Biased Information and Opinion Polarisation

Suen, Richard M. H.

University of Leicester

May 2025

Biased Information and Opinion Polarisation

Richard M. H. Suen*

First Version: 26th May, 2025.

Abstract

Why do people form polarised opinions after receiving the same information? Why does disagreement persist even when public information is abundant? We show that a Bayesian model with potentially biased public signals can answer these questions. When agents are uncertain and disagree about the bias in the signals, persistent disagreement and opinion polarisation can readily emerge. This happens because uncertainty surrounding the bias induces agents with diverse initial beliefs to form drastically different posterior estimates. Prolonged exposure to these signals can in some cases drive the agents' opinions further away from each other and also further away from the truth.

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

JEL Classification: C11, D83

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

1 Introduction

Disagreement is ubiquitous. From political discourses to academic debates to stock market predictions, disagreement exists and persists even when public information is abundant. On some contentious issues (such as immigration and vaccination), people form diverging views even after receiving the same information, leading to opinion polarisation.[1] These observations pose a challenge to the canonical Bayesian learning model which contends that persistent disagreement and opinion divergence will not occur in general. The current study is an attempt to reconcile these conflicts between theory and evidence. Our approach is motivated by a separate but closely related empirical literature which documents the pervasiveness of biased reporting in mass media and a dwindling trust in them among the public.[2] We show that when Bayesian learners are uncertain and disagree about the potential bias in the commonly observed signals, persistent disagreement and opinion polarisation can readily emerge.

Our analysis is built upon a prototypical Bayesian learning model in which economic agents seek to infer an unknown state of the world (a payoff-relevant parameter) from a stream of noisy signals. We enrich this model by assuming that the random signals are not only fraught with errors, they are also *potentially* confounded by another unobserved factor which we refer to as bias. We emphasise on the word "potentially" because it is the agents' perceptions or doubts surrounding the bias (rather than the actual bias) that drives our main results. This point will become clear as we proceed. The unobserved bias is an inherent feature of the signal-generating channel, hence it will affect all the outcoming signals.[3] As an illustration, consider news reports from mass media. The content of these reports may contain factual errors or other forms of unintended mistakes that are very costly (if not impossible) to eradicate. In terms of modelling, these mistakes 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 (editorial bias, 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 bias cannot be easily filtered

[1] Throughout this paper, we use the terms "opinion polarisation" and "opinion divergence" interchangeably to refer to a situation in which two agents receive the same information but form opposite or contradictory opinions.

[2] 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 Gallup (2024) based on US data and Newman *et al.* (2024) based on data from a large group of countries.

[3] We only consider one source of public signal but our model can be readily extended to allow for multiple public signals, each with an unknown bias. We choose the single-signal version to convey the key messages of this study in the most straightforward manner, without burdening the reader with excessive technical details.

out through averaging. Similar examples abound in other contexts.[4] 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. We model systemic bias as a separate additive term embedded in each realisation of the signal and make two important assumptions. First, the magnitude of the bias term is unknown, and second, there is a lack of consensus regarding this magnitude among the agents. We argue that the uncertainty or skepticism surrounding the bias is an important driving force that induces agents with diverse initial beliefs to form drastically different posterior estimates.[5] The effect of this skepticism exists even if the signals are truly unbiased (i.e., the true value of the bias term is identical to zero). Moreover, continuous exposure to these signals can in some cases drive the agents' opinions further away from each other and also further away from the truth.

To formalise these ideas, we hypothesise that when agents face a potentially biased information channel, they form an initial subjective belief on both the hidden state and the unknown bias term. Upon the arrival of new information, they revise their beliefs on these two parameters jointly using Bayes' rule. We examine the implications of this learning process at three different levels: individual, interpersonal and aggregate levels. We begin in Section 2 by considering a single Bayesian learner who observes a sequence of serially independent but potentially biased Gaussian signals. Using a conjugate bivariate normal distribution as the initial belief, the model admits an exact closed-form solution for all subsequently revised beliefs. We show that this sequence of beliefs is convergent and, importantly, the limiting distribution is dependent on the initial belief. In Section 3, we compare two Bayesian learners who observe the same sequence of signals but have different initial beliefs.[6] Here we explore the conditions under which polarised opinions about the hidden state will emerge in the limit. Finally, in Section 4, we extend our analysis to a population of agents with heterogeneous initial beliefs. Here we focus on how pre-existing disagreement about the bias term across the agents will affect the extent of long-term disagreement about the hidden state at the aggregate level.[7]

Before presenting our main results, we first recall the implications of unbiased signals as a

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

[5] By "estimate" and "opinion," we are referring to the posterior mean, which is the minimum variance estimate given the information available to an agent.

[6] When comparing across different agents in Sections 3 and 4, we assume that their initial belief can be represented as a bivariate normal distribution but with different parameters.

[7] In order to single out the effects of the unknown bias term in public signals, we do not consider other sources of information (such as private signals and communications among the agents) in Sections 3 and 4.

point of reference. Consider an environment in which all the signals are truly unbiased, and most importantly, this is *common knowledge* among the agents. This means they all accept that the bias term is identical to zero in their initial belief and there is *no uncertainty* about it.[8] We refer to this simply as the conventional model. With Gaussian noises and a conjugate prior, the conventional model has three main predictions that are most relevant to the current study.[9] First, after observing a sufficiently large sample of signals, any Bayesian learner will be able to infer the true value of the hidden state.[10] Second, it follows immediately that 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, two agents observing the same signal will never revise their estimates in opposite directions and move further apart.

These results may fail to hold once we introduce an unknown bias term. At the individual level, we show that agents' revised beliefs will converge to a non-degenerate distribution which is 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, their long-run estimates are both biased and inconsistent. This happens because uncertainty surrounding the systemic bias creates an identification problem in the learning process: in the limit, all agents (regardless of their initial belief) can fully identify the *sum* of the hidden state and the bias term but not their separate values.[13] In addition, initial belief about the bias will influence how they respond to the signals and revise their estimate for the hidden state. These effects remain even in the limit.

Based on these findings, it is not surprising that any initial disagreement among the agents will continue in the long run. Opinion polarisation, on the other hand, is less straightforward. In

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

[9] Two reasons why we focus on Gaussian model. First, it is by far one of the most commonly used specifications in economics and related fields. Hence, the results mentioned below are most familiar among economists. Second, it facilitates a direct comparison with our results. Further details about these predictions are provided in subsequent footnotes.

[10] This means the agent's limiting belief is degenerate at the true value of the hidden state. In other words, the agent's long-run estimate is both unbiased and consistent. This result remains valid if the hidden state has a finite number of possibel 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 cases, the limiting distribution can be non-degenerate. See, for instance, Chamley (2004, Section 2.4).

[11] Black 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 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) consider 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.

order to make clear 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.** The first one concerns how an agent responds to the observed signals along the convergent path, while the latter concerns the long-run properties of the agent's estimates. 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 high-than-expected signal is an indication to the agent that her current 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 indicated by the signals.[14] 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 presumes a significant negative correlation between the hidden state and the bias in her initial belief. To see how this can lead to defiant learning, consider the following example: Suppose an agent receives a news report on the performance of a new vaccine which exceeds her prior expectation. However, due to a lack of confidence on the news channel, the agent suspects that the report may contain a bias that is negatively correlated with the vaccine's true performance. 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 poorer than expected, the agent will 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 of 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 a long stream of potentially biased 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. Whether misguided learning will happen is ultimately controlled by the parameters in the initial belief: the mean parameters determine whether the agent is initially overestimating or underestimating the unknowns, while those in the covariance matrix determine how the agent responds to the "surprises" (i.e., deviations from prior expecta-

[14] Formally, this means the agent's revised estimate (posterior mean) is a monotonically increasing function in the observed signals. It follows from this property that opinion divergence cannot occur. See Dixit and Weibull (2007, p.7352) for an illustration. Here we emphasise on the *sign* of the correlation between the signals and the hidden state.

tion) in the signals.[15] We show that misguided learning can happen under several combinations of these parameters. This type of learning outcome cannot emerge in the conventional model because any initial misjudgment will be corrected in subsequent revisions. This self-correcting mechanism, however, is disrupted by the perceived bias in our model. Defiant learning is one example of this, but it is neither necessary nor sufficient for misguided learning.[16] Nonetheless, it does present an interesting scenario which showcases the idea that the mere doubt surrounding the systemic bias can lead a learner away from the truth. Suppose now the signals in our model are, unbeknownst to the agents, truly unbiased. Suppose an agent has the correct initial estimate of the bias term, but she is not entirely sure about it (i.e., the variance of the bias term in her initial belief is strictly positive). Over time as more information arrives, the sample mean of the signals (a sufficient statistic) will cluster around the true state. If an agent underestimates [resp., overestimates] the hidden state initially, then she will be receiving higher-than-expected [resp., lower-than-expected] news more frequently than otherwise. If, in addition, the agent is a defiant learner, then she will attribute this to a higher-than-expected [resp., lower-than-expected] bias term and revise her estimate for the hidden state downward [resp., upward], thus moving further away from the truth.[17]

Equipped with these findings at the individual level, it is now straightforward to compare two agents who observe the same sequence of signals but have different initial beliefs. Our main interest is on *permanent* opinion divergence, i.e., divergence in the agents' long-run estimates of the hidden state.[18] We divide the analysis into two parts. In the first one, the two agents have different initial estimates but share the same covariance matrix in their initial beliefs. This means they will respond in the same way to the "surprises" in the signals; in particular, they are either both conventional learners or both defiant learners. The mechanism behind permanent opinion divergence is similar to that of misguided learning, but instead of moving further away from the truth, the two agents are now moving further away from the other's initial estimate. This happens when the arrival of new information leads the agent with the relatively higher [resp.,

[15]For instance, whether defiant learning will happen depends on the correlation parameter in the covariance matrix of the initial belief.

[16]In other words, misguided learning can also happen among the conventional learners in our model.

[17]Our model also predicts a third type of learning behaviour which is not possible in the conventional model. We label this as opinion reversal. At the individual level, this happens when an agent overcorrects her initial estimate after observing the signals. For example, an agent who overestimates the hidden state initially ends up underestimating it in the long run. At the interpersonal level, this happens when the agent with a lower initial estimate eventually "catches up" and overtakes the other agent's estimate. We focus on defiant learning and misguided learning here because they are most directly relevant to our polarisation results. Further details about opinion reversal can be found in Sections 2.4 and 3.

[18]Since the public signals are randomly fluctuating over time, opinion divergence may appear in one period and disappear in another along the convergent path.

lower] initial estimate to revise upward [resp., downward] and thus widening the gap between the two. In the second part, we compare two agents that share the same initial estimates but have different covariance matrices. In this case, permanent opinion divergence happens if and only if one agent is a conventional learner and the other is a defiant learner. The intuition is straightforward: since the two start with the same initial estimates, the only way they can end up having polarised opinions is by responding to the same signals in opposite directions.

From the above analysis, we have learned how the severity and occurrence of permanent opinion divergence depend on the initial beliefs of those in comparison. A natural follow-up question is how these effects will play out in a population of agents with diverse initial beliefs. In particular, whether exposure to a long stream of potentially biased signals will increase the extent of disagreement at the aggregate level. To address this question, we consider a hypothetical population in which all agents' initial beliefs belong to the class of bivariate-normal distributions but with different parameters. We do not impose a specific cross-sectional distribution of these parameters. Instead, we use the cross-sectional variance of estimates as summary measures of disagreement within the population. Three such measures are of particular importance. These are the cross-sectional variance of the initial estimate of the hidden state and the bias term, denoted by $var(s_0)$ and $var(b_0)$, respectively; and the counterpart for the long-run estimate of the hidden state, denoted by $var(\widehat{s}_\infty)$.[19] Aggregate disagreement is said to increase after exposure if $var(\widehat{s}_\infty)$ is greater than $var(s_0)$. Three main lessons emerge from this analysis. The first two are obtained under the assumption that all agents share the same covariance matrix in their initial beliefs, while the third one is under the assumption that they share the same initial estimates. The first result states that, holding other things constant, greater initial disagreement in the bias term will raise the value of $var(\widehat{s}_\infty)$. This is true even when there is no initial disagreement in the hidden state across the agents, i.e., when $var(s_0)$ is zero. This shows that pre-existing disagreement about the bias term can by itself generate long-term disagreement about the hidden state. Second, provided that initial disagreement about the hidden state exists [i.e., $var(s_0)$ is not zero], the mere doubt surrounding the bias term can sustain or even widen disagreement in the long run. This result holds even if there is no initial disagreement on the systemic bias [i.e., when $var(b_0)$ is zero] and regardless of the true value of the bias term.[20] In addition, for

[19] The first two are taken as fundamentals of the model, while $var(\widehat{s}_\infty)$ is an endogenous variable.

[20] A corollary of this result is as follows: Suppose, unbeknownst to the agents, the signals are truly biased, i.e., the true value of the bias term (b) is zero. Suppose all agents share the *same correct* initial estimate of b, so that $var(b_0) = 0$, but they are not entirely sure about this, which means the variance of b in their initial belief $(\sigma_{b,0}^2)$ is strictly positive. Then, initial disagreement in the hidden state will persist in the long run [i.e., $var(s_0) > 0$ implies $var(\widehat{s}_\infty) > 0$]. On the other hand, if $b = var(b_0) = \sigma_{b,0}^2 = 0$ as in the conventional model, then $var(\widehat{s}_\infty) = 0$

a population of conventional learners, disagreement decreases after exposure, i.e., $var(\widehat{s}_\infty)$ is non-zero but strictly less than $var(s_0)$. But for a population of defiant learners, disagreement *increases* after they are exposed to a long stream of potentially biased signals. Finally, even if there is no pre-existing disagreement about the unknowns [i.e., both $var(s_0)$ and $var(b_0)$ are zero], long-term disagreement can exist due to cross-sectional differences in the covariance matrix in the initial belief. This happens because these differences will generate differential responses (e.g., conventional vs defiant learning) to the same signals among the agents.

Related Literature Several other 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 (i.e., the density function of the signal conditional on the hidden state) exhibits monotone likelihood ratio (MLR) property. This observation motivates them to devise models where the MLR property does not hold. Short-run polarisation in the mean estimates can emerge under this approach; however, the authors caution that permanent polarisation is still not possible.[21] Baliga *et al.* (2013) take a different approach and assume that agents display ambiguity-averse preferences. This type of preferences gives rise to a hedging motive when the agents are forming their posterior estimates. These authors show that opinion divergence can emerge between two Bayesian learners with sufficiently extreme and polarised prior 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 have different interpretations about the public signals. In the theoretical model of Andreoni and Mylovanov (2012), these differences arise because different agents receive different private signals which they use to interpret the public signals.[22] Acemoglu *et al.* (2016) show that long-run disagreement can readily emerge when agents are uncertain about the signal-generating process, or more precisely the probability distribution of the signal conditional on the hidden state. They consider a general setup without specifying the reason for this uncertainty and they have not explored the possibility of opinion polarisation. Our study complements and

regardless of $var(s_0)$.

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

[22] 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.

extends this work in three substantial ways. First, in a broad sense, uncertainty about the bias in our model can be seen as one reason why agents are unsure about the conditional distribution of the signals.[23] This provides a specific context under which disagreement will occur and links the theory to the empirical evidence on biased information. Second, we demonstrate the possibility of both permanent disagreement and opinion polarisation in an easily tractable manner. Third, we present novel results, such as defiant learning and misguided learning, which can be of interests in other applications.

Some other recent studies share similar features as the current one, 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 is "overconfident", which means she has erroneous or mis-specified belief about the signal-generating process. In their model, the signals not only include a confounding factor (the agent's ability) but also an endogenous choice variable (action). Liang and Xu (2020) considers a model in which agents receive multiple sources of biased signals. But instead of exploring disagreement and opinion polarisation, they focus on efficient information aggregation among the agents. Bayesian learning models with biased signals have also appeared in the political science literature. Little and Pepinsky (2021) use this type of models to discuss issues related to empirical estimations. Aytimur and Suen (2024) and Little *et al.* (2025) examine electoral competition between two political parties in this kind of learning environment.

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 of the information channel, 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

[23] 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 distribution is evolving over time. This mechanism is absent in Acemoglu *et al.* (2016).

known to the agent at the outset. While the agent's primary concern is on the unobserved state s, the possibility and magnitude of the bias term cannot be ignored. Consequently, 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 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. As mentioned in the Introduction, the parameter $\sigma_{b,0}^2$ represents the agent's uncertainty about the unknown bias. In particular, a higher value of $\sigma_{b,0}^2$ indicates that the agent is more unsure about the bias term. The variance of s (i.e., $\sigma_{s,0}^2$) can be interpreted in a similar fashion. 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.[24] A positive value of ρ_0 means that, in the agent's initial belief, the bias term tends to exaggerate or complement the effect of the unknown state, whereas a negative value means that b tends to contradict the effect of s. We emphasise at the outset that a nonzero value of ρ_0 is not essential in generating opinion polarisation between agents. A negative value of ρ_0, nonetheless, opens up a myriad of intriguing possibilities which we will discuss fully in later sections.

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. The covariances between (s, b) and any individual signal m_t are therefore constant over time and are given by

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

[24] Under these assumptions, the determinant of $\mathbf{\Sigma}_0$ is strictly positive, hence the matrix is invertible. A strictly positive determinant, together with a strictly positive trace $\left(\sigma_{s,0}^2 + \sigma_{b,0}^2 > 0\right)$, implies that the eigenvalues of $\mathbf{\Sigma}_0$ are strictly positive. Since $\mathbf{\Sigma}_0$ is symmetric, this implies positive definiteness.

for all t. 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

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

These conditions make clear 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 because $\lambda_0 + \theta_0 = var\,(s + b)$ 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
$\sigma_{s,0}^2$	Variance of s	λ_0	Covariance between s and m_t
$\sigma_{b,0}^2$	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 $\mathbf{x} = [s \;\; b]'$ 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 will take the form of a bivariate normal distribution with mean vector $\widehat{\mathbf{x}}_t$, which contains the agent's updated estimate of $\mathbf{x}$, i.e.,

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

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 determined by*

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

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

$$\widehat{\omega}_t = \frac{\omega_0\sigma^2_\varepsilon - \left(1 - \rho_0^2\right)\sigma^2_{s,0}\sigma^2_{b,0}t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}, \tag{3}$$

$$\widehat{s}_t = s_0 + \kappa_t\left(\overline{m}_t - s_0 - b_0\right), \tag{4}$$

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

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

$$\kappa_t \equiv \frac{\lambda_0 t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t} \qquad \text{and} \qquad \eta_t \equiv \frac{\theta_0 t}{\sigma^2_\varepsilon + \left(\lambda_0 + \theta_0\right)t}.$$

Unless otherwise stated, all proofs can be found in the Appendix. Here we focus on the interpretation and intuition of the solution. Equations (1)-(3) 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 corresponding 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 (4)-(5) describe how the agent updates the estimates of (s, b) based on the observed signals. In each period, the agent adjusts these estimates based on the "unexpected" component of the signals, i.e., the discrepancy between the sample mean $\overline{m}_t$ and her initial combined estimate $s_0 + b_0$. The *direction* of the adjustment, however, is governed by the *sign* of κ_t and η_t.[25]

[25] The linearity of (4) and (5) greatly simplifies our main analysis. In particular, these equations state that the updated estimates $\left(\widehat{s}_t, \widehat{b}_t\right)$ can be written as a linearly combination of the initial estimates (s_0, b_0) and the average value of the realised signals $\overline{m}_t$. Diaconis and Ylvisaker (1979, Theorem 2) show that this property remains

In order to highlight the novelties brought by the unknown bias term, we compare the current model to one with *known unbiased* signals (the conventional model). In this environment, the true value of the bias term is identical to zero (i.e., $b = 0$) and, more importantly, it is a publicly known fact accepted by the agent so that $b_0 = 0$ and $\sigma_{b,0} = 0$. Setting $b = b_0 = \sigma_{b,0} = 0$ in the above equations yields $\widehat{b}_t = \widehat{\sigma}^2_{b,t} = \widehat{\omega}_t = 0$ for all t. Equations (1) and (4) then simplify to become

$$\widehat{\sigma}^2_{s,t} = \frac{\sigma^2_\varepsilon \sigma^2_{s,0}}{\sigma^2_\varepsilon + \sigma^2_{s,0} t}, \tag{6}$$

$$\widehat{s}_t = s_0 + \frac{\sigma^2_{s,0} t}{\sigma^2_\varepsilon + \sigma^2_{s,0} t} \left(\overline{m}_t - s_0 \right). \tag{7}$$

Equation (6) makes clear that, when the signals are known to be unbiased, the error variance of the agent's estimate will converge to zero as t (the number of observed signals) approaches infinity. This means after observing a sufficiently long stream of unbiased signals, the agent can filter out the noises and identify the true value of s. This well-known result from the conventional model is no longer valid once we introduce an unknown bias term. This can be seen by studying the asymptotic properties of $\{\widehat{\sigma}^2_{s,t}\}_{t=0}^{\infty}$ and $\{\widehat{\sigma}^2_{b,t}\}_{t=0}^{\infty}$. Our next result states that both sequences are monotonically decreasing and convergent. This means the precision of the agent's estimates are improving after each update. Thus, learning occurs even though the signals are confounded by another unknown parameter. But instead of converging to zero, both $\widehat{\sigma}^2_{s,t}$ and $\widehat{\sigma}^2_{b,t}$ converge to a strictly positive value which is determined by the elements in $\mathbf{\Sigma}_0$. This means the agent remains uncertain about the true value of s and b even after observing an infinite stream of signals.

Proposition 2 *In the presence of an unknown bias term, both* $\{\widehat{\sigma}^2_{s,t}\}_{t=0}^{\infty}$ *and* $\{\widehat{\sigma}^2_{b,t}\}_{t=0}^{\infty}$ *are monotonically decreasing sequences that converge to the same limit, which is strictly positive. Formally,*

$$\lim_{t\to\infty} \widehat{\sigma}^2_{s,t} = \lim_{t\to\infty} \widehat{\sigma}^2_{b,t} = \frac{\left(1-\rho_0^2\right) \sigma^2_{s,0}\sigma^2_{b,0}}{\sigma^2_{s,0} + \sigma^2_{b,0} + 2\omega_0} > 0. \tag{8}$$

Even though the agent cannot identify the true value of s and b separately, she can still learn the true value of their sum, $s + b$, in the long run. This follows from Proposition 3 which states

valid if the signals are independently drawn from a distribution in the exponential family (which includes normal, exponential, gamma, binomial and Poisson among others) and a conjugate prior is used as the initial belief (see their remark on p.274).

that the combined estimate $\widehat{s}_t + \widehat{b}_t$ will converge in probability to the true value $s + b$ as the number of observations increases. Intuitively, this means that when t is sufficiently large, the chance that $\widehat{s}_t + \widehat{b}_t$ is somewhat different from $s + b$ can be made arbitrarily small.

Proposition 3 *In the presence of an unknown bias term, the combined estimate* $\widehat{s}_t + \widehat{b}_t$ *converges in probability to the true sum, i.e., for any* $\varepsilon > 0$,

$$\lim_{t \to \infty} \Pr\left[\left|\widehat{s}_t + \widehat{b}_t - s - b\right| \geq \varepsilon\right] = 0.$$

One implication of Proposition 3 is that in the long run, the covariance between s and b in the agent's updated belief (i.e., $\widehat{\omega}_t$) must be strictly negative. This is because, after observing a sufficiently large number of signals, the agent will have a good grasp of the true combined value $s + b$. Any new information 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. This can be shown formally using the results in Proposition 4, which concerns the dynamic properties of $\widehat{\omega}_t$. Starting from any initial value $\widehat{\omega}_0$ that is consistent with $\sigma_{s,0} > 0$, $\sigma_{b,0} > 0$ and $\rho_0 \in (-1, 1)$, the sequence $\{\widehat{\omega}_t\}_{t=0}^{\infty}$ generated by (3) is monotonic and convergent. The limit value $\widehat{\omega}_\infty$ is strictly negative as shown in (9). It follows that the sequence is monotonically decreasing [resp., monotonically increasing] towards $\widehat{\omega}_\infty$ if the initial value $\widehat{\omega}_0$ is greater [resp., lower] than $\widehat{\omega}_\infty$. Equations (8) and (9) together imply

$$\lim_{t \to \infty} \widehat{\rho}_t = \lim_{t \to \infty} \left(\frac{\widehat{\omega}_t}{\widehat{\sigma}_{s,t}\widehat{\sigma}_{b,t}}\right) = -1.$$

Proposition 4 *In the presence of an unknown bias term, the sequence* $\{\widehat{\omega}_t\}_{t=0}^{\infty}$ *is convergent with limit*

$$\widehat{\omega}_\infty \equiv \lim_{t \to \infty} \widehat{\omega}_t = -\frac{\left(1 - \rho_0^2\right) \sigma_{s,0}^2 \sigma_{b,0}^2}{\sigma_{s,0}^2 + \sigma_{b,0}^2 + 2\omega_0} < 0. \tag{9}$$

In addition, $\widehat{\omega}_{t+1} \gtreqless \widehat{\omega}_t$ *for all* $t \geq 0$ *if and only if* $\omega_0 \lesseqgtr \widehat{\omega}_\infty$.

Another interesting observation about the correlation between s and b is that even if they are presumed to be uncorrelated in the initial belief (i.e., $\rho_0 = 0$), the mere uncertainty about the bias term (i.e., $\sigma_{b,0} > 0$) is enough to generate a nonzero correlation in *all* subsequent periods (i.e., $\widehat{\rho}_t \neq 0$ for all t). To see this formally, we use Equations (1)-(3) to derive

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

When evaluated at $\rho_0 = 0$, this simplifies to become

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

which is strictly negative for all t.[26] Note that this result holds regardless of the true value of b and the initial estimate b_0. In other words, uncertainty begets correlation even if the signals are truly unbiased and the agent's initial estimate is correct on this, i.e., $b = b_0 = 0$.

The rest of this section concerns the dynamic and asymptotic properties of $\widehat{s}_t$ and $\widehat{b}_t$. Unlike Equations (1)-(3), the dynamic system for $\widehat{s}_t$ and $\widehat{b}_t$ 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 one way to interpret (4) and (5) is that, over time as new information becomes available, the revised estimates $\widehat{s}_t$ and $\widehat{b}_t$ will fluctuate around the initial estimates s_0 and b_0, respectively. Note also that the sign of κ_t and η_t are solely determined by the sign of λ_0 and θ_0, respectively.

Suppose the agent observes a higher-than-expected sample mean at time t so that $\overline{m}_t > s_0 + b_0$. If the agent perceives the signals as positively correlated with the hidden state, i.e., $\lambda_0 > 0$, then she will interpret the higher-than-expected sample average as an indication that the true state s is greater than her initial estimate s_0. This will motivate the agent to revise her estimate upward, i.e., $\widehat{s}_t > s_0$.[27] We refer to such agent as a *conventional learner*. Contrarily, if the agent presumes a negative correlation between the hidden state and the signals, i.e., $\lambda_0 < 0$, then θ_0 must be strictly positive [recall that λ_0 and θ_0 cannot both be negative]. In this case, the agent will attribute the higher-than-expected sample mean to a higher-than-expected value of b instead. Since a negative correlation between s and b (i.e., $\rho_0 < 0$) is a necessary condition for $\lambda_0 < 0$, a high value of b suggests that the true value of s is likely to be low. This will motivate the agent

[26] This result continues to hold if $\rho_0 \in (-1, 0)$. If $\rho_0 \in (0, 1)$, then there exists a unique value $t^* \geq 0$ such that $\widehat{\rho}_t \gtrless 0$ for $t \lessgtr t^*$.

[27] Note that the comparison here is between $\widehat{s}_t$ and s_0, not between $\widehat{s}_t$ and $\widehat{s}_{t-1}$. It is possible to rewrite (4) and (5) as a sequential update from $\left(\widehat{s}_{t-1}, \widehat{b}_{t-1}\right)$ to $\left(\widehat{s}_t, \widehat{b}_t\right)$. Equations (4) and (5), however, provide an easier route to derive the asymptotic results in Proposition 5.

to revise her estimate for s downward, i.e., $\widehat{s}_t < s_0$. We refer to this kind of agent as a *defiant learner*. If $\lambda_0 = 0$, then in the agent's initial belief the signals are uncorrelated with (and hence uninformative about) the hidden state. It follows that $\kappa_t = 0$ at all times and the agent will never adjust her initial estimate so that $\widehat{s}_t = s_0$ for all t. We will ignore this uninteresting case in the following analysis.

The above description presents another major difference between the current model and the conventional model. As Equation (7) makes clear, an agent learning from unbiased signals will always adjust her estimate in the *same direction* as the unexpected component in the signals, so that

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

In other words, there are only conventional learners in the conventional model. But in the current model, the adjustment can go in either direction depending on the sign of λ_0, so that

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

So there are both conventional and defiant learners in our model.

We can also evaluate the average behaviour of $\widehat{s}_t$ and $\widehat{b}_t$ from an outside observer's perspective. Consider an observer who knows the true value of (s, b), the statistical properties of the noise process and the agent's initial belief $\mathbf{N}(\mathbf{x}_0, \mathbf{\Sigma}_0)$ [hence the observer knows how the agent updates her belief]. To this observer, the expected value of $\overline{m}_t$ is $s + b$ and the average behaviour of $\widehat{s}_t$ and $\widehat{b}_t$ are characterised by

$$E(\widehat{s}_t) = s_0 + \kappa_t (s + b - s_0 - b_0), \tag{10}$$

$$E\left(\widehat{b}_t\right) = b_0 + \eta_t (s + b - s_0 - b_0). \tag{11}$$

The dynamics of $E(\widehat{s}_t)$ and $E\left(\widehat{b}_t\right)$ are entirely driven by the time-varying coefficients κ_t and η_t, respectively. In particular,

$$\frac{d\kappa_t}{dt} = \frac{\lambda_0 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (\lambda_0 + \theta_0) t} \gtrless 0 \qquad \text{if and only if} \qquad \lambda_0 \gtrless 0,$$

and

$$\frac{d\eta_t}{dt} = \frac{\theta_0 \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (\lambda_0 + \theta_0) t} \gtrless 0 \qquad \text{if and only if} \qquad \theta_0 \gtrless 0.$$

This means if $\lambda_0 > 0$ [resp., $\theta_0 > 0$], then κ_t [resp., η_t] is strictly positive and increasing in value over time. Suppose the agent underestimates the true combined value initially, i.e., $s+b > s_0+b_0$, and both λ_0 and θ_0 are strictly positive. Then from the observer's point of view, the average value of both $\widehat{s}_t$ and $\widehat{b}_t$ will be increasing over time. But if either $\lambda_0 < 0$ or $\theta_0 < 0$, then $E\left(\widehat{s}_t\right)$ and $E\left(\widehat{b}_t\right)$ will move in opposite directions.

As more and more signals are realised, the variance of the noises will become less and less important in the learning process. Hence, the limit of κ_t and η_t are given by

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

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 (4) and (5). The results are shown in Proposition 5.

Proposition 5 *In the presence of an unknown bias term,* $\widehat{s}_t$ *and* $\widehat{b}_t$ *converge in probability to* $\widehat{s}_\infty$ *and* $\widehat{b}_\infty$*, respectively, as* $t \to \infty$*. The limits are given by*

$$\widehat{s}_\infty \equiv s_0 + \kappa_\infty \left(s + b - s_0 - b_0\right), \tag{12}$$

$$\widehat{b}_\infty \equiv b_0 + \left(1 - \kappa_\infty\right)\left(s + b - s_0 - b_0\right). \tag{13}$$

The coefficient κ_∞ can be interpreted as follows: Since the agent's initial belief is independent of ε_t, we can write

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

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 (12) and (13) encompass two types of learning behaviour that are incompatible with the conventional model. In the first, learning from an infinite sequence of signals does not bring the agent any closer to the truth. Instead, the agent's long-run estimate of 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.*

Note that 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. Opinion reversal is the subject of Section 2.4. For reasons explained earlier, we do not consider the case when $\lambda_0 = 0$ in the following analysis.

2.3 Misguided Learning

Using (12), it is straightforward to show that $\widehat{s}_\infty > s_0 > s$ happens if and only if

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

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

$$\kappa_\infty (s + b - s_0 - b_0) < 0 \qquad \text{and} \qquad s_0 < s. \tag{15}$$

It suffice to focus on the intuition behind (14) because (15) can be explained symmetrically. Suppose $\lambda_0 > 0$ so that $\kappa_\infty > 0$. This means the agent is a conventional learner. Then the conditions in (14) imply $s + b > s_0 + b_0$ and hence $b - b_0 > s_0 - s > 0$. This means even though the agent overestimates the value of s in the initial estimate, this is counteracted by a substantial underestimation in the bias term so that the initial combined estimate falls below the true value. Over time as new signals arrive, the sample mean $\overline{m}_t$ will get closer and closer to $s + b$ by the law of large numbers. This means the agent will almost always observe $\overline{m}_t > s_0 + b_0$ in the long run. This will induce any conventional learner to maintain a long-run estimate for s that is higher than her initial assessment so that $\widehat{s}_\infty > s_0$.[28]

The above description makes clear that a significant misalignment between b and b_0 is necessary for misguided learning to occur among conventional learners. Interestingly, this is not necessary for a defiant learner [i.e., when $\lambda_0 < 0$]. Suppose, unbeknownst to the agent, the signals are truly unbiased. Suppose the agent's initial estimate b_0 is correct (i.e., $b_0 = b = 0$), but doubt lingers 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 $\overline{m}_t < s_0 + b_0$. Since $\lambda_0 = Cov(s, m_t) < 0$, the agent will take this as an indication that the true value of s is higher than her initial estimate and revise

[28] It is important to note that $s_0 > s$ is just an initial condition, it has no bearing on how the agents will respond to the new information [which is controleld by $\kappa_\infty (s + b - s_0 - b_0)$]. It follows that if $\kappa_\infty (s + b - s_0 - b_0) > 0$, then $\widehat{s}_\infty > s^*$ for any s^* that satisfies $s_0 \geq s^*$. We exploit this argument when considering opinion divergence at the interpersonal level.

her estimate upward. This will again sustain a long-run estimate which is further away from the truth, i.e., $\widehat{s}_\infty > s_0 > s$.

We now provide some simulated examples that can help visualise the above points. 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 (14), 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 $\lambda_0 = -0.0342$, 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 Figures 1 and 2 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.[29]

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
$\sigma_{s,0}^2$	0.1	0.1	0.1	0.5	0.1
$\sigma_{s,0}^2$	0.5	0.5	0.5	0.1	0.5
σ_ε^2	0.03	0.03	0.03	0.03	0.03
λ_0	0.10	-0.045	0.10	-0.045	0.355
θ_0	0.50	0.355	0.50	0.355	-0.045
κ_∞	0.167	-0.147	0.167	-0.147	1.147

[29]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.

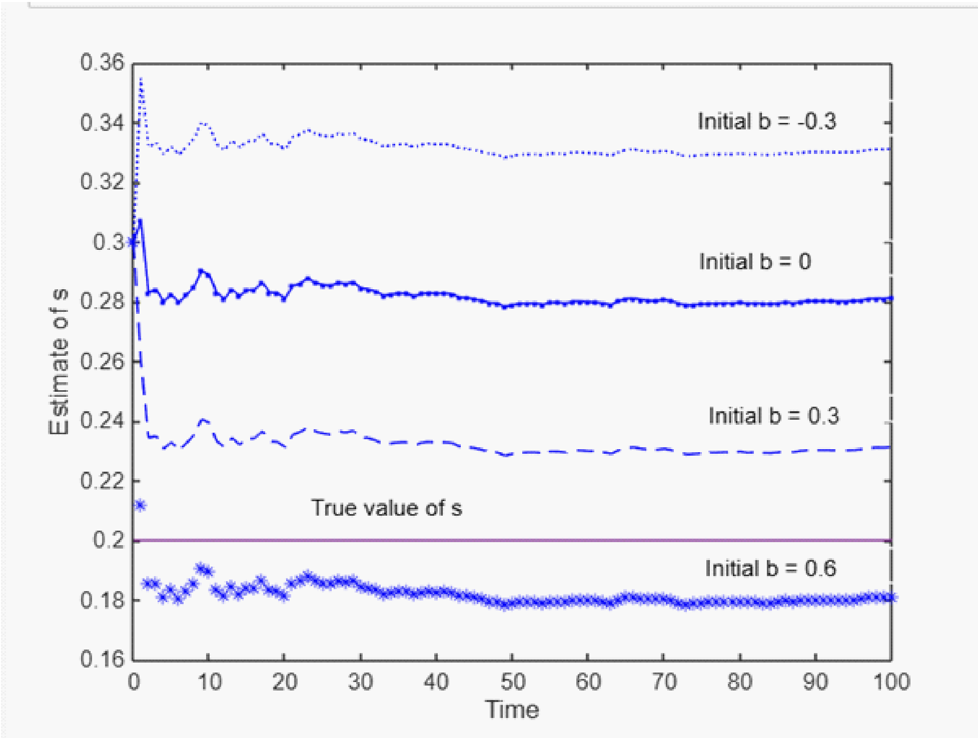

Figure 1

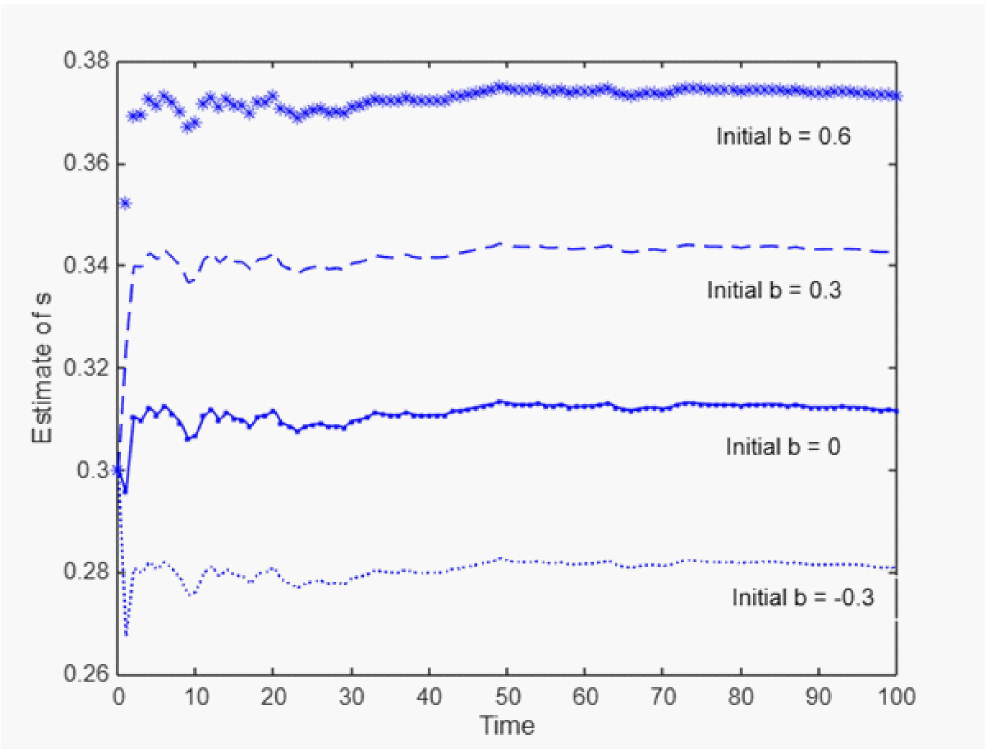

Figure 2

In Figure 1, misguided learning happens in the uppermost sequence which corresponds to the case when $b_0 = -0.3$, so that $s_0 + b_0 = 0 < s + b$. The limit value $\widehat{s}_\infty$ is one-third which confirms the ordering $\widehat{s}_\infty > s_0 > s$. For the second and third sequence (which corresponds to $b_0 = 0$ and $b_0 = 0.3$, respectively), the first inequality in (14) is violated. As a result, the limit value $\widehat{s}_\infty$ is sandwiched between s_0 and s. Learning is incomplete in the sense that $\widehat{s}_\infty \neq s$ but it is closer to the truth than the initial estimate. The fourth sequence in Figure 1 is an example of opinion reversal which we will discuss in the next section.

The only difference between the results in Figure 1 and those in Figure 2 lies in the value of ρ_0. 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. In particular, the third sequence (which corresponds to $b_0 = 0$) confirms that the mere suspicion of a hidden bias in the signals (i.e., $\sigma_{b,0} > 0$) is enough to generate misguided learning when $\lambda_0 < 0$.

2.4 Opinion Reversal

Using (12), it can be readily shown that $\widehat{s}_\infty > s > s_0$ happens if and only if[30]

$$\kappa_\infty (s + b - s_0 - b_0) > s - s_0 > 0. \tag{16}$$

If both λ_0 and θ_0 are strictly positive so that $\kappa_\infty \in (0, 1)$, then (16) is equivalent to

$$b - b_0 > \frac{1 - \kappa_\infty}{\kappa_\infty} (s - s_0) > 0. \tag{17}$$

This has two meanings: First, the agent underestimates both s and b in her initial belief so that, in the long run, the sample mean of realised signals 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 eventually converge to some values that are above the initial estimates so that $\widehat{s}_\infty > s_0$ 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 (17) also mean that the agent has significantly and sufficiently underestimated the bias term in her initial estimate. This ensures that the long-run estimate $\widehat{b}_\infty$ is higher than b_0 but still falls short of the true value b. Since $\widehat{s}_\infty + \widehat{b}_\infty = s + b$, it follows that $\widehat{s}_\infty > s$. A simulated example of this is shown Figure 3 based on the parameter values listed in Table 2.[31] Another example is the

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

[31] 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.

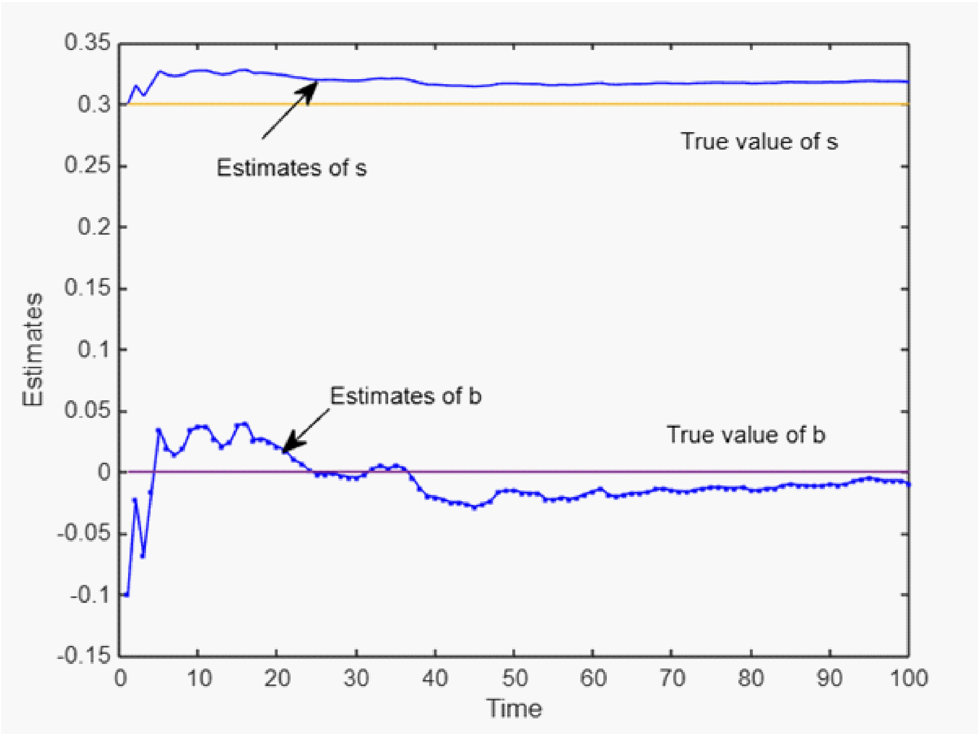

Figure 3

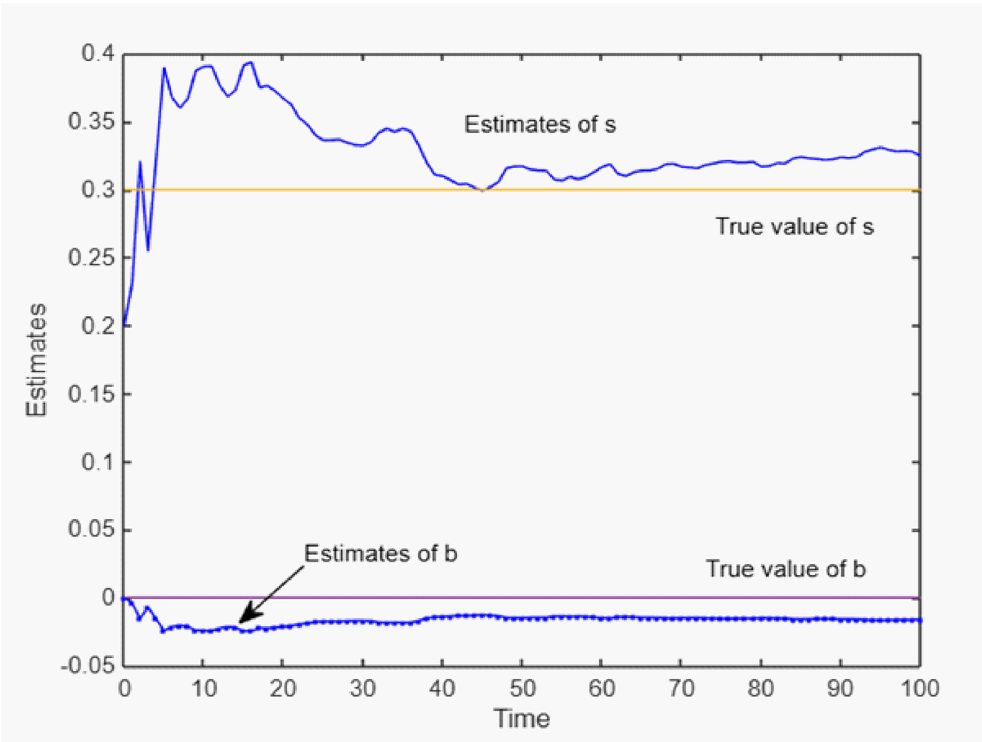

Figure 4

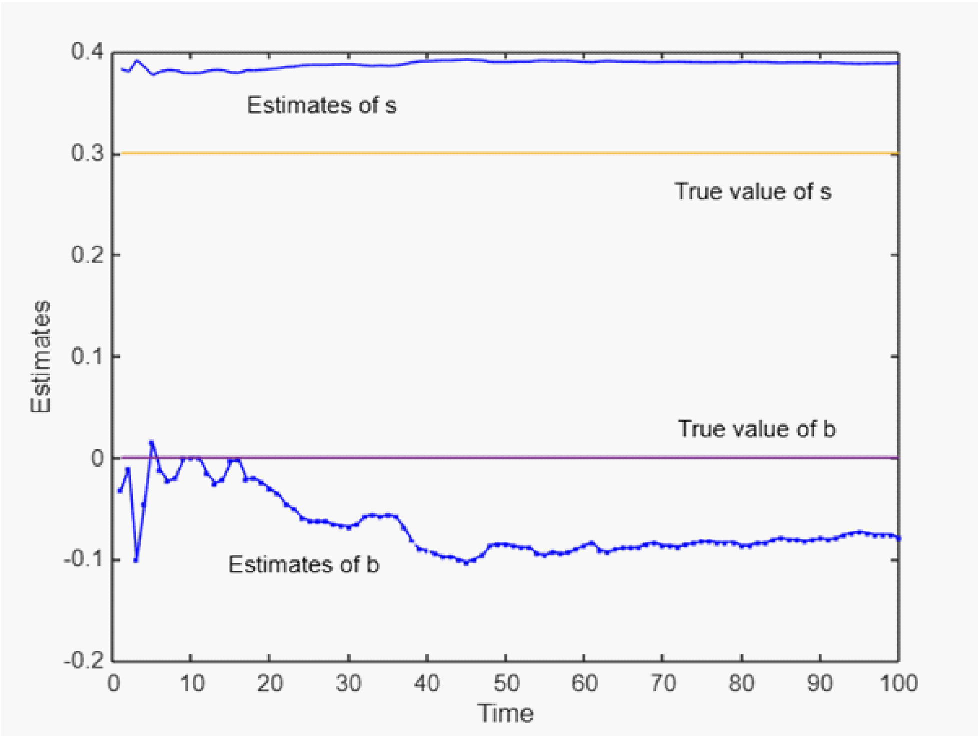

Figure 5

lowermost sequence depicted in Figure 1 (the one with $b_0 = 0.6$). In this case, opinion reversal occurs in the form of $\widehat{s}_\infty < s < s_0$.

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 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 (16) 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 raise the value of $\widehat{s}_\infty$ above s_0 but lower $\widehat{b}_\infty$ below b_0. The fact that $\eta_\infty > 1$ means that the downward adjustment in $\widehat{b}_\infty$ is substantial, so that $\widehat{b}_\infty < b$. This in turn ensures $\widehat{s}_\infty > s$. Figure 5 provides a simulated example for this case.

3 Interpersonal Disagreement

We now compare the learning behaviour of two individuals, referred to as Agent 1 and 2, in the above model. Both agents share the same knowledge about the information channel and observe the same sequence of public signals $\{m_1, m_2, ...\}$. The two agents, 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 (1)-(5). We do not consider any private information in order to single out the effects of the bias in the public signals.

We will refer to $\widehat{s}_t^{\dagger}$ and $\widehat{s}_t^{\ddagger}$ as the agents' opinion about the hidden state s after observing the same history of signals $\mathbf{m}^t$. 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

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

for some t, 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.,

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

Permanent opinion divergence is said to happen when the above inequality holds in the long run, i.e.,

$$\widehat{s}_{\infty}^{\dagger} > s_0^{\dagger} \geq s_0^{\ddagger} > \widehat{s}_{\infty}^{\ddagger}. \tag{19}$$

The main difference between (18) and (19) 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 (18) characterise a random event that happens only 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, so that

$$s_0^\dagger - s_0^\ddagger \geq 0 > \widehat{s}_t^\dagger - \widehat{s}_t^\ddagger. \tag{20}$$

Permanent reversal happens when the ordering between $s_0^\dagger$ and $s_0^\ddagger$ is reversed in the long run, so that

$$s_0^\dagger - s_0^\ddagger \geq 0 > \widehat{s}_\infty^\dagger - \widehat{s}_\infty^\ddagger. \tag{21}$$

We begin with the case in which the two agents share the same covariance matrix in their initial beliefs so that $\mathbf{\Sigma}_0^\dagger = \mathbf{\Sigma}_0^\ddagger = \mathbf{\Sigma}_0$. It follows that the two share the same sequence of coefficients $\{\kappa_t\}_{t=1}^\infty$ as defined in (4) and the same limit κ_∞ as in (12). Proposition 6 provides the conditions under which (18)-(21) will occur.

Proposition 6 *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* $s_0^\dagger \geq s_0^\ddagger$.

(i) *Opinion divergence happens at time* t*, i.e., (18) holds, if and only if*

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

(ii) *Permanent opinion divergence happens, i.e., (19) holds, if and only if*

$$\lambda_0\left(s_0^\ddagger + b_0^\ddagger\right) > \lambda_0\left(s + b\right) > \lambda_0\left(s_0^\dagger + b_0^\dagger\right). \tag{22}$$

(iii) *Opinion reversal happens at time* t*, i.e., (20) holds, if and only if*

$$\kappa_t\left(b_0^\dagger - b_0^\ddagger\right) > (1-\kappa_t)\left(s_0^\dagger - s_0^\ddagger\right). \tag{23}$$

(iv) *Permanent reversal happens, i.e., (21) holds, if and only if*

$$\kappa_\infty\left(b_0^\dagger - b_0^\ddagger\right) > (1-\kappa_\infty)\left(s_0^\dagger - s_0^\ddagger\right).$$

The mechanism behind opinion divergence is similar to that for misguided learning in Section 2.3. Suppose $\lambda_0 > 0$ so that $\kappa_t > 0$ for all t, including the limit. In this case, both agents are conventional learners. Then opinion divergence happens at time t if and only if $s_0^\ddagger + b_0^\ddagger >$

$\overline{m}_t > s_0^\dagger + b_0^\dagger$. This means, even though Agent 1's initial estimate of the hidden state is no less than Agent 2's, her initial estimate of the bias term must be sufficiently lower so that Agent 2's combined estimate $s_0^\ddagger + b_0^\ddagger$ is strictly greater than that of Agent 1. It follows that whenever $\overline{m}_t$ falls between $\left(s_0^\ddagger + b_0^\ddagger\right)$ and $\left(s_0^\dagger + b_0^\dagger\right)$, Agent 1 will adjust her estimate above $s_0^\dagger$ and Agent 2 will lower hers below $s_0^\ddagger$, hence widening the initial disagreement. Similarly to misguided learning, the misalignment between $b_0^\dagger$ and $b_0^\ddagger$ is not needed when $\lambda_0 < 0$ [i.e., when both agents are defiant learners]. 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, Agent 1 will take this as an indication that the true value of s is higher than her initial estimate. As a result, Agent 1 will adjust her estimate above $s_0^\dagger$ while Agent 2 will do the opposite, again widening the initial disagreement. Part (ii) of Proposition 6 can be interpreted in the same fashion by replacing $\overline{m}_t$ with $(s + b)$. In this case, the misguided learning mechanism works on both agents but in opposite directions. In particular, it brings the agents further away from the other agent's initial estimate for s (see Footnote 28). As an example, we can interpret the uppermost and the lowermost sequences in Figure 1 as two agents who share the same initial estimate for s, i.e., $s_0^\dagger = s_0^\ddagger$ but different values of b_0 under $\lambda_0 > 0$. In this case, Agent 1 with $b_0^\dagger = -0.3$ matches the criterion $s + b = 0.2 > s_0^\dagger + b_0^\dagger = 0$, while Agent 2 with a significantly higher value of b_0 (i.e., $b_0^\dagger = 0.6$) satisfies $s_0^\ddagger + b_0^\ddagger = 0.9 > s + b$. The resulting sequences confirm the prediction in part (ii) of Proposition 6.

We now turn to opinion reversal. The conditions for opinion reversal and opinion divergence differ in two material ways. First, the former is independent of the realised signals. To see why this is the case, we first rewrite (4) as

$$\widehat{s}_t = s_0 - \kappa_t \left(s_0 + b_0\right) + \kappa_t \overline{m}_t, \tag{24}$$

where the last term captures the agent's response to the signals. If two agents share the same coefficient κ_t, then the same signals will affect their opinions in exactly the same way. As a result, the difference $\widehat{s}_t^\dagger - \widehat{s}_t^\ddagger$ is independent of $\overline{m}_t$.[32] Note that all the terms in condition (23) are nonrandom and the condition is linear in κ_t (which is the only variable that changes over time). Thus, it is no surprise that opinion reversal can only happen once, if at all. This gives rise to the second major difference between opinion divergence and opinion reversal: once opinion reversal

[32] Oppinion divergence, on the other hand, concerns the directions in which the agents revise their estimates. This is controlled by $\kappa_t \left(\overline{m}_t - s_0 - b_0\right)$ along the convergent path. Hence, the condition in part (i) is dependent on $\overline{m}_t$.

happens at some time $t < \infty$ it is permanent. This result is formally stated in Proposition 7.

Proposition 7 *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, i.e.,* $\widehat{s}_t^\ddagger > \widehat{s}_t^\dagger$, *then* $\widehat{s}_{t+k}^\ddagger > \widehat{s}_{t+k}^\dagger$ *for all* $k > 0$, *including* $\widehat{s}_\infty^\ddagger > \widehat{s}_\infty^\dagger$.

To see the intuition behind these results, consider the following equation which is derived from (24),

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

This breaks down the differences in the agents' opinion into two parts. The first term is due to initial disagreement. The second term captures the differences incurred in the revision at time t. Suppose $\kappa_t > 0$ and $s_0^\dagger + b_0^\dagger > s_0^\ddagger + b_0^\ddagger$. The latter means that Agent 1 has higher expectation about the signals than Agent 2. Over time, it is more likely that $s_0^\dagger + b_0^\dagger > \overline{m}_t$ 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. Note also that if $\kappa_t > 0$, then its magnitude will be increasing over time. This means any downward revision that happens in a later date will be greater in magnitude than one that happens earlier. This makes it possible for Agent 2 to catch up and eventually overtake Agent 1. The case of $\kappa_t < 0$ can be explained similarly.

In the second part of the analysis, we assume that the two agents share the same initial estimates but have different covariance matrices in their initial beliefs, i.e., $\mathbf{x}_0^\dagger = \mathbf{x}_0^\ddagger = \mathbf{x}_0$ and $\mathbf{\Sigma}_0^\dagger \neq \mathbf{\Sigma}_0^\ddagger$.[33] Using (4), it is immediate to see that $\widehat{s}_t^\dagger > \widehat{s}_t^\ddagger$ at any time t if and only if

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

The implications are as follows: Suppose at some time t, the two agents observe a summary statistic $\overline{m}_t$ that is greater than their combined estimate, i.e., $\overline{m}_t > s_0 + b_0$. Then there are three possible scenarios in which Agent 1's revised estimate for s is higher than Agent 2's. In the first one, both $\lambda_0^\dagger$ and $\lambda_0^\ddagger$ are positive but $\lambda_0^\dagger > \lambda_0^\ddagger$. In this case, both agents respond to the higher-than-expected 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 surprise which leads to a higher revised estimate. In the second scenario, both $\lambda_0^\dagger$ and $\lambda_0^\ddagger$ are negative but $\lambda_0^\dagger$ is closer to zero. In this case, both agents revise their opinions about s downward but Agent 1 is less responsive to the surprise. In the final scenario,

[33] Since the two agents start with the same initial estimate s_0, the possibility of opinion reversal is muted in this scenario.

$\lambda_0^\dagger$ is positive but $\lambda_0^\ddagger$ is negative. As a result, Agent 1 will revise her estimate upward while Agent 2 will do the opposite, which results in opinion divergence. Permanent opinion divergence can be similarly explained and characterised by replacing $\overline{m}_t$ with $(s+b)$. These results are summarised below. The proof follows immediately from (4) and (12), hence it is omitted.

Proposition 8 *Consider two agents with initial beliefs* $\mathbf{N}\left(\mathbf{x}_0, \mathbf{\Sigma}_0^\dagger\right)$ *and* $\mathbf{N}\left(\mathbf{x}_0, \mathbf{\Sigma}_0^\ddagger\right)$, *where* $\mathbf{\Sigma}_0^\dagger \neq \mathbf{\Sigma}_0^\ddagger$.

(i) *Opinion divergence happens at time t, i.e.,* $\widehat{s}_t^\dagger > s_0 > \widehat{s}_t^\ddagger$, *if and only if*

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

(ii) *Permanent opinion divergence happens, i.e.,* $\widehat{s}_\infty^\dagger > s_0 > \widehat{s}_\infty^\ddagger$, *if and only if*

$$\lambda_0^\dagger (s + b - s_0 - b_0) > 0 > \lambda_0^\ddagger (s + b - s_0 - b_0).$$

4 Aggregate Disagreement

We now extend the above analysis to a large population of agents who observe the same sequence of public signals $\{m_t\}_{t=1}^\infty$ but have heterogeneous initial beliefs about (s, b). Each agent updates their estimates according to (1)-(5) and the asymptotic value of their estimates are determined by (12). Our main focus in this section is whether continuous exposure to the potentially biased signals will exacerbate disagreement at the aggregate level.

Let $var(s_0)$ and $var(b_0)$ denote, respectively, the variance of s_0 and b_0 in the cross-sectional distribution of initial beliefs.[34] These measure the dispersion of s_0 and b_0 across agents before they are exposed to the signals. To simplify the analysis, we assume that s_0 and b_0 are uncorrelated across agents, i.e., $Cov(s_0, b_0) = 0$. As in Section 3, we first consider the case in which agents share the same covariance matrix $\mathbf{\Sigma}_0$ in their initial beliefs, so that $\{\lambda_0, \theta_0, \kappa_\infty\}$ are the same across individuals. It then follows from (12) that the dispersion of $\widehat{s}_\infty$ within the population is given by

$$var(\widehat{s}_\infty) = (1 - \kappa_\infty)^2 \, var(s_0) + \kappa_\infty^2 var(b_0). \tag{25}$$

The above equation encompasses the conventional model as a special case. If the signals are

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

truly unbiased (i.e., $b = 0$) and this fact is universally accepted by all agents in the society, then there is no pre-existing disagreement about the bias term so that $var(b_0) = 0$, and there is no uncertainty regarding b so that $\sigma_{b,0}^2 = 0$ for all agents. The latter implies $\kappa_\infty = 1$. In the long run, all agents will learn the true value of the hidden state and aggregate disagreement disappears, i.e., $var(\widehat{s}_\infty) = 0$. This result holds regardless of $var(s_0)$ because $\kappa_\infty = 1$.

Suppose now within the society all agents share the same initial estimate s_0, so that $var(s_0) = 0$. But they have different views about the biasedness of the signals, so that $var(b_0) > 0$. In this case, (25) becomes

$$var(\widehat{s}_\infty) = \kappa_\infty^2 var(b_0) = \left[\frac{\sigma_{s,0}^2 + \rho_0\sigma_{s,0}\sigma_{b,0}}{\sigma_{s,0}^2 + \sigma_{b,0}^2 + \rho_0\sigma_{s,0}\sigma_{b,0}}\right]^2 var(b_0).$$

The complicated expression on the right helps us identify the necessary ingredients in generating aggregate disagreement in the long run. Provided that $var(b_0) > 0$, $var(\widehat{s}_\infty) > 0$ if and only if $\lambda_0 \neq 0$. Recall that $\lambda_0 = 0$ [and hence $\kappa_t = 0$ for all t] means the agents never revise their estimates for s. This happens when either $\sigma_{s,0} = 0$ or $\sigma_{s,0} + \rho_0\sigma_{b,0} = 0$. The main message of this observation is that aggregate disagreement can be easily spawned when people disagree about the biasedness of the public signals. In this type of environment, 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., $var(s_0) > 0$, (ii) they share the same b_0 so that $var(b_0) = 0$ [could be $b = b_0 = 0$], but (iii) they are not entirely sure about b_0, i.e., $\sigma_{b,0} > 0$. It follows from (25) that

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

Even though there is no initial disagreement in b_0, the agents' uncertainty about the unknown bias can still generate aggregate disagreement in $\widehat{s}_\infty$. This happens 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., $var(\widehat{s}_\infty) < var(s_0)$, if and only if

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

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 worsen the disagreement within the population, i.e., $var(\widehat{s}_\infty) > var(s_0)$. Note that $\kappa_\infty < 0$ means that all agents in the population are defiant learners.

In general,

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

The main takeaway is that if $\kappa_\infty < 0$ and the agents hold diverse views about the bias term [so that $var(b_0)$ is strictly positive but however small], then being exposed to the potentially biased public signals will widen the initial disagreement within the population. This is true regardless of the extent of initial disagreement $var(s_0)$.

As an illustration of these points, we construct a hypothetical population of 200 agents whose initial estimates (s_0, b_0) are equally spaced on the unit circle centred at the true parameter values $(s, b) = (0.2, 0)$.[35] 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 (s_0, b_0). The resulting values are located on a straight line $L : \widehat{s}_\infty + \widehat{b}_\infty = s + b$ that cuts through the circle. The auxiliary dashed lines link the initial value (s_0, b_0) 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 disagreement. This is confirmed by the reduction in the variances from $var(s_0) = 0.505$ to $var(\widehat{s}_\infty) = 0.365$. 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 $\widehat{s}_\infty$ 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 $var(s_0) = 0.505$ to $var(\widehat{s}_\infty) = 0.675$. Similar to Figure 6, points A and B serve as an example of permanent opinion divergence, while C and D serve as an example of permanent opinion reversal.

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

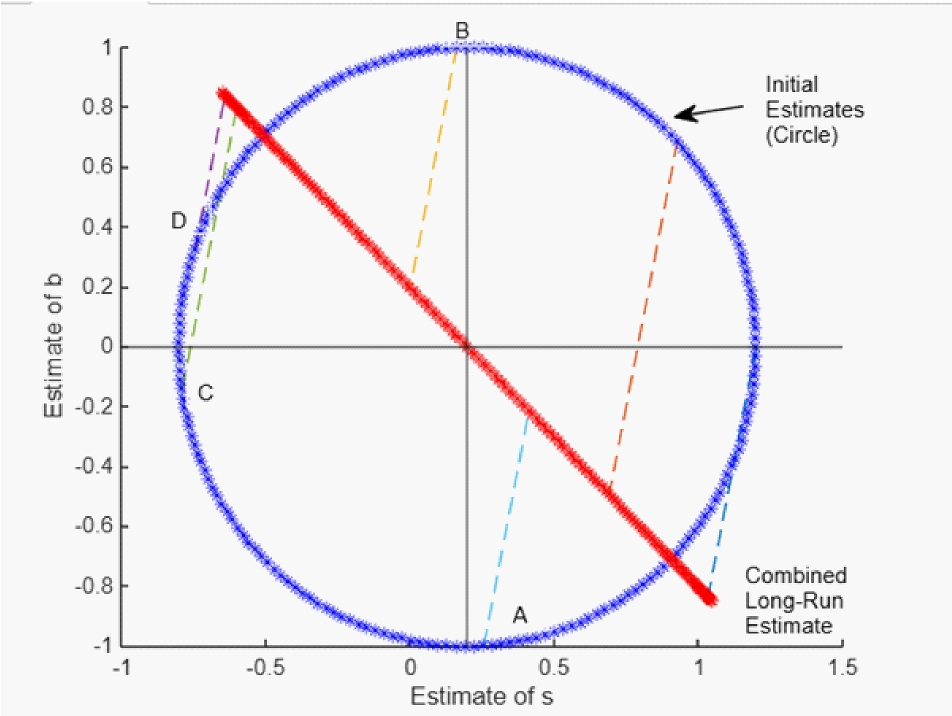

Figure 6

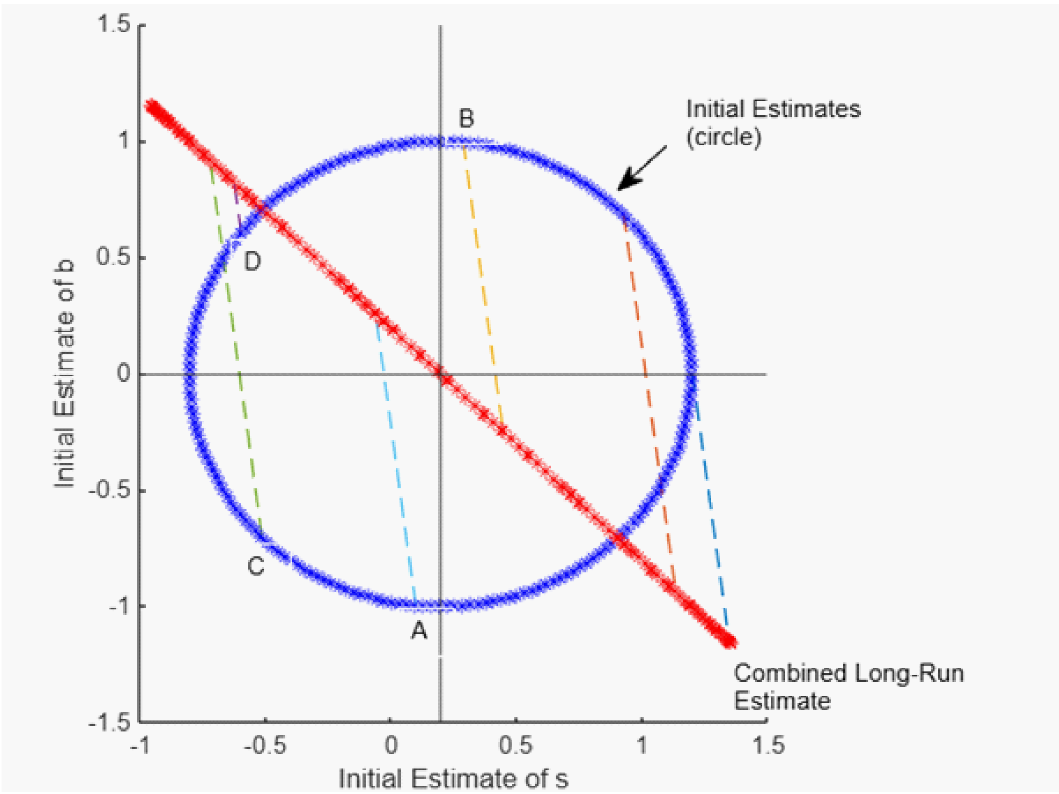

Figure 7

Finally, we consider a society in which all agents share the same initial estimates, i.e., $var(s_0) = var(b_0) = 0$, but have different covariance matrices in their initial beliefs. In particular, the value of κ_∞ varies among them so that $var(\kappa_\infty) > 0$ within the population. Equation (12) then implies

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

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., $var(\widehat{s}_\infty) = 0$, regardless of $var(\kappa_\infty)$. 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 the equation in (4), which is

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

By the law of large number, the sample average $\overline{m}_t$ will eventually converge 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., $\overline{m}_t - s_0 - b_0 \xrightarrow{p} 0$. As a result, any disagreement in κ_t will be 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, disagreement in κ_∞ will pave the way for aggregate disagreement in $\widehat{s}_\infty$. In particular, the greater the discrepancy between $s + b$ and $s_0 + b_0$, the greater the dispersion in $\widehat{s}_\infty$ across agents.

5 Concluding Remarks

In this paper we extend the canonical Bayesian learning model to allow for an unknown bias term in the public signals. What we refer to as "bias" is a highly versatile concept that can be the result of cognitive bias in human judgement, ideological or partisan bias in political information, or systemic bias in research methods. The upshot of this exercise is an easily tractable model that can generate a wide array of learning behaviour and outcomes. Given the popularity of Bayesian learning models in economics, finance, political science and other related fields, and the pervasiveness of persistent disagreement in economic, financial and political discourses, we believe this model can find a lot of applications in these areas.

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

$$\left[\begin{array}{cc} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{array}\right].$$

The probability density function of this distribution satisfies

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

where $\propto$ is the direct proportionality symbol. The terms inside the squared 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

$$h\left(x, y\right) \propto \exp\left\{-\frac{1}{2}\left[\phi_1 x^2 + \phi_2 y^2 - 2\phi_3 xy - 2\phi_4 x - 2\phi_5 y\right]\right\}, \tag{26}$$

where

$$\phi_1 \equiv \frac{1}{\left(1-\rho^2\right)\sigma_x^2}, \qquad \phi_2 \equiv \frac{1}{\left(1-\rho^2\right)\sigma_y^2}, \qquad \phi_3 \equiv \frac{\rho}{\left(1-\rho^2\right)\sigma_x\sigma_y}, \tag{27}$$

$$\phi_4 \equiv \frac{1}{\left(1-\rho^2\right)\sigma_x}\left(\frac{\mu_x}{\sigma_x} - \frac{\rho\mu_y}{\sigma_y}\right), \tag{28}$$

$$\phi_5 \equiv \frac{1}{\left(1-\rho^2\right)\sigma_y}\left(\frac{\mu_y}{\sigma_y} - \frac{\rho\mu_x}{\sigma_x}\right). \tag{29}$$

Conversely, if we are given a density function as in (26), then using the above equations we can 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 (27), we can get

$$\phi_3^2 = \frac{\rho^2}{\left(1-\rho^2\right)^2\sigma_x^2\sigma_y^2} = \rho^2\phi_1\phi_2 \Rightarrow \rho = \frac{\phi_3}{\sqrt{\phi_1\phi_2}}. \tag{30}$$

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 (27) gives

$$\sigma_x^2 = \frac{\phi_2}{\phi_1\phi_2-\phi_3^2} \qquad \text{and} \qquad \sigma_y^2 = \frac{\phi_1}{\phi_1\phi_2-\phi_3^2}. \tag{31}$$

In order to recover μ_x and μ_y, we first rewrite (28) and (29) 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, (30) and (31), we can get

$$\begin{aligned}\mu_x &= \sigma_x^2\phi_4+\rho\sigma_x\sigma_y\phi_5 = \frac{\phi_2\phi_4}{\phi_1\phi_2-\phi_3^2}+\frac{\phi_3}{\sqrt{\phi_1\phi_2}}\frac{\sqrt{\phi_1\phi_2}}{\phi_1\phi_2-\phi_3^2}\phi_5\\ &= \frac{\phi_2\phi_4+\phi_3\phi_5}{\phi_1\phi_2-\phi_3^2},\end{aligned} \tag{32}$$

and

$$\mu_y = \frac{\rho\left(\sigma_x\sigma_y\right)^{-1}}{\left(\sigma_x\sigma_y\right)^{-2}}\phi_4+\frac{\sigma_x^{-2}}{\left(\sigma_x\sigma_y\right)^{-2}}\phi_5 = \frac{\phi_3\phi_4+\phi_1\phi_5}{\phi_1\phi_2-\phi_3^2}. \tag{33}$$

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

$$\phi_1 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2}, \qquad \phi_2 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}^2}, \qquad \phi_3 \equiv \frac{\rho_0}{\left(1-\rho_0^2\right)\sigma_{s,0}\sigma_{b,0}}, \tag{34}$$

$$\phi_4 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}}\left(\frac{s_0}{\sigma_{s,0}} - \frac{\rho_0 b_0}{\sigma_{b,0}}\right), \tag{35}$$

$$\phi_5 \equiv \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}}\left(\frac{b_0}{\sigma_{b,0}} - \frac{\rho_0 s_0}{\sigma_{s,0}}\right). \tag{36}$$

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

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

$$\Rightarrow \pi\left(s,b \mid \mathbf{m}^t\right) \propto \exp\left\{-\frac{1}{2}\left[\frac{\sum_{i=1}^t \left(m_i - (s+b)\right)^2}{\sigma_\varepsilon^2} + \phi_1 s^2 + \phi_2 b^2 - 2\phi_3 sb - 2\phi_4 s - 2\phi_5 b\right]\right\}. \tag{37}$$

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 (37) 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

$$\alpha_{1,t} = \frac{t}{\sigma_\varepsilon^2} + \phi_1, \qquad \alpha_{2,t} = \frac{t}{\sigma_\varepsilon^2} + \phi_2, \tag{38}$$

$$\alpha_{3,t} = -\frac{t}{\sigma_\varepsilon^2} + \phi_3, \tag{39}$$

$$\alpha_{4,t} = \frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_4, \qquad \text{and} \qquad \alpha_{5,t} = \frac{t\overline{m}_t}{\sigma_\varepsilon^2} + \phi_5. \tag{40}$$

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 $\widehat{\mathbf{\Sigma}}_t$.

From (31), we know

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

Combining the equations in (34) and (38) gives

$$\alpha_{1,t} = \frac{t}{\sigma_\varepsilon^2} + \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2} = \frac{\sigma_{b,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{s,0}^2 t\right]}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{41}$$

$$\alpha_{2,t} = \frac{t}{\sigma_\varepsilon^2} + \frac{1}{\left(1-\rho_0^2\right)\sigma_{b,0}^2} = \frac{\sigma_{s,0}^2\left[\sigma_\varepsilon^2 + \left(1-\rho_0^2\right)\sigma_{b,0}^2 t\right]}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{42}$$

Similarly, it can be shown that

$$\begin{aligned} \alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2 &= \left(\frac{t}{\sigma_\varepsilon^2} + \phi_1\right)\left(\frac{t}{\sigma_\varepsilon^2} + \phi_2\right) - \left(-\frac{t}{\sigma_\varepsilon^2} + \phi_3\right)^2 \\ &= \frac{1}{\sigma_\varepsilon^2}\left[\left(\phi_1 + \phi_2 + 2\phi_3\right)t + \left(\phi_1\phi_2 - \phi_3^2\right)\sigma_\varepsilon^2\right]. \end{aligned} \tag{43}$$

Using (34) and after some algebraic manipulations, we can get

$$\phi_1 + \phi_2 + 2\phi_3 = \frac{\sigma_{b,0}^2 + \sigma_{s,0}^2 + 2\rho_0\sigma_{s,0}\sigma_{b,0}}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2} = \frac{\lambda_0 + \theta_0}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{44}$$

and

$$\phi_1\phi_2 - \phi_3^2 = \frac{1}{\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{45}$$

Substituting (44) and (45) into (43) gives

$$\alpha_{1,t}\alpha_{2,t} - \alpha_{3,t}^2 = \frac{\sigma_\varepsilon^2 + \left(\lambda_0 + \theta_0\right)t}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2} \tag{46}$$

Note that the expressions in (41), (42) and (46) all share the same denominator. Combining (42) and (46) gives (1). Similarly, (2) can be obtained by combining (41) and (46).

The covariance term $\widehat{\omega}_t$ can be obtained as follows: Based on (30) and (31), we can get

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

From (34) and (39), 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 (3) can be obtained by combining this with (46).

Based on (32) and (33), the means associated with the posterior density function $\pi\left(s, b \mid \mathbf{m}^t\right)$ can be expressed as

$$\widehat{s}_t \equiv E\left[s \mid \mathbf{m}^t\right] = \frac{\alpha_{2,t}\alpha_{4,t} + \alpha_{3,t}\alpha_{5,t}}{\alpha_{1,t}\alpha_{2,t} - \left(\alpha_{3,t}\right)^2}, \tag{47}$$

$$\widehat{b}_t \equiv E\left[b \mid \mathbf{m}^t\right] = \frac{\alpha_{3,t}\alpha_{4,t} + \alpha_{1,t}\alpha_{5,t}}{\alpha_{1,t}\alpha_{2,t} - \left(\alpha_{3,t}\right)^2}. \tag{48}$$

Using (38)-(40), we can get

$$\alpha_{2,t}\alpha_{4,t} + \alpha_{3,t}\alpha_{5,t} = \frac{1}{\sigma_\varepsilon^2}\left\{\left[\left(\phi_2+\phi_3\right)\overline{m}_t + \left(\phi_4-\phi_5\right)\right]t + \left(\phi_2\phi_4+\phi_3\phi_5\right)\sigma_\varepsilon^2\right\}, \tag{49}$$

$$\alpha_{3,t}\alpha_{4,t} + \alpha_{1,t}\alpha_{5,t} = \frac{1}{\sigma_\varepsilon^2}\left\{\left[\left(\phi_1+\phi_3\right)\overline{m}_t - \left(\phi_4-\phi_5\right)\right]t + \left(\phi_3\phi_4+\phi_1\phi_5\right)\sigma_\varepsilon^2\right\}. \tag{50}$$

Using (34)-(36), 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 (49) and (50) gives

$$\alpha_{2,t}\alpha_{4,t}+\alpha_{3,t}\alpha_{5,t}=\frac{(\lambda_0\overline{m}_t+\theta_0 s_0-\lambda_0 b_0)\,t+s_0\sigma_\varepsilon^2}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}, \tag{51}$$

$$\alpha_{3,t}\alpha_{4,t}+\alpha_{1,t}\alpha_{5,t}=\frac{(\theta_0\overline{m}_t-\theta_0 s_0+\lambda_0 b_0)\,t+b_0\sigma_\varepsilon^2}{\sigma_\varepsilon^2\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}. \tag{52}$$

Finally, substituting (46) and (51) into (47) gives

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

which is (4). Equation (5) can be obtained in a similar fashion by substituting (46) and (52) into (48). This completes the proof of Proposition 1. ■

Proof of Proposition 2

Rewrite (1) as follows

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

where

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

Hence, we can write

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

This shows that $\widehat{\sigma}_{s,t}^2$ is strictly decreasing over time. To derive the limit, it is more straightforward to use (1), which gives

$$\lim_{t\to\infty} \widehat{\sigma}_{s,t}^2 = \lim_{t\to\infty} \left\{ \frac{\frac{\sigma_{s,0}^2 \sigma_\varepsilon^2}{t} + \left(1-\rho_0^2\right)\sigma_{b,0}^2}{\frac{\sigma_\varepsilon^2}{t} + \lambda_0 + \theta_0} \right\} = \frac{\left(1-\rho_0^2\right)\sigma_{b,0}^2}{\sigma_{s,0}^2 + \sigma_{b,0}^2 + 2\omega_0}.$$

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

Proof of Proposition 3

Define α_t according to

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

Then combining (4) and (5) gives

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

$$\Rightarrow \widehat{s}_t + \widehat{b}_t - s - b = \alpha_t \left(\overline{m}_t - s - b\right) + (1-\alpha_t) \underbrace{(s_0 + b_0 - s - b)}_{\Delta}, \tag{53}$$

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

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

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_ε^2/t. Using this fact and (53), we can write

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

It is easy to verify that $\alpha_t \to 1$ as $t \to \infty$. Hence, we can conclude that

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

This completes the proof of Proposition 3. ■

Proof of Proposition 4

Rewrite (3) as

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

Taking the limit $t \to \infty$ gives

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

Equation (3) can also be rewritten as follows

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

After some algebraic manipulations, it can be shown that

$$(\lambda_0+\theta_0)\,\omega_0 + \left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2 = \lambda_0\theta_0.$$

Hence, we can write

$$\widehat{\omega}_t = \omega_0 - \frac{\lambda_0\theta_0 t}{\sigma_\varepsilon^2 + (\lambda_0+\theta_0)\,t}, \qquad \text{for all } t.$$

Using this we can get

$$\widehat{\omega}_{t+1} - \widehat{\omega}_t = \frac{-\lambda_0\theta_0\sigma_\varepsilon^2}{\left[\sigma_\varepsilon^2 + (\lambda_0+\theta_0)\,t\right]\left[\sigma_\varepsilon^2 + (\lambda_0+\theta_0)\,(t+1)\right]}.$$

This proves that $\widehat{\omega}_{t+1} \gtreqless \widehat{\omega}_t$ if and only if $\lambda_0\theta_0 \lesseqgtr 0$. Using the identity

$$\lambda_0\theta_0 = \sigma_{s,0}^2\sigma_{b,0}^2 + \left(\sigma_{s,0}^2+\sigma_{b,0}^2\right)\omega_0 + \omega_0^2,$$

we now have $\lambda_0\theta_0 \lesseqgtr 0$ if and only if

$$\sigma_{s,0}^2\sigma_{b,0}^2 + \left(\sigma_{s,0}^2+\sigma_{b,0}^2\right)\omega_0 + \omega_0^2 \lesseqgtr 0$$

$$\Leftrightarrow \left(\sigma_{s,0}^2+\sigma_{b,0}^2\right)\omega_0 + 2\omega_0^2 \lesseqgtr -\left[\sigma_{s,0}^2\sigma_{b,0}^2 - \omega_0^2\right] = -\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2$$

$$\Leftrightarrow \omega_0 \lessgtr \frac{-\left(1-\rho_0^2\right)\sigma_{s,0}^2\sigma_{b,0}^2}{\sigma_{s,0}^2+\sigma_{b,0}^2+2\omega_0} \equiv \widehat{\omega}_\infty.$$

The direction of the last inequality is preserved because $\sigma_{s,0}^2+\sigma_{b,0}^2+2\omega_0 = var\,(s+b)$ in the initial belief. This completes the proof of Proposition 4. ■

Proof of Proposition 5

Combining (4) and (12) 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 $v_t \equiv \overline{m}_t - s - b$ and $\Delta \equiv (s + b - s_0 - b_0)$ which is just a constant. Before any signal is realised, $\overline{m}_t$ is a normal random variable with mean $s+b$ and variance σ_ε^2/t. Hence, $E\left(v_t\right) = 0$ and $var\left(v_t\right) = E\left(v_t^2\right) = \sigma_\varepsilon^2/t$. Using these, we can now write

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

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

$$E\left[\left(\kappa_t v_t + \left(\kappa_t - \kappa_\infty\right)\Delta\right)^2\right] = \kappa_t^2 E\left(v_t^2\right) + \left(\kappa_t - \kappa_\infty\right)^2 \Delta^2.$$

Substituting this and $E\left(v_t^2\right) = \sigma_\varepsilon^2/t$ back into the inequality using

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

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

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

This completes the proof of Proposition 5. ■

Proof of Proposition 6

These results follow immediately from either (4) or (12). For instance, from (4) we can get

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

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

Hence, $\widehat{s}_t^{\dagger} - s_0^{\dagger} \geq 0 \geq \widehat{s}_t^{\ddagger} - s_0^{\ddagger}$ if and only if

$$\kappa_t \left(s_0^{\ddagger} + b_0^{\ddagger} \right) \geq \kappa_t \overline{m}_t \geq \kappa_t \left(s_0^{\dagger} + b_0^{\dagger} \right).$$

Part (ii) can be established by applying the same argument on (12). As for opinion reversal, as explained in the main text, (4) can be rewritten as (24). It follows that $\widehat{s}_t^{\dagger} < \widehat{s}_t^{\ddagger}$ at any time t if and only if

$$s_0^{\dagger} - \kappa_t \left(s_0^{\dagger} + b_0^{\dagger} \right) < s_0^{\ddagger} - \kappa_t \left(s_0^{\ddagger} + b_0^{\ddagger} \right),$$

which can be rearranged to become (23). Part (iv) can be replacing κ_t with κ_∞ in the above expression. ■

Proof of Proposition 7

First consider the case in which $\lambda_0 > 0$ so that $\kappa_t \in (0, 1)$ for all t. Then opinion reversal happens at time t if and only if

$$\left(b_0^{\dagger} - b_0^{\ddagger} \right) > \frac{1 - \kappa_t}{\kappa_t} \left(s_0^{\dagger} - s_0^{\ddagger} \right), \tag{54}$$

where

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

is decreasing over time if $\lambda_0 > 0$. If both λ_0 and θ_0 are strictly positive so that $\kappa_t \in (0, 1)$ for all t, then the above ratio is strictly positive at all times and decreasing towards zero. If $\lambda_0 > 0$ but $\theta_0 < 0$ so that $\kappa_t > 1$ for all t, then the ratio is strictly negative for all t and becomes more negative over time. In both cases,

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

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

Next, consider the case when $\lambda_0 < 0$ so that $\kappa_t < 0$ for all t. Then opinion reversal happens at time t 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 $(1-\kappa_t)/\kappa_t$ 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 $s_0^\dagger = s_0^\ddagger$. This proves the desired result. ■

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

[3] Aytimur, R. E., and R. M. H. Suen (2024) "Information Quality, Disagreement and Political Polarisation," unpublished manuscript.

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

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

[6] 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.

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

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

[9] Diaconis, P., and D. Ylvisaker (1979) "Conjugate Priors for Exponential Families," *Annals of Statistics*, vol. 7 (2), pp.269-281.

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

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

[12] 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.

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

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

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

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

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

[18] Newman, N., R. Fletcher, C. T. Robertson, A. Ross Arguedas, and R. K. Nielsen (2024) *Reuters Institute digital news report 2024*. The Reuters Institute for the Study of Journalism, Oxford, UK.

[19] 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. Elvesier.