsequent periods the agent will adjust $\widehat{s}_t$ in the same direction as the unexpected component in the signals. Specifically, if the sufficient statistic $\overline{m}_t$ happens to be higher than expected, i.e., $\overline{m}_t > s_0 + b_0$, then she will interpret this as an indication that the true state s is greater than s_0. This in turn motivates the agent to revise her estimate upward so that $\widehat{s}_t > s_0$. We refer to this type of agent as *conventional learner.* Contrarily, if the agent presumes a negative correlation between the hidden state and the signals in her initial belief, i.e., $\lambda_0 < 0$, then θ_0 must be strictly positive (since λ_0 and θ_0 cannot both be negative). In this case, the agent will interpret the higher-than-expected $\overline{m}_t$ as an indication that the true bias b is greater than b_0. This not only induces the agent to revise her estimate for the bias upward, i.e., $\widehat{b}_t > b_0$, it will also induce a *downward* adjustment in the estimate for s. To see why this is the case, recall that a negative correlation between s and b (i.e., $\rho_0 < 0$) is a prerequisite for $\lambda_0 < 0$. Hence, an agent with $\lambda_0 < 0$ believes that the true state tends to be low when the bias is more severe. Consequently, such an agent will take a higher-than-expected $\overline{m}_t$ as an indication that the true state is lower than her initial estimate, leading to a downward revision, i.e., $\widehat{s}_t < s_0$. We refer to this type of agent as *defiant learner.* If $\lambda_0 = 0$, then that means the agent perceives the signals as uncorrelated with (and hence uninformative about) the hidden state. It follows that $\kappa_t = 0$ at all times and the agent will never revise her initial estimate. We will not consider this uninteresting case in the following analysis.

The above description presents another major difference between the current model and the conventional model. As equation (9) makes clear, an agent learning from unbiased signals will always adjust her estimate in the same direction as suggested by the signals, so that

$$\widehat{s}_t \gtrless s_0 \qquad \text{if and only if} \qquad \overline{m}_t \gtrless s_0.$$

Hence, defiant learning will never occur in the conventional model. But in our model, the adjustment can go in either direction depending on the sign of λ_0, so that

$$\left(\widehat{s}_t - s_0\right)\left(\overline{m}_t - s_0 - b_0\right) \gtrless 0 \qquad \text{if and only if} \qquad \lambda_0 \gtrless 0.$$

Hence, both conventional and defiant learning are possible.

As more and more signals are realised, the variance of the noises becomes less and less important in the learning process. This is evident from the limit of κ_t and η_t, which are

$$\kappa_\infty \equiv \lim_{t\to\infty} \kappa_t = \frac{\lambda_0}{\lambda_0 + \theta_0} \qquad \text{and} \qquad \eta_\infty \equiv \lim_{t\to\infty} \eta_t = \frac{\theta_0}{\lambda_0 + \theta_0} = 1 - \kappa_\infty.$$

Using these and the law of large numbers [which implies $\overline{m}_t \xrightarrow{p} (s+b)$], we can derive the asymptotic value of $\widehat{s}_t$ and $\widehat{b}_t$ from (5) and (6). The results are shown in Proposition 4.

Proposition 4 *The revised estimates $\widehat{s}_t$ and $\widehat{b}_t$ converge in probability to $\widehat{s}_\infty$ and $\widehat{b}_\infty$, respectively, where*

$$\widehat{s}_\infty \equiv s_0 + \kappa_\infty \left(s + b - s_0 - b_0\right), \tag{10}$$

$$\widehat{b}_\infty \equiv b_0 + \left(1 - \kappa_\infty\right)\left(s + b - s_0 - b_0\right). \tag{11}$$

Equations (10) and (11) together imply $\widehat{s}_\infty + \widehat{b}_\infty = s + b$, which is consistent with Proposition 3 part (b). The coefficient κ_∞ can be interpreted as follows: Since the agent's initial belief is independent of ε_t,

$$\lambda_0 \equiv Cov\left(s, m_t\right) = Cov\left(s, s+b\right).$$

Thus, κ_∞ indicates the contribution of $Cov\left(s, s+b\right)$ to the uncertainty about $s+b$ in the agent's initial belief, which is $var\left(s+b\right) = \lambda_0 + \theta_0$. Note that κ_∞ is negative if $\lambda_0 < 0$, and greater than one if $\theta_0 < 0$. It is bounded within $[0, 1]$ if and only if both λ_0 and θ_0 are non-negative.

Equations (10) and (11) encompass two types of learning dynamics that are incompatible with the conventional model. In the first one, learning from an infinite sequence of signals does not bring the agent any closer to the truth. Instead, the agent's long-run estimate for the hidden state is further away from the true value than her initial estimate, i.e., either $\widehat{s}_\infty > s_0 > s$ or $\widehat{s}_\infty < s_0 < s$. We refer to this as *misguided learning,* which is analysed in Section 2.3. In the second scenario, the relative position between the agent's estimate and the true value is reversed after learning, i.e., either $\widehat{s}_\infty > s > s_0$ or $\widehat{s}_\infty < s < s_0$. We refer to this as *opinion reversal,* which is the subject of Section 2.4. When opinion reversal happens, the agent can be either further away or closer to the truth after learning, i.e., both $|\widehat{s}_\infty - s| < |s - s_0|$ and $|\widehat{s}_\infty - s| > |s - s_0|$ are possible. We do not further distinguish between these two subcases.

2.3 Misguided Learning

Using (10), it is straightforward to see that misguided learning in the form of $\widehat{s}_\infty > s_0 > s$ happens if and only if

$$\kappa_\infty (s + b - s_0 - b_0) > 0 \qquad \text{and} \qquad s_0 > s. \tag{12}$$

Similarly, $\widehat{s}_\infty < s_0 < s$ is true if and only if

$$\kappa_\infty (s + b - s_0 - b_0) < 0 \qquad \text{and} \qquad s_0 < s. \tag{13}$$

The intuition behind (12) is as follows: Suppose the agent is a conventional learner, i.e., $\lambda_0 > 0$, so that $\kappa_\infty > 0$. Then (12) can be simplified to become $s + b > s_0 + b_0$ and $s_0 > s$, which is equivalent to $b - b_0 > s_0 - s > 0$. This means even though the agent overestimates the value of s in her initial belief, there is also a substantial underestimation in the bias term so that the initial combined estimate falls below the true sum. Over time as more signals are realised, the sample mean $\overline{m}_t$ will get closer and closer to $s + b$. This means the agent will almost always observe $\overline{m}_t > s_0 + b_0$ in the long run. According to (5), this will induce the conventional learner to maintain a long-run estimate for s that is higher than her initial assessment, hence $\widehat{s}_\infty > s_0$ is true.[21]

While a significant misalignment between b and b_0 is necessary for misguided learning to happen to a conventional learner, it is not necessary for a defiant learner. Suppose, unbeknownst to the agent, the signals are truly unbiased and the agent's initial estimate b_0 is correct, i.e.,

---

[21] Note that the initial condition $s_0 > s$ alone does not decide how the agent responds to the new information in the long run, which is controlled by $\kappa_\infty (s + b - s_0 - b_0)$. In general, if $\kappa_\infty (s + b - s_0 - b_0) > 0$, then $\widehat{s}_\infty > \widetilde{s}$ for any $\widetilde{s}$ that satisfies $s_0 \geq \widetilde{s}$. We exploit this argument when considering opinion divergence at the interpersonal level.

$b_0 = b = 0$. However, uncertainty exists so that $\sigma_{b,0} > 0$. Then $s_0 > s$ alone implies $s_0 + b_0 > s + b$, which means in the long run the agent will almost always observe $\overline{m}_t < s_0 + b_0$. Under defiant learning, the agent will revise her estimate for s in the opposite direction as suggested by the signals. Hence, she will persistently maintain a higher estimate for the hidden state than her initial estimate, leading to $\widehat{s}_\infty > s_0 > s$. On the other hand, even though both b_0 and b are identical to zero, subsequent estimates $\widehat{b}_t$ and the long-run value $\widehat{b}_\infty$ are in general different from zero. This means misperception about the signal-generating process is developed once the agent is exposed to the signals. This misperception persists in the long run and distorts the long-run estimate for the hidden state. The condition in (13) can be explained similarly.

We now provide some simulated examples to complement the above discussion. For the hidden parameters, we set $(s, b) = (0.2, 0)$ so that the signals are unbiased. In order to be consistent with the second inequality in (12), we set $s_0 = 0.3 > s$. In the initial covariance matrix, we set $\sigma_{s,0}^2 = 0.1$ and $\sigma_{b,0}^2 = 0.5$. The two key parameters are b_0 and ρ_0. For the correlation coefficient, we consider two possible values: $\rho_0 = 0$ and $\rho_0 = -0.65$. Under the stated value of $\sigma_{s,0}^2$ and $\sigma_{b,0}^2$, these correspond to $\lambda_0 = 0.1$ and $\lambda_0 = -0.045$, respectively. In each case, we consider four possible values of b_0, namely $\{-0.3, 0, 0.3, 0.6\}$. Figure 1 depicts the time paths of $\widehat{s}_t$ obtained under $\rho_0 = 0$, while Figure 2 depicts those obtained under $\rho_0 = -0.65$. All the time paths in these two diagrams share the same initial value $s_0 = 0.3$ and they are based on the same sequence of 100 independent error terms $\{\varepsilon_1, ..., \varepsilon_{100}\}$ drawn from the same distribution $N\left(0, \sigma_\varepsilon^2\right)$, with $\sigma_\varepsilon^2 = 0.03$. Table 2 summarises the parameter values and the resulting values of $\{\lambda_0, \theta_0, \kappa_\infty\}$ in the simulated examples.[22]

In Figure 1, misguided learning happens in the uppermost sequence which corresponds to the case when $b_0 = -0.3$, so that $s + b > s_0 + b_0 = 0$. The long-run estimate $\widehat{s}_\infty$ is one-third which confirms the ranking $\widehat{s}_\infty > s_0 > s$. For the second and third sequences (which correspond to $b_0 = 0$ and $b_0 = 0.3$, respectively), the first inequality in (12) is violated. As a result, $\widehat{s}_\infty$ is sandwiched between s_0 and s. Learning is incomplete in the sense that the long-run estimate is different from the true value, but misguided learning does not happen as $\widehat{s}_\infty$ is closer to the true value than the initial estimate. The fourth sequence in Figure 1 is an example of opinion reversal which we will discuss in the next section.

---

[22] The simulated results are robust to a wide range of parameter values, hence it is easy to construct other examples that can deliver the same messages. The MATLAB codes for generating the numerical results are available from the author's personal website.

Table 2 Parameter Values in Simulated Examples

	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5
s	0.2	0.2	0.3	0.3	0.3
b	0	0	0	0	0
s_0	0.3	0.3	0.2	0.2	0.2
b_0	$[-0.3, 0, 0.3, 0.6]$		-0.6	0	1.4
ρ_0	0	-0.65	0	-0.65	-0.65
$\sigma^2_{s,0}$	0.1	0.1	0.1	0.5	0.1
$\sigma^2_{b,0}$	0.5	0.5	0.5	0.1	0.5
σ^2_{ε}	0.03	0.03	0.03	0.03	0.03
λ_0	0.10	-0.045	0.10	0.355	-0.045
θ_0	0.50	0.355	0.50	-0.045	0.355
κ_∞	0.167	-0.147	0.167	1.147	-0.147

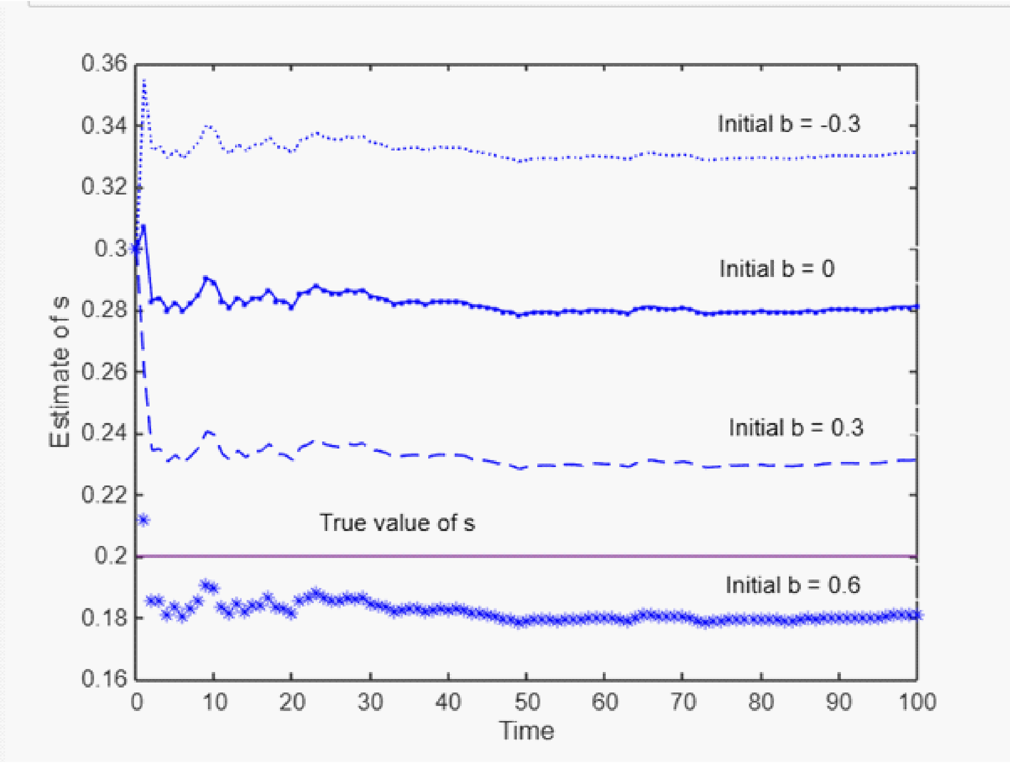

Figure 1

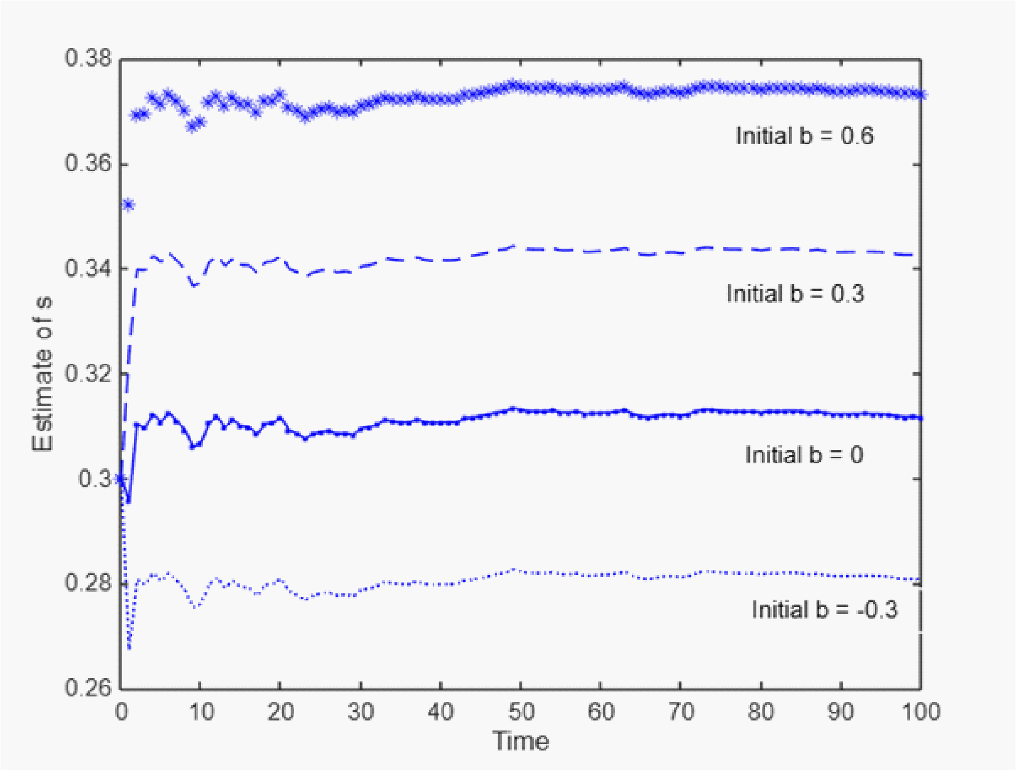

Figure 2

Figure 2 shows that when ρ_0 switches from zero to -0.65, the movement of $\widehat{s}_t$ and the ordering of the sequences are completely reversed. Misguided learning happens in the top three sequences where $\widehat{s}_\infty = 0.403$, 0.359 and 0.315, respectively. The third sequence is an example which shows that misguided learning can happen to a defiant learner even when $b = b_0 = 0$.

2.4 Opinion Reversal

According to (10), opinion reversal in the form of $\widehat{s}_\infty > s > s_0$ happens if and only if[23]

$$\kappa_\infty (s + b - s_0 - b_0) > s - s_0 > 0. \tag{14}$$

If both λ_0 and θ_0 are strictly positive so that $\kappa_\infty \in (0, 1)$, then (14) is equivalent to

$$b - b_0 > \frac{1 - \kappa_\infty}{\kappa_\infty} (s - s_0) > 0. \tag{15}$$

---

[23] Similarly, $\widehat{s}_\infty < s < s_0$ is true if and only if $\kappa_\infty (s + b - s_0 - b_0) < s - s_0 < 0$.

This has two meanings: First, the agent underestimates both s and b in her initial belief so that, in the long run, $\overline{m}_t$ is almost always greater than $s_0 + b_0$. Since κ_t and η_t are both positive for all t, the agent's revised estimates $\widehat{s}_t$ and $\widehat{b}_t$ will be persistently above the initial values, resulting in $\widehat{s}_\infty > s_0$ and $\widehat{b}_\infty > b_0$. This mechanism alone, however, does not guarantee that $\widehat{s}_\infty$ is higher than the true value. The inequalities in (15) also mean that the agent has significantly underestimated the bias term in her initial belief. This ensures that the long-run estimate $\widehat{b}_\infty$ is higher than b_0 but still falls short of the true value so that $b > \widehat{b}_\infty > b_0$. This, together with $\widehat{s}_\infty + \widehat{b}_\infty = s + b$, ensures $\widehat{s}_\infty > s$. A simulated example of this is shown in Figure 3 based on the parameter values listed in Table 2.[24] Another example is the bottom sequence in Figure 1 (which corresponds to $b_0 = 0.6$). In this case, opinion reversal happens in the form of $\widehat{s}_\infty < s < s_0$.

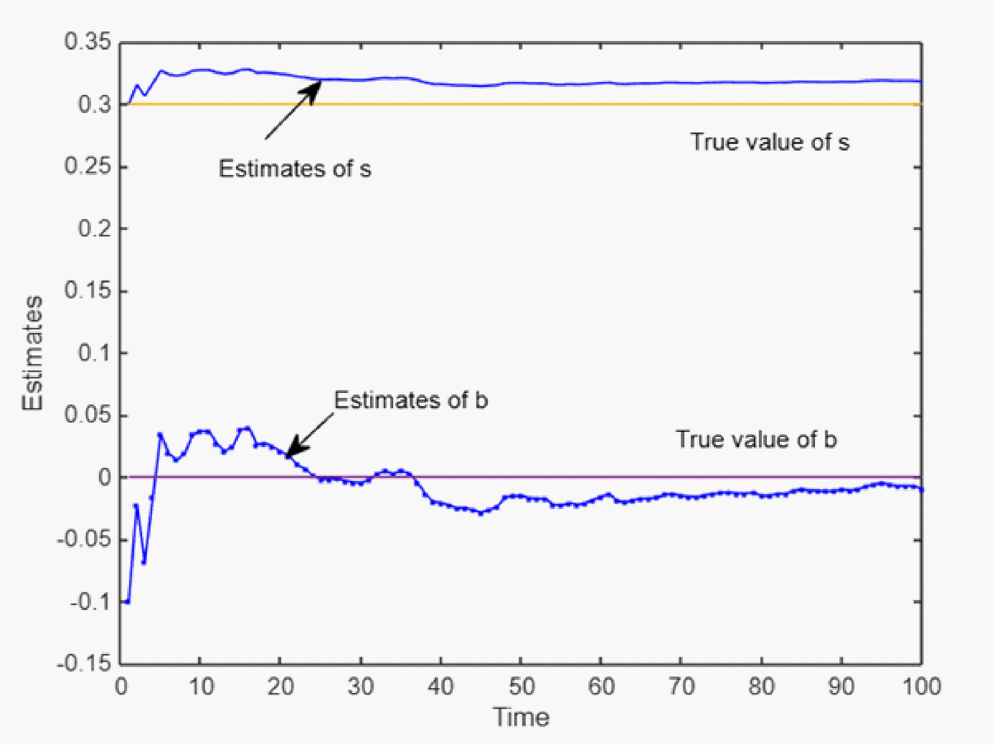

Figure 3

If $\lambda_0 > 0$ and $\theta_0 \leq 0$ so that $\kappa_\infty \geq 1$, then opinion reversal can happen even if $b_0 = b$. In this case, the agent will revise $\widehat{s}_t$ and $\widehat{b}_t$ in opposite directions in each period. In the long run, $\widehat{s}_\infty$ will stay above s_0 while $\widehat{b}_\infty$ falls below b_0, so that $\widehat{b}_\infty < b_0 = b$. This again ensures $\widehat{s}_\infty > s$. A

---

[24] We do not include the initial values (s_0, b_0) in Figures 3 and 5 because b_0 in each case is substantially different from the other values. Including this one point will dwarve the rest of the diagram, making it less comprehensible.

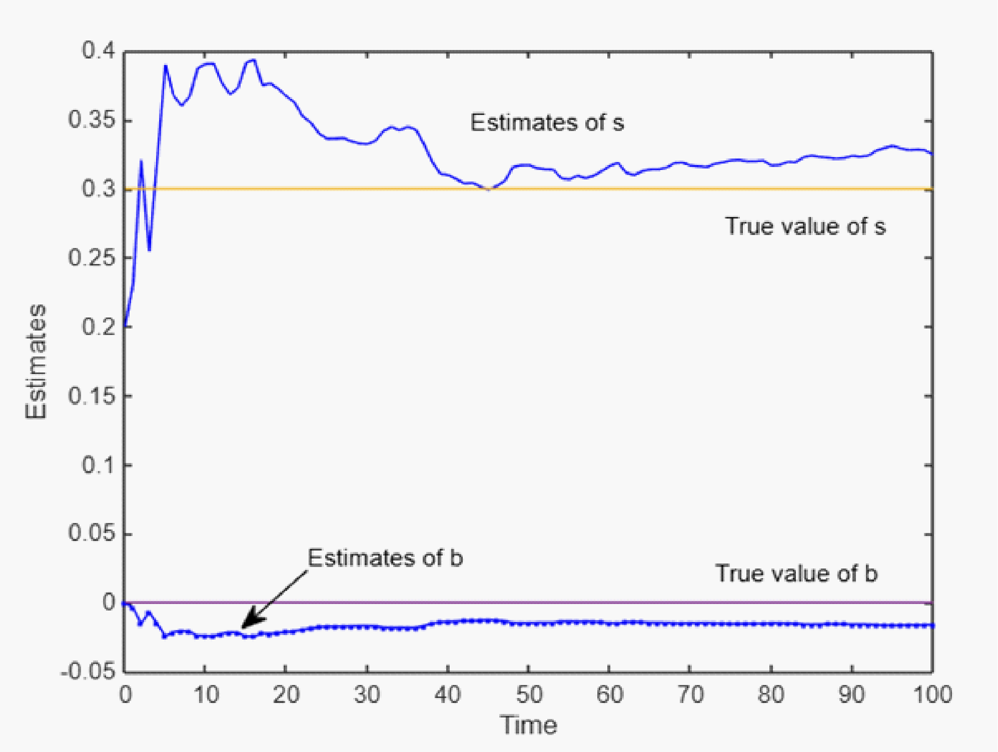

Figure 4

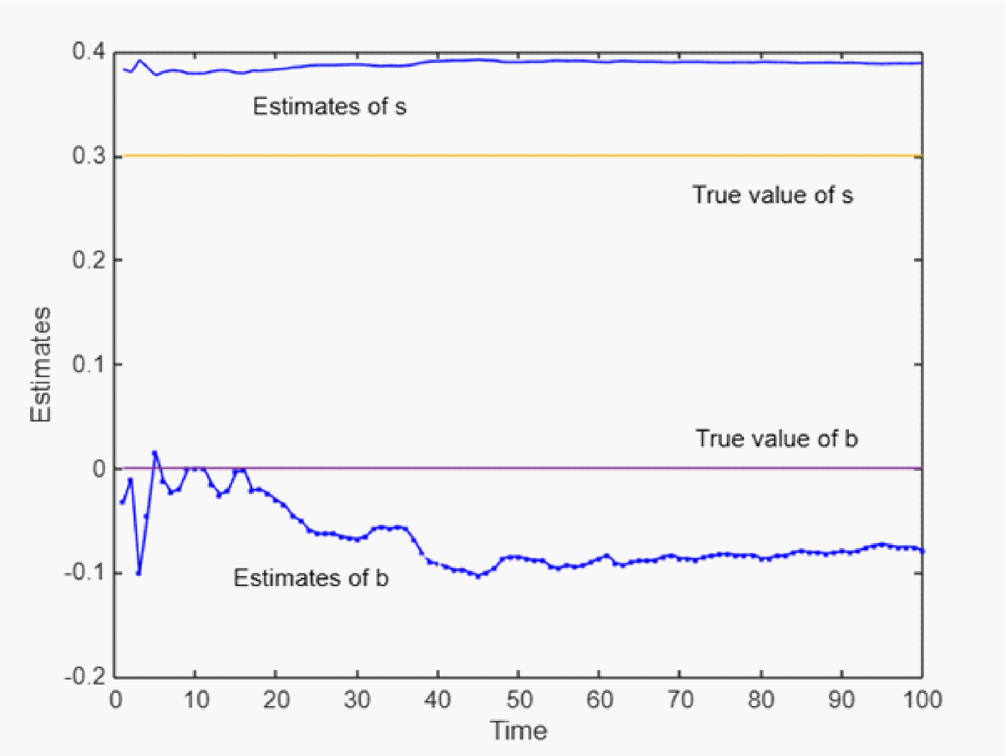

Figure 5

simulated example is shown in Figure 4. Finally, if $\lambda_0 < 0$ so that $\kappa_\infty < 0$ and $\eta_\infty = 1 - \kappa_\infty > 1$, then (14) implies

$$s + b < s_0 + b_0 \qquad \text{and} \qquad b - b_0 < \frac{1 - \kappa_\infty}{\kappa_\infty}(s - s_0) < 0.$$

The first inequality means that the agent will almost always observe $\overline{m}_t < s_0 + b_0$ in the long run. Since $\lambda_0 < 0$ and $\theta_0 > 0$, the agent will persistently maintain $\widehat{s}_t$ above s_0 due to defiant learning, and suppress $\widehat{b}_t$ below b_0 so that $\widehat{s}_\infty > s_0$ and $b_0 > \widehat{b}_\infty$. The fact that $\eta_\infty > 1$ means that the downward adjustment in $\widehat{b}_\infty$ is substantial, leading to an underestimation of b and an overestimation of s in the long run, i.e., $\widehat{b}_\infty < b$ and $\widehat{s}_\infty > s$. An example of this is shown in Figure 5.

We conclude Section 2 by summarising the main messages of Sections 2.3 and 2.4. When the agent is uncertain about the bias in the signals, misguided learning and opinion reversal (large swing in opinion) are possible for both conventional and defiant learners. These unconventional learning patterns can emerge even when the signals are truly unbiased. Misguided learning is problematic as it goes against the original intention of inferring the hidden state from the signals. For a conventional learner, this happens only if there is a significant misperception about the bias term in the initial belief (before observing the signals). But for a defiant learner, misguided learning can happen even if the agent has the correct initial estimate for the bias. In this case, misperception about the signal-generating process is developed only after the agent is exposed to the signals. This misperception persists in the long run, and in turn, distorts the long-run estimate for the hidden state.

3 Interpersonal Disagreement

We now compare the learning patterns of two individuals, referred to as Agent 1 and 2, in the above model. Both agents share the same knowledge about the signal-generating process and observe the same sequence of public signals. The two, however, have different initial beliefs about (s, b) which are given by $\mathbf{N}\left(\mathbf{x}_0^\dagger, \mathbf{\Sigma}_0^\dagger\right)$ and $\mathbf{N}\left(\mathbf{x}_0^\ddagger, \mathbf{\Sigma}_0^\ddagger\right)$, respectively. Upon the arrival of new information, both agents update their beliefs according to (2)-(6). Their long-run estimates for the hidden state are determined according to (10).

We will refer to $\widehat{s}_t^\dagger$ and $\widehat{s}_t^\ddagger$ as the agents' opinion about the hidden state after observing the same history of signals up to time t. Disagreement is said to occur at time t if $\widehat{s}_t^\dagger \neq \widehat{s}_t^\ddagger$. Given the initial differences in belief, disagreement is bound to happen except in the knife-edge case when

$$s_0^\dagger + \kappa_t^\dagger \left(\overline{m}_t - s_0^\dagger - b_0^\dagger \right) = s_0^\ddagger + \kappa_t^\ddagger \left(\overline{m}_t - s_0^\ddagger - b_0^\ddagger \right),$$

or in the long run

$$s_0^\dagger + \kappa_\infty^\dagger \left(s + b - s_0^\dagger - b_0^\dagger \right) = s_0^\ddagger + \kappa_\infty^\ddagger \left(s + b - s_0^\ddagger - b_0^\ddagger \right).$$

We will not indulge in these special cases. Our main interest is whether the initial disagreement will widen or reverse after the two agents observe an infinite stream of public signals.

Without loss of generality, assume $s_0^\dagger \geq s_0^\ddagger$. **Opinion divergence** or polarisation refers to a situation in which the two agents update their estimates in *opposite* directions and as a result disagreement widens over time, i.e.,

$$\widehat{s}_t^\dagger > s_0^\dagger \geq s_0^\ddagger > \widehat{s}_t^\ddagger, \qquad \text{for some } t \geq 1. \tag{16}$$

Permanent opinion divergence happens when the above inequalities hold in the long run, i.e.,

$$\widehat{s}_\infty^\dagger > s_0^\dagger \geq s_0^\ddagger > \widehat{s}_\infty^\ddagger. \tag{17}$$

The main difference between (16) and (17) is that in the former, both $\widehat{s}_t^\dagger$ and $\widehat{s}_t^\ddagger$ are random variables driven by the signals. Hence, the inequalities in (16) define a random event that happens with some probability. On the contrary, permanent divergence either happens or not, depending on the parameter values.

Opinion reversal refers to a situation in which the ordering between $s_0^\dagger$ and $s_0^\ddagger$ is reversed at some time $t \geq 1$, so that

$$s_0^\dagger - s_0^\ddagger \geq 0 > \widehat{s}_t^\dagger - \widehat{s}_t^\ddagger. \tag{18}$$

Permanent reversal happens when this condition holds in the long run.

We first consider two agents who share the same covariance matrix in their initial beliefs so that $\mathbf{\Sigma}_0^\dagger = \mathbf{\Sigma}_0^\ddagger = \mathbf{\Sigma}_0$. This means they share the same sequence of coefficients $\{\kappa_t\}_{t=1}^\infty$. Using (5), it can be shown that opinion divergence happens at time $t \geq 1$ if and only if[25]

$$\lambda_0 \left(s_0^{\ddagger} + b_0^{\ddagger} \right) > \lambda_0 \overline{m}_t > \lambda_0 \left(s_0^{\dagger} + b_0^{\dagger} \right). \tag{19}$$

If the agents are conventional learners, i.e., $\lambda_0 > 0$, then (19) can be reduced to $\left(s_0^{\ddagger} + b_0^{\ddagger} \right) > \overline{m}_t > \left(s_0^{\dagger} + b_0^{\dagger} \right)$. This means opinion divergence happens only when Agent 2 has a higher initial expectation for the signals than Agent 1. Given $s_0^{\dagger} \geq s_0^{\ddagger}$, this requires $b_0^{\ddagger} > b_0^{\dagger}$. Hence, for conventional learners, a significant difference in the agents' initial estimates for the hidden bias is necessary for opinion divergence to occur. Suppose $\left(s_0^{\ddagger} + b_0^{\ddagger} \right) > \overline{m}_t > \left(s_0^{\dagger} + b_0^{\dagger} \right)$ holds in some time period. Then by the reasoning described in Section 2.3, Agent 1 will adjust her initial estimate for the hidden state upward while Agent 2 revises hers downward, widening the initial disagreement (see Footnote 21). This shows that opinion divergence happens when the misguided learning mechanism described in Section 2.3 works on both agents but in *opposite* directions.

For defiant learners, a mismatch between $b_0^{\dagger}$ and $b_0^{\ddagger}$ is not necessary for opinion divergence to happen. To see this, suppose $s_0^{\dagger} > s_0^{\ddagger}$, $b_0^{\ddagger} = b_0^{\dagger}$ and $\lambda_0 < 0$. Then whenever $s_0^{\dagger} > \overline{m}_t > s_0^{\ddagger}$ holds, the misguided learning mechanism will direct Agent 1 to revise her initial estimate for s upward, and direct Agent 2 to do the opposite, again widening the initial disagreement.[26]

Permanent opinion divergence can be explained and characterised in the same way by replacing $\overline{m}_t$ with $(s + b)$ in the above discussion. As an illustration, consider the top and bottom sequences in Figure 1 which can be viewed as the estimates produced by two agents who share the same s_0 and $\lambda_0 > 0$, but have different values of b_0. In particular, Agent 1 with $b_0^{\dagger} = -0.3$ satisfies $s + b = 0.2 > s_0^{\dagger} + b_0^{\dagger} = 0$, while Agent 2 with a significantly higher value of b_0 ($b_0^{\ddagger} = 0.6$) satisfies $s_0^{\ddagger} + b_0^{\ddagger} = 0.9 > s + b$. Since both agents are conventional learners, initial disagreement in b_0 alone is enough to generate permanent divergence in the agents' long-run estimate for the hidden state.

Remark Before proceeding further, we make a remark about expected disagreement. Suppose Agent 1 and Agent 2 share the same coefficients $\{\kappa_t\}_{t=1}^{\infty}$ but have different initial estimates with $s_0^{\dagger} \geq s_0^{\ddagger}$. Let $E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]$ be the expected value of $\widehat{s}_t^{\ddagger}$ based on Agent 1's initial belief. Using (5), we

---

[25] From (5) we can get $\widehat{s}_t^{\dagger} - s_0^{\dagger} = \kappa_t \left(\overline{m}_t - s_0^{\dagger} - b_0^{\dagger} \right)$ and $\widehat{s}_t^{\ddagger} - s_0^{\ddagger} = \kappa_t \left(\overline{m}_t - s_0^{\ddagger} - b_0^{\ddagger} \right)$. Hence, $\widehat{s}_t^{\dagger} - s_0^{\dagger} \geq 0 \geq \widehat{s}_t^{\ddagger} - s_0^{\ddagger}$ if and only if $\kappa_t \left(s_0^{\ddagger} + b_0^{\ddagger} \right) \geq \kappa_t \overline{m}_t \geq \kappa_t \left(s_0^{\dagger} + b_0^{\dagger} \right)$, which can be simplified to become (19).

[26] Note that differences in s_0 alone are enough to generate differences in $\widehat{b}_t$ according to (6). In other words, even if the two agents share the same b_0, their perception about the signal-generating process will diverge once they start receiving the signals.

can derive

$$E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]=\left(1-\kappa_t\right)s_0^{\ddagger}+\kappa_t s_0^{\dagger}+\kappa_t\left(b_0^{\dagger}-b_0^{\ddagger}\right).$$

If $\kappa_t\in(0,1)$ and $b_0^{\dagger}=b_0^{\ddagger}$ (as in the conventional model), then $s_0^{\dagger}\geq E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]\geq s_0^{\ddagger}$, which is consistent with the "direction" and "undershooting" results in Nartik *et al.* (2021, p.169). This means, before any signals are realised, Agent 1 expects Agent 2's revised estimate at time t to move closer to her own initial estimate $\left(s_0^{\dagger}\right)$ but not surpass it. Similarly, Agent 2 will have the same expectation. In other words, before observing any signals, both agents expect that the initial disagreement will be reduced subsequently. However, κ_t can be negative or greater than one in the current model. If $\kappa_t<0$ and $b_0^{\dagger}=b_0^{\ddagger}$, then $E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]$ will fall below $s_0^{\ddagger}$, violating the direction result. If $\kappa_t>1$ and $b_0^{\dagger}=b_0^{\ddagger}$, $E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]$ will surpass $s_0^{\dagger}$, violating the undershooting result. Likewise, if $\kappa_t\in(0,1)$ but $b_0^{\dagger}\neq b_0^{\ddagger}$, then both $E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]>s_0^{\dagger}$ and $s_0^{\ddagger}>E^{\dagger}\left[\widehat{s}_t^{\ddagger}\right]$ are possible. This represents another difference between the conventional model and the current model. ■

Next, we turn to opinion reversal. The conditions for opinion reversal and opinion divergence differ in two material ways. First, the conditions for opinion reversal are independent of the realised signals. To see why this is the case, we first rewrite (5) as

$$\widehat{s}_t=s_0-\kappa_t\left(s_0+b_0\right)+\kappa_t\overline{m}_t, \tag{20}$$

where the last term captures the agent's response to the signals. It is clear that if two agents share the same coefficient κ_t and observe the same signals, then the difference in their revised estimates $\widehat{s}_t^{\dagger}-\widehat{s}_t^{\ddagger}$ must be independent of $\overline{m}_t$. It follows that opinion reversal happens at time $t\geq 1$ if and only if[27]

$$\kappa_t\left(b_0^{\dagger}-b_0^{\ddagger}\right)>\left(1-\kappa_t\right)\left(s_0^{\dagger}-s_0^{\ddagger}\right)\geq 0. \tag{21}$$

The second major difference between opinion divergence and opinion reversal is that once opinion reversal happens, it becomes permanent. In other words, opinion reversal can only happen once, if at all. This follows from the fact that the condition in (21) is deterministic and linear in κ_t. This result is formally stated in Proposition 5.

Proposition 5 *Consider two agents with initial beliefs* $\mathbf{N}\left(\mathbf{x}_0^{\dagger},\mathbf{\Sigma}_0\right)$ *and* $\mathbf{N}\left(\mathbf{x}_0^{\ddagger},\mathbf{\Sigma}_0\right)$, *where* $s_0^{\dagger}\geq s_0^{\ddagger}$. *If opinion reversal happens at some time* $t\geq 1$, *i.e.,* $\widehat{s}_t^{\ddagger}>\widehat{s}_t^{\dagger}$, *then* $\widehat{s}_{t+k}^{\ddagger}>\widehat{s}_{t+k}^{\dagger}$ *for all*

---

[27] Using (20), we can show that $\left(\widehat{s}_t^{\dagger}-\widehat{s}_t^{\ddagger}\right)<0$ if and only if $s_0^{\dagger}-\kappa_t\left(s_0^{\dagger}+b_0^{\dagger}\right)<s_0^{\ddagger}-\kappa_t\left(s_0^{\ddagger}+b_0^{\ddagger}\right)$. This, together with $s_0^{\dagger}-s_0^{\ddagger}\geq 0$, is equivalent to (21).

$k \geq 1$, *including* $\widehat{s}_{\infty}^{\ddagger} > \widehat{s}_{\infty}^{\dagger}$.

The intuition behind Proposition 5 can be explained as follows: From (20), we can get

$$\widehat{s}_t^{\dagger} - \widehat{s}_t^{\ddagger} = s_0^{\dagger} - s_0^{\ddagger} - \kappa_t \left(s_0^{\dagger} + b_0^{\dagger} - s_0^{\ddagger} - b_0^{\ddagger} \right).$$

This breaks down the differences in the agents' opinions into two parts. The first part is due to initial disagreement. The second part captures the differences incurred in the revision at time t. Suppose both agents are conventional learners so that $\kappa_t > 0$, and Agent 1 has a higher initial expectation for the signals than Agent 2 so that $s_0^{\dagger} + b_0^{\dagger} > s_0^{\ddagger} + b_0^{\ddagger}$. Then over time it is more likely to have $s_0^{\dagger} + b_0^{\dagger} > \overline{m}_t$ than $s_0^{\ddagger} + b_0^{\ddagger} > \overline{m}_t$. This, together with $\kappa_t > 0$, means that Agent 1 is more likely to revise her estimate downward than Agent 2. In addition, the magnitude of κ_t is growing over time whenever it is strictly positive.[28] Therefore, Agent 1 not only revises her estimate downward more frequently than Agent 2, the magnitude of those revisions also become larger over time. This explains why Agent 1 ends up having a lower long-term estimate than Agent 2. The case where $\kappa_t < 0$ can be explained similarly.

In the remaining part of this section, we consider two agents who share the same initial estimates but have different covariance matrices in their initial beliefs, i.e., $s_0^{\dagger} = s_0^{\ddagger} = s_0$ and $b_0^{\dagger} = b_0^{\ddagger} = b_0$, but $\mathbf{\Sigma}_0^{\dagger} \neq \mathbf{\Sigma}_0^{\ddagger}$.[29] Using (5), it is immediate to see that $\widehat{s}_t^{\dagger} > \widehat{s}_t^{\ddagger}$ happens if and only if

$$\left(\kappa_t^{\dagger} - \kappa_t^{\ddagger} \right) \left(\overline{m}_t - s_0 - b_0 \right) > 0.$$

The implications are as follows: Suppose at some time t, the summary statistic $\overline{m}_t$ is higher than the agents' initial expectation, i.e., $\overline{m}_t > s_0 + b_0$. Then there are three possible scenarios that can lead to $\widehat{s}_t^{\dagger} > \widehat{s}_t^{\ddagger}$. In the first one, both $\kappa_t^{\dagger}$ and $\kappa_t^{\ddagger}$ are positive but $\kappa_t^{\dagger} > \kappa_t^{\ddagger}$. In this case, both agents respond to the higher-than-expected summary statistic by revising their estimates upward, i.e., $\widehat{s}_t^{\dagger} > s_0$ and $\widehat{s}_t^{\ddagger} > s_0$, but Agent 1 is more responsive to the signals, hence she makes a larger upward revision than Agent 2. In the second scenario, both $\kappa_t^{\dagger}$ and $\kappa_t^{\ddagger}$ are negative but $\kappa_t^{\dagger}$ is closer to zero. In this case, both agents revise their opinions downward but this time Agent 1 is less responsive. In the final scenario, $\kappa_t^{\dagger}$ is positive and $\kappa_t^{\ddagger}$ is negative. It follows that Agent 1

---

[28] Differentiating the expression for κ_t in (7) with respect to time gives

$$\frac{d\kappa_t}{dt} = \frac{\lambda_0 \sigma_\varepsilon^2}{\left[\sigma_\varepsilon^2 + (\lambda_0 + \theta_0)\, t\right]^2} = \frac{\sigma_\varepsilon^2 \kappa_t}{\left[\sigma_\varepsilon^2 + (\lambda_0 + \theta_0)\, t\right] t}.$$

Hence κ_t is strictly increasing over time whenever it is strictly positive.

[29] Since the two agents start with the same s_0, the possibility of opinion reversal is precluded.

will revise her estimate upward while Agent 2 does the opposite, leading to opinion divergence. This is the only scenario in which the agents' opinions diverge. Hence, opinion divergence in the form of $\widehat{s}_t^\dagger > s_0 > \widehat{s}_t^\ddagger$ happens if and only if

$$\kappa_t^\dagger \left(\overline{m}_t - s_0 - b_0\right) > 0 > \kappa_t^\ddagger \left(\overline{m}_t - s_0 - b_0\right),$$

which can be further simplified to become

$$\lambda_0^\dagger \left(\overline{m}_t - s_0 - b_0\right) > 0 > \lambda_0^\ddagger \left(\overline{m}_t - s_0 - b_0\right).$$

Permanent divergence can be similarly explained and characterised by replacing $\overline{m}_t$ with $s + b$.

To summarise, opinion divergence can happen in three situations. First, if one agent is a conventional learner and the other is a defiant learner, then they will respond to the same signals in opposite directions as we have shown in Section 2. Second, if both agents are conventional learners, opinion divergence happens only if there is a significant disagreement in the hidden bias before the signals are realised. In other words, the two agents must have significantly different perception about the signal-generating process *ex ante*. Third, if both agents are defiant learners, then polarisation can emerge even if they share the same initial perception about the signal source. In this case, *ex ante* disagreement about the hidden state alone is enough to generate permanent opinion polarisation.

4 Aggregate Disagreement

We now extend the above analysis to a large population of agents who observe the same sequence of public signals but have different initial beliefs about (s, b). Agents update their beliefs according to (2)-(6) and their long-run estimates for the hidden state are determined by (10). Our main focus in this section is whether continuous exposure to the potentially biased signals will exacerbate disagreement at the aggregate level.

Let $var\left(s_0\right)$ and $var\left(b_0\right)$ denote, respectively, the variance of s_0 and b_0 in the cross-sectional distribution of initial beliefs.[30] These measure the dispersion of s_0 and b_0 within the population before the agents are exposed to the signals. To simplify the analysis, we assume that s_0 and b_0 are uncorrelated across agents. As in Section 3, we first consider the case in which agents share

---

[30] The exact distribution of (s_0, b_0) within the population is irrelevant for our analysis.

the same covariance matrix $\mathbf{\Sigma}_0$ in their initial beliefs, so that $\{\lambda_0, \theta_0, \kappa_\infty\}$ are the same across individuals. Based on (10), the dispersion of $\widehat{s}_\infty$ within the population is given by

$$var\left(\widehat{s}_\infty\right) = \left(1 - \kappa_\infty\right)^2 var\left(s_0\right) + \kappa_\infty^2 var\left(b_0\right). \tag{22}$$

This quantifies the extent of aggregate disagreement after exposure to the signals.

Equation (22) encompasses the conventional model as a special case. If the signals are truly unbiased and this is common knowledge within the population, there is no initial disagreement about the bias term so that $var\left(b_0\right) = 0$, and there is no uncertainty regarding b so that $\sigma_{b,0} = \theta_0 = 0$ and $\kappa_\infty = 1$ for all agents. Then, regardless of the value of $var\left(s_0\right)$, aggregate disagreement disappears in the long run, i.e., $var\left(\widehat{s}_\infty\right) = 0$, as all agents learn the true value of the hidden state.

Suppose now within the society all agents share the same initial estimate s_0 so that $var\left(s_0\right) = 0$, but they disagree on the biasedness of the signals so that $var\left(b_0\right) > 0$. Equation (22) then becomes

$$var\left(\widehat{s}_\infty\right) = \kappa_\infty^2 var\left(b_0\right) = \left(\frac{\lambda_0}{\lambda_0 + \theta_0}\right)^2 var\left(b_0\right).$$

Provided that $var\left(b_0\right) > 0$, long-term aggregate disagreement exists whenever $\lambda_0 \neq 0$. The main message is that aggregate disagreement can easily emerge when people disagree about the biasedness of the public signals. In this environment, persistent disagreement is the norm and consensus is a special case.

Now consider a different society in which (i) agents disagree about the hidden state *ex ante*, i.e., $var\left(s_0\right) > 0$, (ii) they share the same b_0 so that $var\left(b_0\right) = 0$ [could be $b = b_0 = 0$], but (iii) they are (equally) uncertain about the bias so that $\sigma_{b,0} > 0$. Equation (22) now becomes

$$var\left(\widehat{s}_\infty\right) = \left(1 - \kappa_\infty\right)^2 var\left(s_0\right).$$

Even though there is no initial disagreement in b_0, the agents' uncertainty about the unknown bias is enough to generate long-term aggregate disagreement because $\sigma_{b,0} > 0$ implies $\kappa_\infty \neq 1$. Learning from the public signals can help reduce the initial disagreement within the population, i.e., $var\left(\widehat{s}_\infty\right) < var\left(s_0\right)$, if and only if

$$\left(1 - \kappa_\infty\right)^2 < 1 \qquad \Leftrightarrow \qquad \kappa_\infty \equiv \frac{\lambda_0}{\lambda_0 + \theta_0} \in (0, 2).$$

A necessary condition for this is $\lambda_0 > 0$, which means all the agents are conventional learners. If instead, $\kappa_\infty < 0$ or $\kappa_\infty > 2$, then the public signals will exacerbate aggregate disagreement, i.e., $var\left(\widehat{s}_\infty\right) > var\left(s_0\right).$ Note that $\kappa_\infty < 0$ happens only when all the agents in the population are defiant learners.

In general, (22) implies that

$$var\left(\widehat{s}_\infty\right) \gtrless var\left(s_0\right) \qquad \text{iff} \qquad var\left(b_0\right) \gtrless \left(\frac{2-\kappa_\infty}{\kappa_\infty}\right) var\left(s_0\right).$$

A key message is that if the agents are defiant learners and hold diverse views about the bias term so that $var\left(b_0\right)$ is strictly positive (however small), then exposure to the potentially biased public signals will widen the initial disagreement within the population. This is true regardless of the extent of initial disagreement $var\left(s_0\right).$

As an illustration of these points, we construct a hypothetical population of 200 agents whose initial estimates $\left(s_0, b_0\right)$ are equally spaced on the unit circle centred at the true parameter values $(s, b) = (0.2, 0)$.[31] This circular distribution is depicted in Figures 6 and 7. Using the same values of $\left\{\sigma_{s,0}^2, \sigma_{b,0}^2, \sigma_\varepsilon^2, \rho_0\right\}$ as in Figures 1 and 2 (see Table 2), we compute the asymptotic values $\left(\widehat{s}_\infty, \widehat{b}_\infty\right)$ for each $\left(s_0, b_0\right)$. The resulting values are located on the straight line $L : \widehat{s}_\infty + \widehat{b}_\infty = s + b$ which cuts through the centre of the circle. The auxiliary dashed lines link the initial value $\left(s_0, b_0\right)$ to the corresponding long-run value $\left(\widehat{s}_\infty, \widehat{b}_\infty\right)$ for a selected sample of agents. Figure 6 shows the results obtained under $\rho_0 = 0$. The first thing to note is that the range of $\widehat{s}_\infty$ is smaller than that of s_0, which suggests a reduction in aggregate disagreement. This is confirmed by the cross-sectional variance which reduces from $var\left(s_0\right) = 0.505$ to $var\left(\widehat{s}_\infty\right) = 0.365$. Both misguided learning and opinion reversal are at play. The former happens for any agent whose initial estimates either (i) lie below the straight line L and satisfy $s_0 > s$ (e.g., point A) or (ii) lie above L and satisfy $s_0 < s$ (e.g., point B). In particular, the long-run estimate associated with point A is further to the right than its initial value, while point B's long-run estimate is further to the left. Opinion reversal, on the other hand, happens at the periphery of the circle. For instance, point C has a lower value of s_0 than point D but a higher value of $\widehat{s}_\infty$. Figure 7 shows the results obtained under $\rho_0 = -0.65$. In this case, the dispersion of opinions increases from $var\left(s_0\right) = 0.505$ to $var\left(\widehat{s}_\infty\right) = 0.675$. Similar to Figure 6, points A and B serve as examples of permanent opinion divergence, while C and D serve as examples of permanent opinion reversal.

---

[31] When projected onto the x-axis, the distribution of s_0 is not uniform since there will be more points clustered around the endpoints of the range.

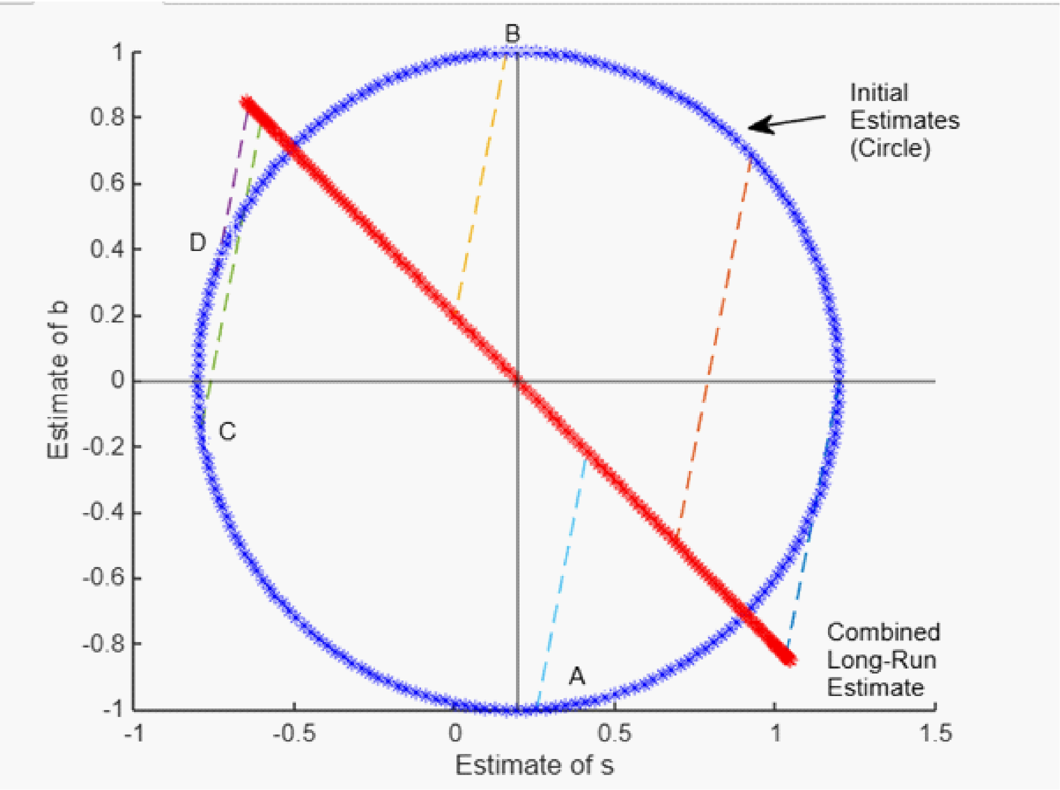

Figure 6

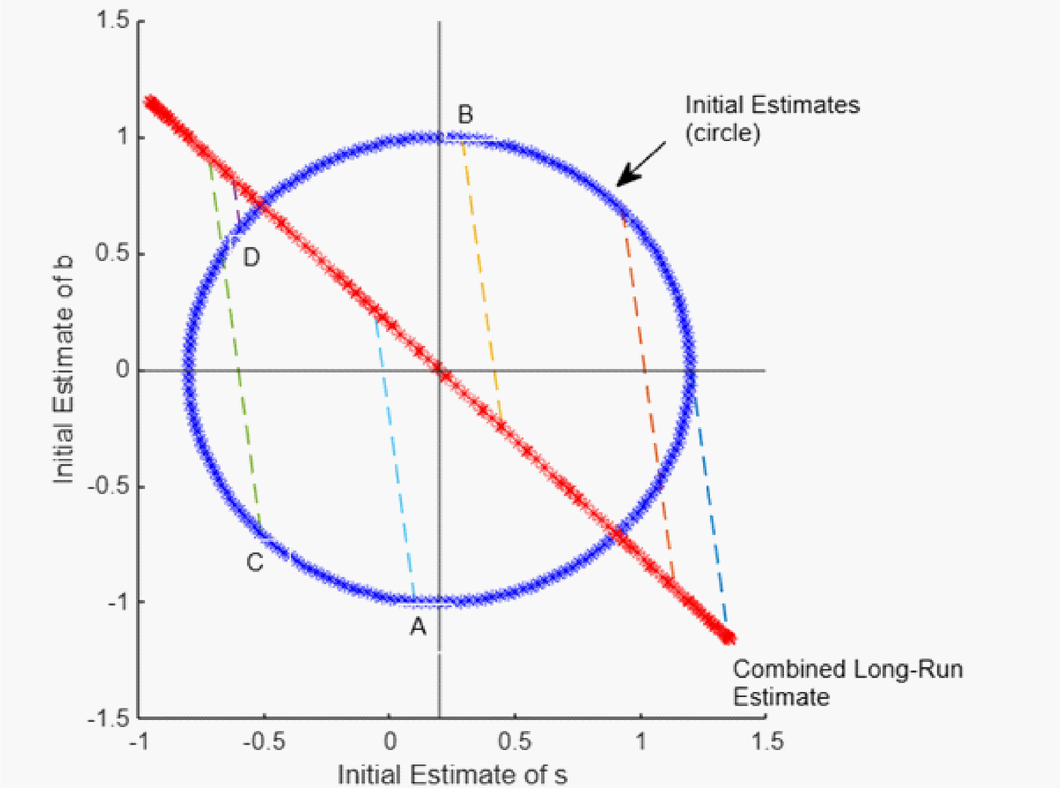

Figure 7

Finally, we consider a society in which all agents share the same initial estimates, i.e., $var\left(s_0\right) = var\left(b_0\right) = 0$, but have different covariance matrices in their initial beliefs. In particular, there is heterogeneity in κ_∞ so that $var\left(\kappa_\infty\right) > 0$ within the population. Equation (10) then implies

$$var\left(\widehat{s}_\infty\right) = var\left(\kappa_\infty\right)\left(s + b - s_0 - b_0\right)^2 .$$

If the agents' initial combined estimate is correct, i.e., $s_0 + b_0 = s + b$, then they will eventually reach a consensus about s, i.e., $var\left(\widehat{s}_\infty\right) = 0$, regardless of $var\left(\kappa_\infty\right)$. Does this mean that the agents have reached a consensus because they have learned the true value of s from the signals? Quite the opposite. Recall that agents' revised estimates are formed according to

$$\widehat{s}_t = s_0 + \kappa_t\left(\overline{m}_t - s_0 - b_0\right).$$

As t increases indefinitely, the sample average $\overline{m}_t$ converges to $s + b$. If $s_0 + b_0 = s + b$, then eventually the agents' estimates will respond less and less to the signals as there is no more surprises, i.e., $\overline{m}_t \xrightarrow{p} \left(s_0 + b_0\right)$. As a result, the cross-sectional differences in κ_t are irrelevant in the long run. The agents' estimate for s will then converge in probability to s_0, which means there is no learning at all.

Barring from this knife-edge case, heterogeneity in κ_∞ will lead to aggregate disagreement in $\widehat{s}_\infty$. In addition, the greater the discrepancy between $s+b$ and s_0+b_0, the greater the disagreement in $\widehat{s}_\infty$ at the aggregate level.

5 Concluding Remarks

This paper presents a variant of the Bayesian learning model that allows for an unknown bias in the public signals. We show that uncertainty and disagreement regarding this bias (a confounding factor) can result in persistent disagreement and opinion polarisation about the hidden state (the parameter of interest). Our model offers new insights as to how polarisation can emerge due to two novel mechanisms, namely defiant learning and misguided learning. When applied to the context of public discourse, the main message is that polarised views among fully rational agents can readily emerge when they have different perceptions about the biasedness of the mass media (the source of public signals). Our model is admittedly stylised, but its tractability allows us to explain the various mechanisms behind opinion polarisation in a transparent manner. We believe this is a necessary first step before enriching the model with more realistic features, such as private signals and communications among the agents.

Appendix

Preliminary Results

We begin with some preliminary results which will be useful in deriving the closed-form solution in Proposition 1. Suppose (X, Y) is a random vector that follows a bivariate normal distribution with mean $\left(\mu_x, \mu_y\right)$ and covariance matrix

$$\begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}.$$

The probability density function of this distribution satisfies

$$h(x, y) \propto \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\},$$

where $\propto$ is the direct proportionality symbol. The terms inside the square brackets can be expanded and regrouped to become

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - \frac{2\rho xy}{\sigma_x\sigma_y} - \frac{2}{\sigma_x}\left(\frac{\mu_x}{\sigma_x} - \frac{\rho\mu_y}{\sigma_y}\right)x - \frac{2}{\sigma_y}\left(\frac{\mu_y}{\sigma_y} - \frac{\rho\mu_x}{\sigma_x}\right)y + \left[\frac{\mu_x^2}{\sigma_x^2} - \frac{2\rho\mu_x\mu_y}{\sigma_x\sigma_y} + \frac{\mu_y^2}{\sigma_y^2}\right].$$

Using this, we can write

$$h(x, y) \propto \exp\left\{-\frac{1}{2}\left[\phi_1 x^2 + \phi_2 y^2 - 2\phi_3 xy - 2\phi_4 x - 2\phi_5 y\right]\right\}, \tag{23}$$

where

$$\phi_1 \equiv \frac{1}{(1-\rho^2)\sigma_x^2}, \qquad \phi_2 \equiv \frac{1}{(1-\rho^2)\sigma_y^2}, \qquad \phi_3 \equiv \frac{\rho}{(1-\rho^2)\sigma_x\sigma_y}, \tag{24}$$

$$\phi_4 \equiv \frac{1}{(1-\rho^2)\sigma_x}\left(\frac{\mu_x}{\sigma_x} - \frac{\rho\mu_y}{\sigma_y}\right), \tag{25}$$

$$\phi_5 \equiv \frac{1}{(1-\rho^2)\sigma_y}\left(\frac{\mu_y}{\sigma_y} - \frac{\rho\mu_x}{\sigma_x}\right). \tag{26}$$

Conversely, given a density function as in (23), we can use the above equations to recover the underlying moments $\left(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho\right)$ as follows. First, from the three equalities in (24), we can get

$$\phi_3^2 = \frac{\rho^2}{(1-\rho^2)^2\sigma_x^2\sigma_y^2} = \rho^2\phi_1\phi_2 \Rightarrow \rho = \frac{\phi_3}{\sqrt{\phi_1\phi_2}}. \tag{27}$$

This in turn implies

$$1-\rho^2=\frac{\phi_1\phi_2-\phi_3^2}{\phi_1\phi_2}.$$

Substituting this into the first two equalities in (24) gives

$$\sigma_x^2=\frac{\phi_2}{\phi_1\phi_2-\phi_3^2}\qquad\text{and}\qquad\sigma_y^2=\frac{\phi_1}{\phi_1\phi_2-\phi_3^2}. \tag{28}$$

In order to recover μ_x and μ_y, we first rewrite (25) and (26) in matrix form

$$\left(1-\rho^2\right)\begin{bmatrix}\phi_4\\ \phi_5\end{bmatrix}=\begin{bmatrix}\sigma_x^{-2} & -\rho\left(\sigma_x\sigma_y\right)^{-1}\\ -\rho\left(\sigma_x\sigma_y\right)^{-1} & \sigma_y^{-2}\end{bmatrix}\begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}$$

$$\Rightarrow\begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}=\frac{\left(1-\rho^2\right)}{\left(1-\rho^2\right)\left(\sigma_x\sigma_y\right)^{-2}}\begin{bmatrix}\sigma_y^{-2} & \rho\left(\sigma_x\sigma_y\right)^{-1}\\ \rho\left(\sigma_x\sigma_y\right)^{-1} & \sigma_x^{-2}\end{bmatrix}\begin{bmatrix}\phi_4\\ \phi_5\end{bmatrix}.$$

Using this, (27) and (28), we can get

$$\begin{aligned}\mu_x &= \sigma_x^2\phi_4+\rho\sigma_x\sigma_y\phi_5=\frac{\phi_2\phi_4}{\phi_1\phi_2-\phi_3^2}+\frac{\phi_3}{\sqrt{\phi_1\phi_2}}\frac{\sqrt{\phi_1\phi_2}}{\phi_1\phi_2-\phi_3^2}\phi_5\\ &= \frac{\phi_2\phi_4+\phi_3\phi_5}{\phi_1\phi_2-\phi_3^2},\end{aligned} \tag{29}$$

and

$$\mu_y=\rho\sigma_x\sigma_y\phi_4+\sigma_y^2\phi_5=\frac{\phi_3\phi_4+\phi_1\phi_5}{\phi_1\phi_2-\phi_3^2}. \tag{30}$$

Proof of Proposition 1

Recall that the agent's initial belief about (s,b) is given by a bivariate normal distribution with mean (s_0,b_0) and variance-covariance matrix

$$\begin{bmatrix}\sigma_{s,0}^2 & \rho_0\sigma_{s,0}\sigma_{b,0}\\ \rho_0\sigma_{s,0}\sigma_{b,0} & \sigma_{b,0}^2\end{bmatrix}.$$

As shown in the preliminary findings, the probability density function of this distribution satisfies

$$h\left(s,b\right)\propto\exp\left\{-\frac{1}{2}\left[\phi_1 s^2+\phi_2 b^2-2\phi_3 sb-2\phi_4 s-2\phi_5 b\right]\right\},$$

where

$$\phi_1 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2}, \qquad \phi_2 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}^2}, \qquad \phi_3 \equiv \frac{\rho_0}{\left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}}, \tag{31}$$

$$\phi_4 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}}\left(\frac{s_0}{\sigma_{s,0}} - \frac{\rho_0 b_0}{\sigma_{b,0}}\right), \tag{32}$$

$$\phi_5 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}}\left(\frac{b_0}{\sigma_{b,0}} - \frac{\rho_0 s_0}{\sigma_{s,0}}\right). \tag{33}$$

Conditional on (s, b), the history of signals $\mathbf{m}^t = \{m_1, ..., m_t\}$ forms an i.i.d. sequence of normal random variables with mean $(s+b)$ and variance σ_ε^2. The sequence $\mathbf{m}^t$ has a joint probability density function $f\left(\mathbf{m}^t \mid s, b\right)$ that satisfies

$$f\left(\mathbf{m}^t \mid s, b\right) \propto \exp\left\{-\frac{1}{2}\frac{\sum_{i=1}^t \left[m_i - (s+b)\right]^2}{\sigma_\varepsilon^2}\right\},$$

where

$$\sum_{i=1}^t \left[m_i - (s+b)\right]^2 = \sum_{i=1}^t m_i^2 - 2(s+b)\, t\overline{m}_t + t\left(s^2 + 2sb + b^2\right),$$

and $\overline{m}_t = \sum_{i=1}^t m_i / t$. Therefore, after observing $\mathbf{m}^t$, the agent's updated belief has a probability density function $\pi\left(s, b \mid \mathbf{m}^t\right)$ that satisfies

$$\pi\left(s, b \mid \mathbf{m}^t\right) \propto f\left(\mathbf{m}^t \mid s, b\right) h(s, b)$$

$$\Rightarrow \pi\left(s, b \mid \mathbf{m}^t\right) \propto \exp\left\{-\frac{1}{2}\left[\frac{\sum_{i=1}^t \left(m_i - (s+b)\right)^2}{\sigma_\varepsilon^2} + \phi_1 s^2 + \phi_2 b^2 - 2\phi_3 sb - 2\phi_4 s - 2\phi_5 b\right]\right\}. \tag{34}$$

The terms inside the square brackets can be regrouped to become

$$\frac{\sum_{i=1}^t m_i^2}{\sigma_\varepsilon^2} + \left(\frac{t}{\sigma_\varepsilon^2} + \phi_1\right)s^2 + \left(\frac{t}{\sigma_\varepsilon^2} + \phi_2\right)b^2 - 2\left(-\frac{t}{\sigma_\varepsilon^2} + \phi_3\right)sb - 2\left(\frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_4\right)s - 2\left(\frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_5\right)b.$$

Note that the first term in the above expression does not depend on (s, b), which means it can be included in the constant of proportionality. We can then rewrite (34) as

$$\pi\left(s, b \mid \mathbf{m}^t\right) \propto \exp\left\{-\frac{1}{2}\left[\alpha_{1,t}s^2 + \alpha_{2,t}b^2 - 2\alpha_{3,t}sb - 2\alpha_{4,t}s - 2\alpha_{5,t}b\right]\right\},$$

where

$$\alpha_{1,t} = \frac{t}{\sigma_\varepsilon^2} + \phi_1, \qquad \alpha_{2,t} = \frac{t}{\sigma_\varepsilon^2} + \phi_2, \tag{35}$$

$$\alpha_{3,t} = -\frac{t}{\sigma_\varepsilon^2} + \phi_3, \tag{36}$$

$$\alpha_{4,t} = \frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_4, \qquad \text{and} \qquad \alpha_{5,t} = \frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_5. \tag{37}$$

Our task now is to recover the mean vector and the covariance matrix associated with $\pi\left(s, b \mid \mathbf{m}^t\right)$, which are $\widehat{\mathbf{x}}_t$ and $\widehat{\boldsymbol{\Sigma}}_t$.

From (28), we know that

$$\widehat{\sigma}_{s,t}^2 = \frac{\alpha_{2,t}}{\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2} \qquad \text{and} \qquad \widehat{\sigma}_{b,t}^2 = \frac{\alpha_{1,t}}{\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2}.$$

Combining the equations in (31) and (35) gives

$$\alpha_{1,t} = \frac{t}{\sigma_\varepsilon^2} + \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2} = \frac{\sigma_{b,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{s,0}^2 t\right]}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{38}$$

$$\alpha_{2,t} = \frac{t}{\sigma_\varepsilon^2} + \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}^2} = \frac{\sigma_{s,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{b,0}^2 t\right]}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{39}$$

Similarly, it can be shown that

$$\begin{aligned}\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2 &= \left(\frac{t}{\sigma_\varepsilon^2} + \phi_1\right)\left(\frac{t}{\sigma_\varepsilon^2} + \phi_2\right) - \left(-\frac{t}{\sigma_\varepsilon^2} + \phi_3\right)^2 \\ &= \frac{1}{\sigma_\varepsilon^2}\left[\left(\phi_1 + \phi_2 + 2\phi_3\right)t + \left(\phi_1\phi_2 - \phi_3^2\right)\sigma_\varepsilon^2\right]. \end{aligned} \tag{40}$$

Using (31) and after some algebraic manipulations, we can get

$$\phi_1 + \phi_2 + 2\phi_3 = \frac{\sigma_{b,0}^2 + \sigma_{s,0}^2 + 2\rho_0\sigma_{s,0}\sigma_{b,0}}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2} = \frac{\lambda_0 + \theta_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{41}$$

and

$$\phi_1\phi_2 - \phi_3^2 = \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{42}$$

Substituting (41) and (42) into (40) gives

$$\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2 = \frac{\sigma_\varepsilon^2 + \left(\lambda_0 + \theta_0\right)t}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2} \tag{43}$$

Note that the expressions in (38), (39) and (43) all share the same denominator. Combining (39) and (43) gives (2). Similarly, (3) can be obtained by combining (38) and (43).

The covariance term $\widehat{\omega}_t$ can be obtained as follows: Based on (27), we can write

$$\widehat{\rho}_t = \frac{\alpha_{3,t}}{\sqrt{\alpha_{1,t}\alpha_{2,t}}}.$$

Combining this with (28) gives

$$\widehat{\omega}_t \equiv \widehat{\rho}_t \widehat{\sigma}_{s,t} \widehat{\sigma}_{b,t} = \frac{\alpha_{3,t}}{\sqrt{\alpha_{1,t}\alpha_{2,t}}} \cdot \frac{\sqrt{\alpha_{2,t}} \cdot \sqrt{\alpha_{1,t}}}{\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2} = \frac{\alpha_{3,t}}{\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2}.$$

From (31) and (36), we can get

$$\alpha_{3,t} = -\frac{t}{\sigma_\varepsilon^2} + \frac{\rho_0}{\left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}} = \frac{\rho_0 \sigma_{s,0}\sigma_{b,0}\sigma_\varepsilon^2 - \left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2 t}{\sigma_\varepsilon^2 \left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2},$$

where $\rho_0\sigma_{s,0}\sigma_{b,0} = \omega_0$. Equation (4) can be obtained by combining this with (43).

Using (2)-(4), we can get derive an expression for the correlation between s and b in the revised belief. This is given by

$$\begin{aligned}
\widehat{\rho}_t &= \frac{\widehat{\omega}_t}{\widehat{\sigma}_{s,t}\widehat{\sigma}_{b,t}} \\
&= \frac{\sigma_{s,0}\sigma_{b,0}\left[\rho_0\sigma_\varepsilon^2 - \left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}t\right]}{\sqrt{\sigma_{b,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{s,0}^2 t\right]} \cdot \sqrt{\sigma_{s,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{b,0}^2 t\right]}} \\
&= \frac{\rho_0\sigma_\varepsilon^2 - \left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}t}{\sqrt{\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{s,0}^2 t\right]\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{b,0}^2 t\right]}}. \qquad (44)
\end{aligned}$$

Based on (29) and (30), the means associated with the posterior density function $\pi\left(s, b \mid \mathbf{m}^t\right)$ can be expressed as

$$\widehat{s}_t \equiv E\left[s \mid \mathbf{m}^t\right] = \frac{\alpha_{2,t}\alpha_{4,t} + \alpha_{3,t}\alpha_{5,t}}{\alpha_{1,t}\alpha_{2,t} - \left(\alpha_{3,t}\right)^2}, \qquad (45)$$

$$\widehat{b}_t \equiv E\left[b \mid \mathbf{m}^t\right] = \frac{\alpha_{3,t}\alpha_{4,t} + \alpha_{1,t}\alpha_{5,t}}{\alpha_{1,t}\alpha_{2,t} - \left(\alpha_{3,t}\right)^2}. \qquad (46)$$

Using (35)-(37), we can get

$$\alpha_{2,t}\alpha_{4,t} + \alpha_{3,t}\alpha_{5,t} = \frac{1}{\sigma_\varepsilon^2}\left\{\left[\left(\phi_2+\phi_3\right)\overline{m}_t + \left(\phi_4 - \phi_5\right)\right]t + \left(\phi_2\phi_4 + \phi_3\phi_5\right)\sigma_\varepsilon^2\right\}, \qquad (47)$$

$$\alpha_{3,t}\alpha_{4,t} + \alpha_{1,t}\alpha_{5,t} = \frac{1}{\sigma_\varepsilon^2}\left\{\left[\left(\phi_1+\phi_3\right)\overline{m}_t - \left(\phi_4 - \phi_5\right)\right]t + \left(\phi_3\phi_4 + \phi_1\phi_5\right)\sigma_\varepsilon^2\right\}. \qquad (48)$$

Using (31)-(33), we can derive

$$\phi_2 + \phi_3 = \frac{\lambda_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2},$$

$$\phi_4 - \phi_5 = \frac{\theta_0 s_0 - \lambda_0 b_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2},$$

$$\phi_2\phi_4 + \phi_3\phi_5 = \frac{s_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2},$$

$$\phi_1 + \phi_3 = \frac{\theta_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2},$$

$$\phi_3\phi_4 + \phi_1\phi_5 = \frac{b_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}.$$

Substituting these into (47) and (48) gives

$$\alpha_{2,t}\alpha_{4,t} + \alpha_{3,t}\alpha_{5,t} = \frac{\left(\lambda_0 \overline{m}_t + \theta_0 s_0 - \lambda_0 b_0\right)t + s_0\sigma_\varepsilon^2}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{49}$$

$$\alpha_{3,t}\alpha_{4,t} + \alpha_{1,t}\alpha_{5,t} = \frac{\left(\theta_0 \overline{m}_t - \theta_0 s_0 + \lambda_0 b_0\right)t + b_0\sigma_\varepsilon^2}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{50}$$

Finally, substituting (43) and (49) into (45) gives

$$\begin{aligned}
\widehat{s}_t &= \frac{\left(\lambda_0 \overline{m}_t + \theta_0 s_0 - \lambda_0 b_0\right)t + s_0\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \left(\lambda_0+\theta_0\right)t} \\
&= \frac{\left(\lambda_0 \overline{m}_t - \lambda_0 s_0 - \lambda_0 b_0\right)t + s_0\left[\sigma_\varepsilon^2 + \left(\lambda_0+\theta_0\right)\right]t}{\sigma_\varepsilon^2 + \left(\lambda_0+\theta_0\right)t} \\
&= s_0 + \frac{\lambda_0 t}{\sigma_\varepsilon^2 + \left(\lambda_0+\theta_0\right)t}\left(\overline{m}_t - s_0 - b_0\right),
\end{aligned}$$

which is (5). Equation (6) can be obtained in a similar fashion by substituting (43) and (50) into (46). This completes the proof of Proposition 1. ■

Proof of Proposition 2

At any time $t \geq 1$, after observing the signals $\mathbf{m}^t = \{m_1, ..., m_t\}$, the agent's revised belief is given by $\mathbf{N}\left(\widehat{\mathbf{x}}_t, \widehat{\mathbf{\Sigma}}_t\right)$. Under this belief, the expected value of any future signal m_{t+k}, with $k \geq 1$, is

$$E_t\left(m_{t+k}\right) = \widehat{s}_t + \widehat{b}_t,$$

and the covariance between s and m_{t+k} is

$$\begin{aligned}
\widehat{\lambda}_t &\equiv Cov_t\left(s, m_{t+k}\right) = E_t\left[\left(s - \widehat{s}_t\right)\left(m_{t+k} - \widehat{s}_t - \widehat{b}_t\right)\right] \\
&= E_t\left[\left(s - \widehat{s}_t\right)\left(s - \widehat{s}_t + b - \widehat{b}_t - \varepsilon_{t+k}\right)\right] \\
&= E_t\left[\left(s - \widehat{s}_t\right)^2 + \left(s - \widehat{s}_t\right)\left(b - \widehat{b}_t\right)\right] \\
&= \widehat{\sigma}^2_{s,t} + \widehat{\omega}_t.
\end{aligned}$$

The third line uses the fact that $E_t\left[\left(s - \widehat{s}_t\right)\varepsilon_{t+k}\right] = 0$. Substituting (2) and (4) into the above equation gives

$$\widehat{\lambda}_t = \frac{\sigma^2_\varepsilon\left(\sigma^2_{s,0} + \omega_0\right)}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t} = \frac{\sigma^2_\varepsilon\lambda_0}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}.$$

Hence, $\widehat{\lambda}_t \gtreqless 0$ if and only if $\lambda_0 \gtreqless 0$.

Using the same line of argument, we can get

$$\widehat{\theta}_t \equiv Cov_t\left(b, m_{t+k}\right) = \frac{\sigma^2_\varepsilon\theta_0}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}.$$

Hence, $\widehat{\theta}_t \gtreqless 0$ if and only if $\theta_0 \gtreqless 0$. This completes the proof of Proposition 2. ■

Proof of Proposition 3

Part (a) Rewrite (2) as follows

$$\begin{aligned}
\widehat{\sigma}^2_{s,t} &= \frac{\left[\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t\right]\sigma^2_{s,0} + \left[\left(1 - \rho_0^2\right)\sigma^2_{b,0} - \left(\lambda_0 + \theta_0\right)\right]\sigma^2_{s,0}t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t} \\
&= \sigma^2_{s,0} + \frac{\left[\left(1 - \rho_0^2\right)\sigma^2_{b,0} - \left(\lambda_0 + \theta_0\right)\right]\sigma^2_{s,0}t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t},
\end{aligned}$$

where

$$\begin{aligned}
\left[\left(1 - \rho_0^2\right)\sigma^2_{b,0} - \left(\lambda_0 + \theta_0\right)\right]\sigma^2_{s,0} &= -\sigma^2_{s,0}\left(\sigma^2_{s,0} + 2\rho_0\sigma_{b,0}\sigma_{s,0} + \rho_0^2\sigma^2_{b,0}\right) \\
&= -\left(\sigma^2_{s,0} + \rho_0\sigma_{b,0}\sigma_{s,0}\right)^2 = -\lambda_0^2.
\end{aligned}$$

Hence, we can write

$$\widehat{\sigma}^2_{s,t} = \sigma^2_{s,0} - \frac{\lambda_0^2 t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t} = \sigma^2_{s,0} - \frac{\lambda_0^2}{\frac{\sigma^2_\varepsilon}{t} + \left(\lambda_0 + \theta_0\right)}.$$

This shows that $\widehat{\sigma}^2_{s,t}$ is strictly decreasing over time. To derive the limit, it is more straightforward to use (2), which gives

$$\lim_{t\to\infty} \widehat{\sigma}^2_{s,t} = \lim_{t\to\infty} \left\{ \frac{\left[\frac{\sigma^2_\varepsilon}{t} + \left(1-\rho^2_0\right)\sigma^2_{b,0}\right]\sigma^2_{s,0}}{\frac{\sigma^2_\varepsilon}{t} + (\lambda_0+\theta_0)} \right\} = \frac{\left(1-\rho^2_0\right)\sigma^2_{b,0}\sigma^2_{s,0}}{\sigma^2_{s,0}+\sigma^2_{b,0}+2\omega_0}.$$

The proof for $\widehat{\sigma}^2_{b,t}$ is essentially identical, hence it is omitted.

Part (b) Define α_t according to

$$\alpha_t \equiv \kappa_t + \eta_t = \frac{(\lambda_0+\theta_0)\,t}{\sigma^2_\varepsilon + (\lambda_0+\theta_0)\,t}.$$

Then combining (5) and (6) gives

$$\widehat{s}_t + \widehat{b}_t = s_0 + b_0 + \alpha_t\left(\overline{m}_t - s_0 - b_0\right) = \alpha_t\overline{m}_t + (1-\alpha_t)(s_0+b_0).$$

$$\Rightarrow \widehat{s}_t + \widehat{b}_t - s - b = \alpha_t\left(\overline{m}_t - s - b\right) + (1-\alpha_t)\underbrace{(s_0+b_0-s-b)}_{\Delta}, \tag{51}$$

where Δ is a constant. For any $\nu > 0$,

$$\begin{aligned} \Pr\left(\left|\widehat{s}_t + \widehat{b}_t - s - b\right| \geq \nu\right) &= \Pr\left[\left(\widehat{s}_t + \widehat{b}_t - s - b\right)^2 \geq \nu^2\right] \\ &\leq \frac{E\left[\left(\widehat{s}_t + \widehat{b}_t - s - b\right)^2\right]}{\nu^2}. \end{aligned}$$

The second line follows from Markov's inequality. The expectation is taken over the distribution of $\overline{m}_t$ which is normal with mean $(s+b)$ and variance σ^2_ε/t. Using this fact and (51), we can write

$$\begin{aligned} E\left[\left(\widehat{s}_t + \widehat{b}_t - s - b\right)^2\right] &= \alpha^2_t E\left[\left(\overline{m}_t - s - b\right)^2\right] + (1-\alpha_t)^2\Delta \\ &= \frac{\alpha^2_t\sigma^2_\varepsilon}{t} + (1-\alpha_t)^2\Delta. \end{aligned}$$

It is easy to verify that $\alpha_t \to 1$ as $t \to \infty$. Hence, for any $\nu > 0$

$$\lim_{t\to\infty} \Pr\left(\left|\widehat{s}_t + \widehat{b}_t - s - b\right| \geq \nu\right) \leq \frac{1}{\nu^2}\cdot\lim_{t\to\infty}\left[\frac{\alpha^2_t\sigma^2_\varepsilon}{t} + (1-\alpha_t)^2\Delta\right] = 0.$$

Part (c) Rewrite (4) and (44) as

$$\widehat{\omega}_t = \frac{\frac{\omega_0 \sigma_\varepsilon^2}{t} - \left(1-\rho_0^2\right)\sigma_{s,0}^2 \sigma_{b,0}^2}{\frac{\sigma_\varepsilon^2}{t} + (\lambda_0 + \theta_0)},$$

$$\widehat{\rho}_t = \frac{\frac{\rho_0 \sigma_\varepsilon^2}{t} - \left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}}{\sqrt{\left[\frac{\sigma_\varepsilon^2}{t} + \left(1-\rho_0^2\right)\sigma_{s,0}^2\right]\left[\frac{\sigma_\varepsilon^2}{t} + \left(1-\rho_0^2\right)\sigma_{b,0}^2\right]}},$$

Taking the limit $t \to \infty$ gives

$$\lim_{t\to\infty} \widehat{\omega}_t = -\frac{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}{\sigma_{s,0}^2 + \sigma_{b,0}^2 + 2\omega_0} < 0,$$

$$\lim_{t\to\infty} \widehat{\rho}_t = \frac{-\left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}}{\sqrt{\left(1-\rho_0^2\right)\sigma_{s,0}^2 \cdot \left(1-\rho_0^2\right)\sigma_{b,0}^2}} = -1.$$

This completes the proof of Proposition 3. ■

Proof of Proposition 4

Combining (5) and (10) gives

$$\begin{aligned} \widehat{s}_t - \widehat{s}_\infty &= \kappa_t\left(\overline{m}_t - s - b + s + b - s_0 - b_0\right) - \kappa_\infty\left(s + b - s_0 - b_0\right) \\ &= \kappa_t\left(\overline{m}_t - s - b\right) + \left(\kappa_t - \kappa_\infty\right)\left(s + b - s_0 - b_0\right). \end{aligned}$$

Define $\xi_t \equiv \overline{m}_t - s - b$ and $\Delta \equiv (s + b - s_0 - b_0)$. Before any signal is realised, $\overline{m}_t$ is a normal random variable with mean $s + b$ and variance σ_ε^2/t. Hence, $E(\xi_t) = 0$ and $var(\xi_t) = \sigma_\varepsilon^2/t$. Then for any $\nu > 0$,

$$\begin{aligned} \Pr\left(\left|\widehat{s}_t - \widehat{s}_\infty\right| \geq \nu\right) &= \Pr\left(\left|\kappa_t \xi_t + (\kappa_t - \kappa_\infty)\Delta\right| \geq \nu\right) \\ &= \Pr\left[\left(\kappa_t \xi_t + (\kappa_t - \kappa_\infty)\Delta\right)^2 \geq \nu^2\right] \\ &\leq \frac{E\left[\left(\kappa_t \xi_t + (\kappa_t - \kappa_\infty)\Delta\right)^2\right]}{\nu^2}. \end{aligned}$$

The last inequality follows from Markov's inequality. The expectation can be simplified as follows:

$$E\left[\left(\kappa_t \xi_t + (\kappa_t - \kappa_\infty)\Delta\right)^2\right] = \frac{\kappa_t^2 \sigma_\varepsilon^2}{t} + (\kappa_t - \kappa_\infty)^2 \Delta^2.$$

Substituting this back into the inequality gives

$$\Pr\left(|\widehat{s}_t - \widehat{s}_\infty| \geq \nu\right) \leq \left(\frac{\kappa_t \sigma_\varepsilon}{\nu}\right)^2 \frac{1}{t} + \frac{\left(\kappa_t - \kappa_\infty\right)^2 \Delta^2}{\nu^2}.$$

As mentioned in the main text, $\kappa_t \to \kappa_\infty$ as $t \to \infty$. Therefore, we can get

$$\lim_{t\to\infty} \Pr\left(|\widehat{s}_t - \widehat{s}_\infty| \geq \nu\right) = 0.$$

Using the same line of argument, we can show that $\widehat{b}_t \xrightarrow{p} \widehat{b}_\infty$. ■

Proof of Proposition 5

Suppose $s_0^\dagger \geq s_0^\ddagger$. First consider the case in which $\lambda_0 > 0$ so that $\kappa_t > 0$ for all $t \geq 1$. Then $\widehat{s}_t^\ddagger > \widehat{s}_t^\dagger$ holds at some time $t \geq 1$ if and only if

$$\left(b_0^\dagger - b_0^\ddagger\right) > \frac{1-\kappa_t}{\kappa_t}\left(s_0^\dagger - s_0^\ddagger\right),$$

where

$$\frac{1-\kappa_t}{\kappa_t} = \frac{1}{\lambda_0}\left(\frac{\sigma_\varepsilon^2}{t} + \theta_0\right),$$

is decreasing over time as $\lambda_0 > 0$. It follows that

$$\left(b_0^\dagger - b_0^\ddagger\right) > \frac{1-\kappa_t}{\kappa_t}\left(s_0^\dagger - s_0^\ddagger\right) \geq \frac{1-\kappa_{t+k}}{\kappa_{t+k}}\left(s_0^\dagger - s_0^\ddagger\right),$$

with strictly equality holds only if $s_0^\dagger = s_0^\ddagger$. The second inequality implies $\widehat{s}_{t+k}^\ddagger > \widehat{s}_{t+k}^\dagger$, for all $k \in \{1, 2, ...\}$.

Next, consider the case when $\lambda_0 < 0$ so that $\kappa_t < 0$ for all t. Then $\widehat{s}_t^\ddagger > \widehat{s}_t^\dagger$ holds if and only if

$$\left(b_0^\dagger - b_0^\ddagger\right) < \frac{1-\kappa_t}{\kappa_t}\left(s_0^\dagger - s_0^\ddagger\right).$$

When $\lambda_0 < 0$, the ratio $\left(1-\kappa_t\right)/\kappa_t$ is strictly increasing over time so that

$$\left(b_0^\dagger - b_0^\ddagger\right) < \frac{1-\kappa_t}{\kappa_t}\left(s_0^\dagger - s_0^\ddagger\right) \leq \frac{1-\kappa_{t+k}}{\kappa_{t+k}}\left(s_0^\dagger - s_0^\ddagger\right),$$

with strictly equality holds only if $s_0^\dagger = s_0^\ddagger$. The second inequality implies $\widehat{s}_{t+k}^\ddagger > \widehat{s}_{t+k}^\dagger$, for all $k \in \{1, 2, ...\}$. This completes the proof. ■

References

[1] Acemoglu, D., V. Chernozhukov and M. Yildiz (2016) "Fragility of Asymptotic Agreement under Bayesian Learning," *Theoretical Economics*, vol. 11, pp. 187-225.

[2] Abramitzky, R., L. Boustan, E. Jácome, S. Pérez and J. D. Torres (2024) "Law-Abiding Immigrants: The Incarceration Gap between Immigrants and the US-Born, 1870–2020," *American Economic Review: Insights*, vol. 6 (4), pp. 453–71.

[3] Andreoni, J., and T. Mylovanov (2012) "Diverging Opinions," *American Economic Journal: Microeconomics*, vol. 4 (1), pp. 209-232.

[4] Aytimur, R. E., and R. M. H. Suen (2025) "Information Quality, Disagreement and Political Polarisation," unpublished manuscript.

[5] Baliga, S., E. Hanany and P. Klibanoff (2013) "Polarization and Ambiguity," *American Economic Review*, vol. 103 (7), pp. 3071-3083.

[6] Barberis, N. (2018) "Psychology-Based Models of Asset Prices and Trading Volume," in *Handbook of Behavioral Economics,* vol. 1, edited by B. D. Bernheim, S. DellaVigna and D. Laibson, pp. 79-174. North-Holland, Amsterdam.

[7] Benjamin, D. J. (2019) "Errors in Probabilistic Reasoning and Judgement Biases," in *Handbook of Behavioral Economics,* vol. 2, edited by B. D. Bernheim, S. DellaVigna and D. Laibson, pp. 69-186. North-Holland, Amsterdam.

[8] Blackwell, D., and L. Dubins (1962) "Merging of Opinions with Increasing Information," *Annals of Mathematical Statistics*, vol. 33 (3), pp. 882-886.

[9] Brenan, M. (2024) "Americans' Trust in Media Remains at Trend Low," Gallup. https://news.gallup.com/poll/651977/ americans-trust-media-remains-trend-low.aspx.

[10] Bullock, J. G. (2009) "Partisan Bias and the Bayesian Ideal in the Study of Public Opinion," *Journal of Politics*, vol. 71 (3), pp. 1109-1124.

[11] Chamley, C. P. (2004) *Rational Herds.* Cambridge University Press. Cambridge, UK.

[12] DeGroot, M. H. (1970) *Optimal Statistical Decisions.* John Wiley & Sons. New Jersey.

[13] Dixit, A. K., and J. W. Weibull (2007) "Political Polarization," *Proceedings of the National Academy of Sciences*, vol. 104 (18), pp. 7351-7356.

[14] Garz, M., G. Sood, D. F. Stone and J. Wallace (2020) "The Supply of Media Slant across Outlets and Demand for Slant within Outlets: Evidence from US Presidential Campaign News," *European Journal of Political Economy*, vol. 63, article no. 101877.

[15] Heidhues, P., B. Köszegi and P. Strack (2018) "Unrealistic Expectations and Misguided Learning," *Econometrica*, vol. 86 (4), pp. 1159-1214.

[16] Jones, J. M. (2024) "Far Fewer in U.S. Regard Childhood Vaccinations as Important," Gallup. https://news.gallup.com/ poll/ 648308/ far-fewer-regard-childhood-vaccinations - important.aspx.

[17] Kandel, E., and B. Zilberfarb (1999) "Differential Interpretation of Information in Inflation Forecasts," *Review of Economics and Statistics*, vol. 81 (2), pp. 217-226.

[18] Kartik, N., F. X. Lee and W. Suen (2021) "Information Validates the Prior: A Theorem on Bayesian Updating and Applications," *American Economic Review: Insights*, vol. 3 (2), pp. 165-182.

[19] Kline, B., and E. Tamer (2023) "Recent Developments in Partial Identification," *Annual Review of Economics*, vol. 15, pp. 125-150.

[20] Kondor, P. (2012) "The More We Know about the Fundamental, The Less We Agree on the Price," *Review of Economic Studies*, vol. 79 (3), pp. 1175-1207.

[21] Lahiri, K., and X. Sheng (2008) "Evolution of Forecast Disagreement in a Bayesian Learning Model," *Journal of Econometrics*, vol. 144, pp. 325-340.

[22] Leiserowitz, A., J. Kotcher, S. Rosenthal, E. Goddard, J. Carman, M. Verner, T. Myers, J. Ettinger, J. Fine, E. Richards, M. Goldberg, J. Marlon and E. Maibach (2026) *Climate Change in the American Mind: Beliefs & Attitudes, Fall 2025.* Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.

[23] Liang, A., and X. Mu (2020) "Complementary Information and Learning Traps," *Quarterly Journal of Economics*, vol. 135 (1), pp. 389-448.

[24] Little, A. T., and T. B. Pepinsky (2021) “Learning from Biased Research Designs,” *Journal of Politics*, vol. 83 (2), pp. 602-616.

[25] Little, A. T., A. Mack and T. B. Pepinsky (2025) “Biased Learning from Elections,” unpublished manuscript.

[26] Marie, O., and P. Pinotti (2024) “Immigration and Crime: An International Perspective,” *Journal of Economic Perspectives*, vol. 38 (1), pp. 181–200.

[27] Newman, N., and R. Fletcher (2017) *Bias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media*. The Reuters Institute for the Study of Journalism, Oxford, UK.

[28] Newman, N., R. Fletcher, C. T. Robertson, A. R. Arguedas, and R. K. Nielsen (2024) *Reuters Institute Digital News Report 2024*. The Reuters Institute for the Study of Journalism, Oxford, UK.

[29] Puglisi, R., and J. M. Snyder Jr. (2015) “Empirical Studies of Media Bias,” in *Handbook of Media Economics*, vol. 1, edited by S. P. Anderson, J. Waldfogel and D. Strömberg, pp. 647-667. North-Holland, Amsterdam.