FTG Working Paper Series

Trust in Signals and the Origins of Disagreement*

by

Ing-Haw Cheng
Alice Hsiaw

Working Paper No. 00047-00

Finance Theory Group

www.financetheory.com

Trust in Signals and the Origins of Disagreement*

Ing-Haw Cheng† Alice Hsiaw‡

December 2018

Abstract

Why do individuals interpret the same information differently? We propose that individuals follow Bayes' Rule when forming posteriors with one exception: when assessing the credibility of signal sources, they "double-dip" the data and use already-updated beliefs instead of their priors. Individuals who make this mistake either over- or underreact to new information depending on the order in which they received previous signals. Traders engage in excessive speculation associated with price bubbles and crashes. Our model provides a theory of the origins of disagreement: individuals disagree about both unknown states and credibility despite sharing common priors and information.

Keywords: disagreement, learning, expectations, speculation, bubbles

*The authors thank Roland Bénabou, Dan Feiler, Teresa Fort, Jens Großer, Adam Kleinbaum, Botond Kőszegi, Jon Lewellen, Juhani Linnainmaa, Ted O'Donoghue, Martin Oehmke, Carol Osler, Davide Pettenuzzo, Uday Rajan, Tanya Rosenblat, Kathy Spier, Phillip Stocken, Courtney Stoddard, Dustin Tingley, Wei Xiong, Muhamet Yildiz, seminar participants at Brandeis University, Cornell University, Harvard University, University of Chicago, MIT Sloan, and workshop participants at CSWEP CeMENT, the Duke Behavioral Models of Politics conference, the 2017 North American Summer Meeting of the Econometric Society, the Stanford Institute for Theoretical Economics, and the Behavioral Economics Annual Meeting for comments. This version subsumes a previous version titled "Distrust in Experts and the Origins of Disagreement."

†Dartmouth College, Tuck School of Business. Email: ing-haw.cheng@tuck.dartmouth.edu.

‡Brandeis University, International Business School. Email: ahsiaw@brandeis.edu.

Ing-Haw Cheng has no relevant or material financial interests that relate to the research described in this paper. This research was supported by internal research funds from the Tuck School of Business, Dartmouth College.

Alice Hsiaw has no relevant or material financial interests that relate to the research described in this paper. This research was supported by internal research funds from the International Business School, Brandeis University.

"We all have the same information, and we're just making different conclusions about what the future will hold." - Henry Blodget, *Former analyst (Lewis, 2002)*

Why do investors interpret the same information differently? From a practical standpoint, this question is more relevant than ever: Despite widely shared information, investors today disagree over the value of everything from Bitcoin to technology stocks. Several theories show that disagreement about the informativeness of signals helps explain disagreement about an unknown state of the world, leading to speculative trading, price volatility, and bubbles (Harrison and Kreps, 1978; Harris and Raviv, 1993; Kandel and Pearson, 1995; Scheinkman and Xiong, 2003; Kyle et al., 2017). However, these theories often do not model why agents disagree over interpretations, leaving several fundamental questions open. For example, why might some agents trust some signals more, while other agents trust the same signals less?

This paper develops a model of why identical economic agents with common priors endogenously interpret the same information differently. We start with the premise that agents are not only uncertain of the state of the world (e.g., of an investment's future payoffs), but also of the data-generating process that produces signals and hence the informativeness of signals for the state. For example, investors may be uncertain whether a given valuation metric, or an analyst's recommendations, has any value in forecasting a firm's future performance. To set terminology, we refer to signals as coming from *sources* (financial metrics, analysts, advisors) who have uncertain *credibility* (informativeness for the true state).

With uncertain credibility, an agent faces a key inference problem in determining how much weight to apply to a source's signals. Bayesian agents use only their priors to determine this weight. They treat uncertainty about source credibility no differently from how they treat uncertainty about the state. Thus, endowing Bayesians with exogenous heterogeneous beliefs or priors about credibility generates disagreement about the state. With common priors and common information, however, Bayesian agents would agree.

Our theory proposes that, when faced with this key inference problem, agents make a mistake we call *pre-screening*. Pre-screeners mistakenly use updated beliefs about the credibility of signal sources when weighting signals instead of priors. Intuitively, if new signals suggest low credibility, pre-screeners think: "I now think the source is not credible, and my beliefs should reflect that I saw non-credible signals. After all, if the source is probably not credible, then all of its signals are more questionable than I originally thought." They update in two steps. In the first step, they form a first-stage belief about credibility based on newly-observed signals. In the second step, they use this first-stage belief instead of their prior to weight all signals they have ever seen from the source. Pre-screening's departure from Bayes' Rule is this substitution of updated beliefs about source credibility for priors.

This two-step procedure leads pre-screeners to "double-dip" the data. Consider the following example. An individual who is reasonably sure that he weighs 200 pounds steps on a scale that he also believes is likely reliable, and the scale reads 300 pounds. Surprised by the reading, a Bayesian's posterior is that the scale is likely unreliable but that there is some chance he weighs 300 pounds, as he carefully combines the likelihood of the data with his prior belief that the scale was reliable. A pre-screener, in contrast, first infers that the scale is unreliable upon seeing the reading, and then combines the likelihood of the data with this updated belief as though he knew the scale was unreliable all along. This leads the pre-screener to discount the possibility that he might be 300 pounds too much, on the premise that the scale is unreliable.

Pre-screening draws on two insights. First, because agents must weight the data when updating beliefs, they may naturally (but erroneously) think credibility is a fundamental parameter that they: 1) seek to learn first, and 2) then apply to all the data due to its essential importance. Mistakenly double-dipping the data to infer essential parameters analogously occurs in criticized forms of Empirical Bayes inference methods (Lindley, 1969; Carlin and Louis, 2000). Second, using updated beliefs to judge past data is consistent with evidence from cognitive psychology that individuals who have seen information often fail to ignore it when the context requires using only prior beliefs. Prominent examples include hindsight bias (e.g., Fischhoff, 1975; Hawkins and Hastie, 1990), where individuals who have seen data are prone to think they have "known it all along,'' and the curse of knowledge (Camerer et al., 1989), the difficulty of conceptualizing what it was like to be uninformed in the past. We link these two insights and suggest that agents may be prone to use updated beliefs about credibility instead of their priors, as if they had updated beliefs about credibility all along.

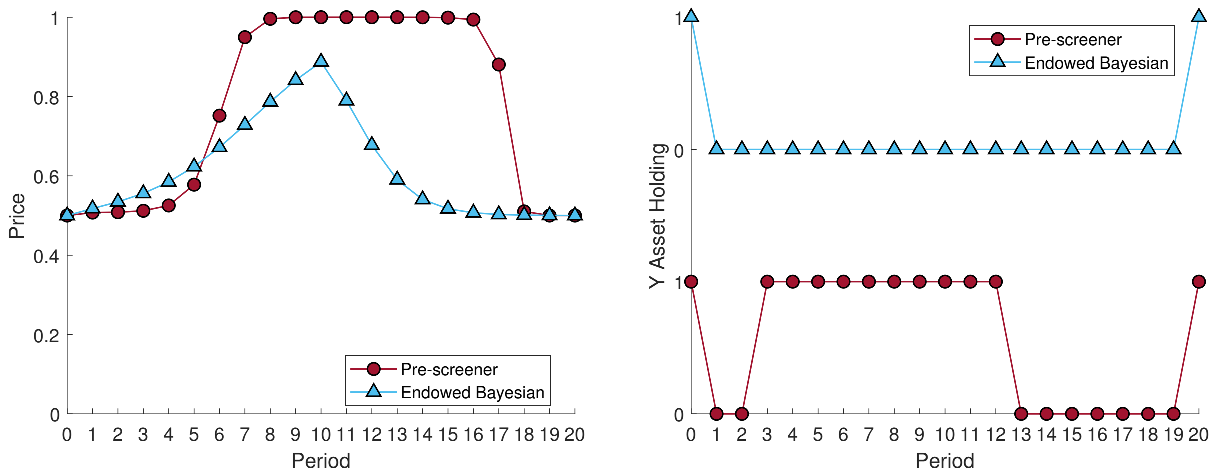
We introduce pre-screening in Section 1. To isolate the effects of biased learning, we assume that signal sources are data-generating processes that produce signals about an uncertain state and abstract from strategic motives (e.g., Morgan and Stocken, 2003; Hong, Scheinkman and Xiong, 2008). We model a source's credibility as its underlying ability to discern the true state of the world, which is either A or B.

We characterize three predictions of pre-screening in Section 2. First, pre-screeners who share common priors and who have seen the same objective information disagree if they received the signals in different order, even though Bayesians would agree. This is because the information in early signals about source credibility, or "first impressions of credibility," have an outsized influence on subsequent beliefs. Second, pre-screeners with differing first impressions of credibility have correlated disagreements about the unknown state of the world and source credibility. A pre-screener who thinks the state is more likely than objectively warranted also thinks the source is more credible than objectively warranted.

Third, pre-screeners endogenously over-react or underreact to new information, behavior akin to overconfidence (Scheinkman and Xiong, 2003), confirmation bias (Rabin and Schrag, 1999), and their opposites. However, in contrast to these theories, agents can endogenously over- and underreact to signals that either confirm or contradict beliefs about the state, depending on how these signals interact with first impressions about credibility.

More broadly, pre-screening provides a theory of the origins of disagreement, since pre-screeners disagree despite sharing common priors and information. In Section 3, we show that endogenizing these origins helps explain speculative trading and bubbles. We ask: How do pre-screeners with differing first impressions about credibility trade, compared to Bayesians with heterogeneous priors? We build on the trading game of Harris and Raviv (1993) where agents observe common signals about an asset's unknown payoff and "agree to disagree" about source credibility, but we allow agents to learn about credibility rather than know it for certain. We compare a game where all traders are pre-screeners with a game where all traders are Bayesians. At the opening of trade, disagreement about credibility is the same across the two games. However, in the game with pre-screeners, this initial disagreement originates from differing first impressions of credibility from signals prior to trade, whereas we endow Bayesians with heterogeneous beliefs.

We first show that pre-screeners trade more than Bayesians trade. Pre-screeners engage



(a) Prices **(b)** Asset holdings

Figure 1: Trading game. This figure plots outcomes from a trading game between two groups of traders, X and Y, described in Section 3, when they are either both pre-screeners or both endowed Bayesians. Prices are in Panel (a) and the asset holdings of group Y are in Panel (b). Realized signals in periods 1-10 are good cash flow news ('a'), while periods 11-20 have bad cash flow news ('b'). Trader beliefs at period 0 are equal to the beliefs that pre-screeners would have after observing '$aabb$' for X and '$abab$' for Y, starting from common priors. Parameters are $\{q_H, q_L, \omega^A_{-\tau}, \omega^H_{-\tau}, T, \tau\} = \{0.8, 0.5, 0.5, 0.3, 20, 4\}$.

in "speculative trading" as defined by Harrison and Kreps (1978): "an investor may buy the stock now so as to sell it later for more than he thinks it is actually worth, thereby reaping capital gains." In our model, this occurs if some traders believe that other pre-screeners will over-react to positive news about fundamental value, even if they are relatively skeptical about fundamental value themselves. This type of speculative trade is not a necessary feature of disagreement models and is notably absent in the model of Harris and Raviv (1993) upon which we build, and is difficult to generate in rational expectations models (Tirole, 1982).

We then show that pre-screening generates bubbles, crashes, and speculation in ways that correspond to several empirical features described by Barberis (2018). Figure 1 illustrates one example: In response to a string of good cash flow news, agents develop too much trust in the source and become over-optimistic about cash flows, leading asset prices to rise sharply above the Bayesian benchmark and become a bubble. If bad cash flow news then follows, prices are initially resistant and do not appreciably fall, because traders view the previous good news as very credible. However, after enough bad cash flow news arrives, agents begin to doubt whether they believe anything the source reported before, due to the contradiction with previously-reported good cash flow news. Beliefs about credibility collapse, and prices fall quickly. There is excess volume both during the bubble and crash. During the price rise, traders who are relatively skeptical about fundamental value speculatively buy the asset to "ride the bubble" (Abreu and Brunnermeier, 2003; Brunnermeier and Nagel, 2005). We show more generally that endogenous over- and under-reaction to cash flow news and differing first impressions of credibility are key mechanisms that explain these behaviors.

Our theory has broad implications. In Section 4, we introduce two further applications of pre-screening and ask: 1) When can new information resolve disagreement? and 2) How do sources systematically "slanted" towards one state affect disagreement? We also discuss how our framework differs from other biases such as gradual information flow (Hong and Stein, 1999), inattention or selective attention (Hirshleifer and Teoh, 2003; Peng and Xiong, 2006; Schwartzstein, 2014; Suen, 2004), bounded rationality, and others.

Our key contribution is to offer a new micro-foundation for the origins of why individuals disagree about the interpretation of the same data. Research ranging from Kandel and Pearson (1995) and Cookson and Niessner (2016) document that differences in signal interpretation are an important component of overall disagreement about stock prices. Our approach complements the heterogeneous priors approach (Acemoglu et al., 2016; Morris, 1995) by providing a theory for the origins of such priors. Our model speaks to fundamental questions about why disagreement about states and source credibility are correlated, how disagreement originates and is resolved, and the origins of speculation and bubbles.

Our theory also connects to the broader literature on polarized disagreement in economics (Gentzkow and Shapiro, 2006; Mullainathan and Shleifer, 2005). A core feature of many disagreements is that individuals disagree not just about their positions ("Do humans affect climate change?"), but also about the credibility of information sources that inform those positions ("How reliable are scientists and their data?"). In debates over questions of economics ("What is the value of stimulus spending?"), medicine ("Are vaccinations safe for children?"), and politics ("Why is it hard to debunk fake news?"), one side typically expresses supreme confidence in their preferred experts while dismissing the other side's trusted sources. We speculate on how our theory helps explain this disagreement in the conclusion.

1 Model

1.1 Information environment

An agent learns about an unknown state $\theta \in \{A, B\}$ by observing binary signals $s_t \in \{a, b\}$ in each period t from a signal source. Sources are data-generating processes (e.g., financial metrics, analysts, advisors) and are not strategic. A source's credibility is a type c that describes the informativeness of its signals. Nature draws true source credibility independently from the true state. Conditional on state and credibility, signals are independently and identically distributed. We first assume the agent observes one signal per period from a single source, and later generalize to multiple sources and multiple signals per period.

We focus on the case where credibility summarizes the reliability, or underlying skill, of the source in determining the state. Specifically, a source has credibility $c \in \{L(ow), H(igh)\}$. A high-quality source has a higher probability of correctly reporting the state than a low-quality source and is more informative: $P(s_t = a|c, A) = P(s_t = b|c, B) = q_c$. We require $q_c \in [1/2, 1)$: the least reliable signal is noise, while even the most reliable signal is not perfectly correlated with the true state. We assume $q_L < q_H$. In Section 4.2, we generalize credibility to include the possibility that the source is biased towards one state.

The agent does not know credibility c, but does know that it is either high or low. We take this as given, but highlight three (among possibly more) informational frictions that make this assumption realistic. First, agents often have insufficient data to establish a source's credibility about a given state for certain. Second, agents often rely on opinions from experts with uncertain credibility, and third-party information often cannot resolve this uncertainty. Finally, the ability to learn about credibility by comparing predictions to outcomes through repeated controlled experiments is often limited.

Consider, for example, disagreement over valuations during the dot-com boom. High price/dividend ratios led Campbell and Shiller (1998, 2001) to warn of future price declines. However, other observers doubted whether these ratios were a reliable guide to future returns. As Campbell and Shiller discuss, this uncertainty arose because agents had insufficient history to truly know the P/D ratio's reliability in forecasting returns, especially in the context of the technology boom.

In practice, investors often rely not on data directly, but on recommendations and opinions from financial advisors, fund managers, and analysts, who have uncertain credibility themselves. An extensive literature examines both the underlying skill among these professionals (Clement, 1999; Fama and French, 2010; Kacperczyk et al., 2014) as well as the possibility that they are biased (Lin and McNichols, 1998; Hong and Kubik, 2003; Malmendier and Shanthikumar, 2014; McNichols and O'Brien, 1997). Agents face significant inference problems in these settings, which may make mistakes more likely (Malmendier and Shanthikumar, 2007; Hong, Scheinkman and Xiong, 2008).

Why don't independent signals about credibility, such as an expert's "credentials" (CFA charters, PhDs, and so on), resolve this gap? General credentials such as degrees and certifications may not be very informative about the credibility of a source's opinions for a specific state. Consistent with this, DellaVigna and Pope (forthcoming) run a large experiment where they ask economists to forecast the effectiveness of different incentive treatments on subjects, and find that objective measures of expertise are unrelated to forecast accuracy.

Furthermore, credibility may be uncertain to non-experts even if it is more certain to other experts. For example, a financial expert may be able to ascertain whether another expert is high or low quality based on their résumé, but this is of no help to non-experts because that expert's quality is also unknown. Uncertain credibility may thus remain even if one expands the set of signal sources an agent observes, because each source carries its own uncertainty about credibility.

Finally, perhaps the most useful tool for evaluating credibility—the ability to compare predictions to outcomes through repeated controlled experiments—is limited in many settings. An expert or source may predict that the dot-com bubble will burst, but they may have lucked out, or were vague in their timing. Given that outcomes must be compared to predictions outside of a laboratory setting, it seems reasonable to assume that credibility has residual uncertainty and that agents are learning about it.

In summary, we view the assumption of uncertain credibility as more plausible than that of known credibility, particularly given information frictions. Because credentials are of uncertain informativeness themselves, we focus on a stark environment with no signals that explicitly convey source credibility.

1.2 Learning

Suppose the agent has the prior that the state and credibility are independent with marginal probabilities ω_0^θ for the state and ω_0^c for each credibility type c. Let the agent observe a sequence of n signals, denoted $\mathbf{s}_n = (s_1, s_2, \ldots s_n)$, where one signal is observed each period.

A Bayesian's posterior belief $P^u(c, \theta|\mathbf{s}_n)$ equals:

$$P^u(c, \theta|\mathbf{s}_n) = \frac{\left(\prod_{t=1}^n P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c}{\sum_c \sum_\theta \left(\prod_{t=1}^n P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c}. \tag{1}$$

The Bayesian uses her prior belief ω_0^c about source credibility to weight the likelihood of signals $\left(\prod_{t=1}^n P(s_t|c, \theta)\right)$, and infers her posterior belief $P^u(c, \theta|\mathbf{s}_n)$ in one step based on the *information content* of signals, defined as follows:

Definition 1 (Information content) *The* ***information content*** *of any sequence of signals $\mathbf{s}_n$ is given by the number of "a" signals n_a and the number of "b" signals n_b.*

We propose that individuals make a mistake that we call *pre-screening* when faced with the problem of determining how much weight to apply to a source's signals. A pre-screener mistakenly uses updated beliefs about credibility when weighting signals instead of priors. She updates in two steps. First, she forms an updated first-stage belief about credibility, denoted $\kappa_c(\mathbf{s}_n)$, using Bayes' Rule.[1] Second, she uses this updated belief $\kappa_c(\mathbf{s}_n)$ to weight the signals $\mathbf{s}_n$ in forming her joint posterior of state and credibility, denoted $P^b(c, \theta|\mathbf{s}_n)$. The key mistake is in using $\kappa_c(\mathbf{s}_n)$ to evaluate all signals, whereas a Bayesian uses her prior ω_0^c.

To illustrate, suppose a pre-screener observes two signals, one in each period. After

[1] More generally, all signals observed in a given time period are used to form the first-stage belief about credibility. In Cheng and Hsiaw (2017) we extend the model to allow multiple signals per period and further explore the implications of signal timing. Here, we confine the signals to one per period for simplicity.

observing the first signal (s_1), the first-stage updated belief about credibility, $\kappa_c(s_1)$, is:

$$\kappa_c(s_1) = \frac{\omega_0^c \sum_\theta P(s_1|c,\theta)\omega_0^\theta}{\sum_c \sum_\theta P(s_1|c,\theta)\omega_0^\theta \omega_0^c}.$$

Using $\kappa_c(s_1)$ to form the joint posterior belief on the state and credibility, $P^b(c,\theta|s_1)$, yields the pre-screener's posterior beliefs after the first signal:

$$P^b(c,\theta|s_1) = \frac{P(s_1|c,\theta)\kappa_c(s_1)\omega_0^\theta}{\sum_c \sum_\theta P(s_1|c,\theta)\kappa_c(s_1)\omega_0^\theta}.$$

After observing the second signal (s_2), the pre-screener's first-stage updated belief about credibility, $\kappa_c(s_1, s_2)$, is:

$$\kappa_c(s_1, s_2) = \frac{\sum_\theta P(s_2|c,\theta)P^b(c,\theta|s_1)}{\sum_c \sum_\theta P(s_2|c,\theta)P^b(c,\theta|s_1)}.$$

The pre-screener then uses $\kappa_c(s_1, s_2)$ to form her joint posterior belief on the state and credibility by re-weighting all the information from the source. The posterior, $P^b(c,\theta|s_1,s_2)$, equals:

$$P^b(c,\theta|s_1,s_2) = \frac{P(s_2|c,\theta)P(s_1|c,\theta)\kappa_c(s_1,s_2)w_0^\theta}{\sum_c \sum_\theta P(s_2|c,\theta)P(s_1|c,\theta)\kappa_c(s_1,s_2)\omega_0^\theta}.$$

Iterating on the pre-screener's process of repeatedly substituting newly-updated beliefs about credibility for priors allows us to characterize her posterior beliefs.

Definition 2 (Pre-screener's beliefs) *After observing a sequence of n signals $\mathbf{s}_n$ from a source, the **pre-screener's first-stage updated belief** about source credibility, $\kappa_c(\mathbf{s}_n)$, is given by:*

$$\kappa_c(\mathbf{s}_n) = \frac{\kappa_c(\mathbf{s}^{n-1}) \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\omega_0^\theta\right)}{\sum_c \kappa_c(\mathbf{s}^{n-1}) \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\omega_0^\theta\right)}, \quad (2)$$

*where $\kappa_c(\emptyset) = \omega_0^c$. The **pre-screener's joint posterior** on credibility and the state,*

$P^b(c, \theta|\mathbf{s}_n)$, *is given by:*

$$P^b(c, \theta|\mathbf{s}_n) = \frac{\left(\prod_{t=1}^n P(s_t|c, \theta)\right) \kappa_c(\mathbf{s}_n)\omega_0^\theta}{\sum_c \sum_\theta \left(\prod_{t=1}^n P(s_t|c, \theta)\right) \kappa_c(\mathbf{s}_n)\omega_0^\theta}. \tag{3}$$

Comparing Equations 1 and 3, the pre-screener's key mistake is to use the first-stage updated belief $\kappa_c(\mathbf{s}_n)$ to evaluate the signals in Equation 3, whereas a Bayesian uses her prior ω_0^c in Equation 1. Note that if there is no uncertainty about source credibility, the first step becomes innocuous, so the pre-screener's posterior beliefs are identical to the Bayesian's.

The above definition of a pre-screener's beliefs assumes ex-ante independence of states and credibility. We maintain this assumption both for simplicity and because it isolates the effect of pre-screening on joint beliefs about the state and credibility without assuming any correlation ex-ante. We provide a generalized definition in the Internet Appendix.

Pre-screening draws on two insights. First, because agents must weight the data when forming updated beliefs, agents may naturally (but erroneously) treat credibility as a fundamental parameter that they seek to learn first and then mistakenly apply to all the data due to its essential importance. "Double-dipping" the data this way leads to problematic inference. For example, in criticized forms of Empirical Bayes research methods (Carlin and Louis, 2000), researchers sometimes first calibrate model hyperparameters governing their priors using the data before performing a full analysis on the same data. Lindley (1969) famously noted that "there is no one less Bayesian than an empirical Bayesian."

In our context, a pre-screener uses signals to calibrate the weight in a first step, $\kappa_c(\mathbf{s}_n)$. In the second step, they apply the weight to all signals they have seen from the source by substituting $\kappa_c(\mathbf{s}_n)$ for their prior, ω_0^c. Intuitively, if new signals suggest low credibility and $\kappa_c(\mathbf{s}_n)$ is low, pre-screeners think: "I now think the source is not credible, and my beliefs should reflect that I saw non-credible signals. After all, if the source is not credible, then all of its signals are more questionable than I originally thought." This reasoning mistakenly double-dips the data and uses updated beliefs to judge current and past signals.[2]

[2] As another example of why using updated beliefs might seem plausible, Subramanyam (1996) notes that, when the error precision of a normally-distributed signal about a normally-distributed unobserved random variable is unknown, a Bayesian can calculate the posterior mean by applying the updated signal-gain in the linear updating equation for the mean, due to the Law of Iterated Expectations. This is due to the

Second, the idea of erroneously using updated beliefs to judge past data is consistent with longstanding findings in cognitive psychology that individuals who have seen information often fail to ignore it when the context requires using only prior beliefs. Intuitively, once an individual has seen the data, she thinks she knew it all along. One prominently related phenomenon is hindsight bias, or the tendency for individuals with "outcome knowledge to overestimate what they would have known without outcome knowledge" (Fischhoff, 1975; Fischhoff and Beyth, 1975; Hawkins and Hastie, 1990). A core tenet of leading cognitive models for hindsight bias (Hoffrage et al., 2000; Hertwig et al., 2003) is that "If knowledge is constantly updated"—as it is in our setting—then "inferences based on updated knowledge may be different from those based on past knowledge" (Hoffrage and Hertwig, 1999, p.196).

For example, according to the RAFT cognitive process model (Hoffrage et al., 2000), hindsight bias is generated "if individuals are unable to directly retrieve their initial judgment but try to reconstruct it by repeating the original judgment process, this time, however, on the basis of the updated knowledge base" (Blank and Nestler, 2007). An implication of an individual's belief that she "knew it all along" in the past is that she also currently thinks that she "knew it all along" and behaves accordingly. For example, Biais and Weber (2009) show that individuals exhibit more hindsight bias when not explicitly reminded of their prior beliefs, and that hindsight bias is correlated with lower performance among bankers.

The curse of knowledge is the failure of well-informed individuals to accurately anticipate the judgments of less-informed agents (Camerer et al., 1989), as they are unable to ignore their own additional information and imagine what it was like to be uninformed in the past. Relatedly, Madarász (2012) suggests that people overestimate the extent to which others know what they know.

In our context, individuals face the key inference problem of assessing how much weight to apply to a source's signals. A pre-screener infers source credibility first by forming first-stage belief $\kappa_c(\mathbf{s}_n)$ and then erroneously thinks she "knew it all along": Having observed the

linear relationship between the posterior mean and the realized signal in a Gaussian environment, and the procedure does not recover the joint posterior of the mean and error precision. The non-monotone reaction to surprises in that paper occurs due to how the likelihood combines with a Bayesian's prior beliefs when signal precision is uncertain.

data, she believes she had a prior belief of $\kappa_c(\mathbf{s}_n)$ about credibility all along, and forms her posterior beliefs using Bayes' Rule over past and current signals. This creates behavior akin to hindsight bias and the curse of knowledge, as a pre-screener fails to appropriately use her prior about credibility.

The information processing mechanism of erroneously using updated beliefs to form posterior beliefs was also conjectured by Lord, Ross and Lepper (1979, p.2106–2107). In an experimental setting examining how subjects update beliefs in response to information about capital punishment, they write that "Our subjects' main inferential shortcoming, in other words, did not lie in their inclination to process evidence in a biased manner...Rather, their sin lay in their readiness to use evidence already processed in a biased manner to bolster the very theory or belief that initially 'justified' the processing bias."

As Rabin and Schrag (1999, p.46–47) discuss, the "sin" is analogous to a teacher who first assigns a student a low grade because she unfavorably interprets an unclear answer from the student as consistent with priors about low ability, but then goes on to erroneously use the low grade as *further* or *additional* evidence of low ability. In our context, a pre-screener errs in using the data to formulate a first stage belief about credibility before using that belief as further evidence about the informativeness of signals when forming her posterior.

Pre-screening does assume agents have memory over previous signals, as pre-screeners substitute first-stage updated beliefs $\kappa_c(\mathbf{s}_n)$ for time-0 priors ω_0^c. One can incorporate limited memory by assuming that agents substitute $\kappa_c(\mathbf{s}_n)$ for priors K-periods in the past, where $K > 0$ represents some fixed window or time period (pre-screening corresponds to $K = n$). Under this alternative, agents still believe credibility is a fundamental parameter they need to know and double-dip the data. They mistakenly update on credibility in a first stage and then act as if they had updated beliefs "all along" (for the past K signals). This leads to qualitatively similar effects we describe later, such as belief path-dependence and endogenous over- and under-reaction to signals.[3]

[3] One difference is with the extreme case of $K = 0$. This case preserves the first-stage update but makes beliefs path-independent (though a form of over- and under-reaction continue to occur). However, we view this as less consistent with our motivation for why agents would perform the first-stage update in the first place: they think credibility is a fundamental parameter that they seek to learn first and then mistakenly apply to all the data due to its essential importance.

1.3 An example

An individual who is reasonably sure that he weighs 200 pounds steps on a scale with unknown credibility, and the scale reads 300 pounds. A second reading also shows 300 pounds. What would a Bayesian and pre-screener infer after each signal?

In the following example, the Bayesian's belief about weight moves progressively towards 300 pounds after each signal, even though he also believes the scale might not be reliable. In contrast, a pre-screener's belief about his weight moves very little towards 300 pounds, and will move *back towards* 200 pounds after the second signal, as he concludes the scale is almost certainly unreliable. The stark difference in this illustration occurs because the pre-screener erroneously acts as if he had updated beliefs about credibility all along.

Let the individual's weight $\theta \in \{200, 300\}$ pounds be the unknown state of the world, and suppose the scale can read either 200 or 300. Both the Bayesian and pre-screener are uncertain about the scale's reliability (probability of reporting the truth), which can be high ($q_H = 0.9$, reliable) or low ($q_L = 0.5$, noise). They share the same priors that their weight is probably 200 pounds ($\omega_0^{200} = 0.98$) and that the scale is probably reliable ($\omega_0^H = 0.80$).

Given these beliefs, the first reading of $\{s_1\} = \{300\}$ is quite a surprise. The Bayesian's marginal posterior beliefs equal $P^u(\theta = 200|\{300\}) = 0.91$ and $P^u(c = H|\{300\}) = 0.48$. Even though the Bayesian's posterior belief is that the scale is likely unreliable, he is careful to reach his joint posterior beliefs by combining the likelihood of the data with his *prior* belief that the scale is reliable, $\omega_0^H = 0.80$, following Equation 1.

A pre-screener's marginal posteriors equal $P^b(200|\{300\}) = 0.96$ and $P^b(H|\{300\}) = 0.18$. He reaches his joint posterior beliefs by erroneously combining the likelihood of the data with an *updated* belief that the scale is likely *un*reliable, $\kappa_H(\{300\}) = 0.48$ (Equations 2 and 3). This "double-dipping" leads him to update insufficiently towards the belief that he weighs 200 pounds and too much in the direction that the scale is unreliable.

After a second reading, the Bayesian's marginal posterior beliefs equal $P^u(200|\{300, 300\}) = 0.80$ and $P^u(H|\{300, 300\}) = 0.29$. Notice that the Bayesian's belief about his weight progressively moves away from 200 pounds after each signal (from 0.98 to 0.91 to 0.80), even though his trust in the scale progressively drops.

In contrast, the pre-screener's beliefs equal $P^b(200|\{300, 300\}) = 0.97$ and $P^b(H|\{300, 300\}) = 0.01$. His posterior probability that he weighs 200 pounds increases (from 0.96 to 0.97), which is the wrong way relative to the Bayesian. This is because he erroneously thinks that his updated belief that the scale is unreliable, $\kappa_H(\{300, 300\}) = 0.09$, should apply to all of the scale's readings, leaving him fairly confident that he is 200 pounds, close to what he initially believed. After all, if the scale is probably unreliable, then he thinks that all of its readings are more questionable than he originally thought. Put yet another way, the pre-screener thinks: I now think the scale is not credible, and my beliefs should reflect that any concern I had about being overweight stemmed from not-credible readings. Therefore, he is now (erroneously) less concerned about being overweight.

With more readings of 300 pounds, it takes the Bayesian only three signals to begin inferring that the scale is likely reliable. His belief about weight progressively moves away from 200 pounds with each signal. It takes the pre-screener six readings to begin believing that the scale might be reliable, and during this time, his belief about weight is moving towards 200 pounds—the wrong way—before reversing afterwards.

In Section 4.2, we generalize credibility to include the possibility that a signal is "slanted" or "tilted" towards one state. Changing the above example so that the Bayesian and pre-screener are uncertain about whether the scale is slanted towards reporting 300 pounds yields qualitatively similar insights under several scenarios. Both examples illustrate that the pre-screener's error is to use the signals to infer source credibility first, and then use this updated credibility belief to weight current and past signals when forming her posterior.

2 Pre-screening generates disagreement

In this section, we develop three key predictions about the effect of pre-screening on disagreement. We maintain the common priors assumption throughout.

To fix ideas, consider the case where the prior on the state is neutral ($\omega_0^\theta = 1/2$) and the agent observes only one signal ($n = 1$). No disagreement occurs, either between a Bayesian and a pre-screener ($P^b(c,\theta|\mathbf{s}_n) = P^u(c,\theta|\mathbf{s}_n)$), or between two pre-screeners, because $\kappa_c(\mathbf{s}_n) = w_0^c$. This simple example shows that disagreement is not exogenously built into pre-screening.

With more signals, disagreement can occur. Consider disagreement between a Bayesian and a pre-screener. The average disagreement about the state is zero as long as both agents begin with the prior that both states are equally likely. This is because beliefs about states are ex-ante symmetric around A and B and are ex-ante independent of reliability. However, there is ex-post disagreement along realized sequences: the average squared difference in marginal posteriors about θ is strictly positive. Similarly, average disagreement between two pre-screeners about the state is zero even though pre-screeners disagree along realized sequences. Furthermore, pre-screeners disagree even if they see signal sequences that share the same information content, so long as they see signals in different order.

Proposition 1 (Ex-ante disagreement about θ) *Let all agents share a common prior of $(\omega_0^A, \omega_0^H) = (1/2, \hat{\omega})$ for any $\hat{\omega} \in (0,1)$, and suppose this represents the true distribution from which nature draws $(\theta, \hat{\omega})$.*

1. *Average disagreement is zero*

   $E_0[P^b(\theta = A|\mathbf{s}_n) - P^u(\theta = A|\mathbf{s}_n)] = 0$*, where the expectation E_0 is taken over this distribution and all signal sequences $\mathbf{s}_n$.*

2. *Average squared disagreement is positive*

   (a) $E_0\left[\left(P^b(\theta = A|\mathbf{s}_n) - P^u(\theta = A|\mathbf{s}_n)\right)^2\right] > 0$*, where the expectation E_0 is taken over this distribution and all signal sequences $\mathbf{s}_n$.*

   (b) $E_0\left[\left(P^b(\theta = A|\mathbf{s}_n^J) - P^b(\theta = A|\mathbf{s}_n^M)\right)^2\right] > 0$*, where the expectation E_0 is taken over the distribution of all pairs of signal sequences $\{\mathbf{s}_n^J, \mathbf{s}_n^M\}$ that have identical information content but where signals occur in different order.*

2.1 First impressions about credibility generate disagreement

The key reason disagreement occurs is that a pre-screener's final posteriors depend on signal order. "First impressions about credibility" matter: early signals color the interpretation of later signals through their effect on beliefs about credibility.

For example, suppose $\omega_0^\theta = 1/2$ and consider the signal sequences $\{a, a, b\}$ and $\{b, a, a\}$, which have identical information content. A Bayesian's posterior beliefs are the same irrespective of order: $P^u(c, \theta|\{a, a, b\}) = P^u(c, \theta|\{b, a, a\})$. But for a pre-screener, $P^b(c, \theta|\{a, a, b\}) \neq P^b(c, \theta|\{b, a, a\})$, since $\kappa_c(\{a, a, b\}) \neq \kappa_c(\{b, a, a\})$. The reason the κ_c values differ is as follows. After each new signal, the pre-screener substitutes first-stage updated beliefs about credibility κ_c for prior ω_0^c. This repeated process means that each new first-stage updated belief includes previous substitutions for ω_0^c. Thus, early signals accumulate in κ_c and disproportionately influence beliefs. Since signal order affects κ_c, it also affects posteriors and generates disagreement among pre-screeners who see the same set of signals in different order.

To characterize first impressions about credibility, we first define the following terms:[4]

Definition 3 (Optimism and trust) *Fix the information content with $n_a > n_b$ without loss of generality. Given a signal sequence $\mathbf{s}_n$,*

1. *A pre-screener is **optimistic** if $Pr^b(\theta = A|\mathbf{s}_n) > Pr^u(\theta = A|\mathbf{s}_n)$, and **pessimistic** if strictly less than ($<$).*

2. *A pre-screener **overtrusts** if $Pr^b(c = H|\mathbf{s}_n) > Pr^u(c = H|\mathbf{s}_n)$, and **under-trusts** if strictly less than ($<$).*

We show that there is a unique sequence of signals that generates the maximal over- and under-trust for any fixed information content:

Proposition 2 (First impressions about credibility) *Let $(\omega_0^A, \omega_0^H) = (1/2, \hat{\omega})$ for any $\hat{\omega} \in (0, 1)$. Consider a given combination of n_a a signals and n_b b signals, where $n_a > n_b \geq 1$. The sequence in which n_a consecutive a signals is followed by n_b consecutive b signals generates the maximal degree of trust in the source. The sequence in which n_b pairs of (a, b) signals is followed by $n_a - n_b$ "a" signals generates the minimal degree of trust in the source.*

[4]Because there is no sense in which A is a better outcome than B until Section 3, a more precise definition would replace optimism with "overestimates the likelihood of A" and pessimism with "underestimates the likelihood of A." We choose "optimistic" and "pessimistic" purely for brevity.

Proposition 2 shows that, holding information content fixed, pre-screeners erroneously believe that the timing of signal reversals is itself informative, in that a pattern of few (more) initial reversals inflates (deflates) their beliefs about source reliability. Holding information content fixed, re-ordering the signals so that the longest consistent streak appears first generates the most trust in the source, while alternating the signals first generates the least trust. In contrast, a Bayesian's final beliefs are independent of signal order.

There is an asymmetry between mixed versus identical signals: mixed signals are worse news for credibility than identical signals are good news. For example, in the extreme case of $q_L = 1/2$ and $q_H \approx 1$, an (a, b) pair almost immediately rules out the possibility of a high type, while (a, a) is less indicative of high reliability since the low type also could have produced it by chance. Early mixed signals lead agents to overweight this more-informative negative news for credibility whereas early identical signals lead agents to overweight less-informative positive news, making negative first impressions more robust.

In Proposition 3, we show that this asymmetry affects the degree to which first impressions persist in the limit. Enough mixed signals can always unravel a positive first impression. In contrast, arbitrarily high levels of persistence can occur for negative first impressions.

Proposition 3 (Long-run persistence and asymmetry of first impressions) *Let* $(\omega_0^A, \omega_0^H) = (1/2, \hat{\omega})$ *for any* $\hat{\omega} \in (0, 1)$*. Positive first impressions can eventually be undone, but negative first impressions may be arbitrarily persistent:*

1. *Positive first impressions: Suppose the agent observes* $n_a \geq 1$ *consecutive* a *signals, followed by* $m \geq 1$ *pairs of* (b, a) *signals:* $\mathbf{s}_n = (a, a, a, \ldots, b, a, b, a)$*. For a given* n_a*, there exists* $\hat{m}$ *such that when* $m > \hat{m}$*, the pre-screener under-trusts and is pessimistic about the most likely state for any* (q_L, q_H)*.*

2. *Negative first impressions: Suppose the agent observes* $n_b \geq 1$ *pairs of* (a, b) *signals, followed by* $m \geq 1$ *consecutive* a *signals:* $\mathbf{s}_n = (a, b, a, b, \ldots, a, a, a)$*. For a given* $n_b \geq 1$ *and* $m \geq 1$*, there exists some* $\check{q} > \frac{1}{2}$ *and* $\hat{q} < 1$ *such that the pre-screener under-trusts and is pessimistic about the most likely state if* (q_L, q_H) *satisfies one of the following sufficient conditions: (a)* $\hat{q} \leq q_L < q_H$*, or (b)* $q_L \leq \check{q}$ *and* $q_H > \hat{q}$*.*

2.2 Disagreements about states and credibility are correlated

Learning errors such as confirmation bias (Rabin and Schrag, 1999; Fryer et al., 2016) can also generate path-dependent beliefs where first impressions matter. Proposition 2 shows that, in our theory, first impressions *about credibility* drive disagreement. This leads to a key implication, distinct from several other theories, that disagreement about states and credibility are endogenously correlated.

Proposition 4 (Correlated disagreement) *Suppose all agents share a common prior* $(\omega_0^A, \omega_0^H) = (\hat{\theta}, \hat{\omega})$ *for any* $\hat{\theta} \in (0,1)$ *and* $\hat{\omega} \in (0,1)$*.*

1. *Suppose a pre-screener and a Bayesian observe a signal sequence* $\mathbf{s}_n^J$ *with* $n_a > n_b$*. The pre-screener under-trusts the source if and only if she is pessimistic about the more likely state:* $P^b(c = H|\mathbf{s}_n^J) < P^u(c = H|\mathbf{s}_n^J)$ *if and only if* $P^b(\theta = A|\mathbf{s}_n^J) < P^u(\theta = A|\mathbf{s}_n^J)$*. The pre-screener overtrusts the source if and only if she is optimistic in beliefs about the more likely state:* $P^b(c = H|\mathbf{s}_n^J) > P^u(c = H|\mathbf{s}_n^J)$ *if and only if* $P^b(\theta = A|\mathbf{s}_n^J) > P^u(\theta = A|\mathbf{s}_n^J)$*.*

2. *Suppose two pre-screeners,* J *and* M*, observe signal sequences* $\mathbf{s}_n^J$ *and* $\mathbf{s}_n^M$ *that have identical information content but different signal orders, where and* $n_a > n_b$*.*

   Agent J *trusts the source more than agent* M *does if and only if agent* J *believes state* A *is more likely than agent* M *does:* $P^b(c = H|\mathbf{s}_n^J) > P^b(c = H|\mathbf{s}_n^M)$ *if and only if* $P^b(\theta = A|\mathbf{s}_n^J) > P^b(\theta = A|\mathbf{s}_n^M)$*. Likewise, Agent* J *trusts the source less than agent* M *if and only if agent* J *believes state* A *is less likely than agent* M *does:* $P^b(c = H|\mathbf{s}_n^J) < P^b(c = H|\mathbf{s}_n^M)$ *if and only if* $P^b(\theta = A|\mathbf{s}_n^J) < P^b(\theta = A|\mathbf{s}_n^M)$*.*

The intuition is as follows. Consider first disagreement between a Bayesian and a pre-screener (Part 1), and suppose the signals objectively favor state A ($n_a > n_b$). If a pre-screener under-trusts a source, then she will be pessimistic about A, because she places too little weight on the information content. Conversely, if the pre-screener is pessimistic about A, she under-trusts the source.

Two pre-screeners also have correlated disagreement about the state and credibility (Part 2). Suppose again that the signals objectively favor state A. A pre-screener J who trusts the source more (less) than pre-screener M must also believe A is more (less) likely, and vice versa. These effects occur even when the two pre-screeners share common priors (Proposition 4 holds for any $\omega_0^\theta \in (0,1)$). Thus, our framework generates disagreement even when agents share identical information content, learning errors, and priors.

2.3 Endogenous over- and underreaction

Pre-screeners endogenously over- and under-react to signals, depending on how signals interact with their first-stage beliefs about credibility.

To build intuition, we contrast more closely with theories of confirmation bias and overconfidence. Confirmation bias is the tendency for individuals to interpret new information as confirming existing beliefs (Lord, Ross and Lepper, 1979; Griffin and Tversky, 1992; Rabin and Schrag, 1999; Fryer, Harms and Jackson, 2016). In Rabin and Schrag (1999), agents under-react to contradictory information because they probabilistically flip signals that oppose current beliefs about the state. With pre-screening, the key error is in how signals distort beliefs about credibility, irrespective of whether they confirm or contradict beliefs about the state. This can generate both over- and under-reaction to contradictory information, as we show below.

Overconfidence is the tendency for agents to over-react to signals that they believe are more informative than objectively warranted (e.g., Hirshleifer, 2015; Scheinkman and Xiong, 2003). Pre-screening explains why agents are endogenously overconfident in some sources while underconfident in others, based on first impressions of credibility (Proposition 2). Gervais and Odean (2001) endogenize overconfidence through a form of self-attribution bias so that successful financial traders become overconfident in their financial trading skills. In our framework, no actions are required, and underconfidence can also occur.

Proposition 5 formalizes these distinctions by characterizing when pre-screeners over- and under-react to new information. Consider the following thought experiment: Suppose a pre-screener begins with prior $\omega_0^A = 1/2$, observes signals $\mathbf{s}_n$ from one source, and has posterior ω_n^b. To contrast with confirmation bias, assume that the existing evidence $\mathbf{s}_n$ objectively strictly suggests A, which implies that $\{\mathbf{s}_n, s_{n+1}\}$ weakly suggests A. Does the pre-screener's beliefs about the state over- or under-react in response to s_{n+1}, compared to a Bayesian endowed with prior ω_n^b?

If the next signal s_{n+1} is a (b) and thus confirms (contradicts) beliefs, we say the agent has over-reacted (under-reacted) if $P^b[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] > P^u[\theta = A|\text{prior} = \omega_n^b]$, and under-reacted (over-reacted) if $P^b[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] < P^u[\theta = A|\text{prior} = \omega_n^b]$. As a benchmark, an agent in Rabin and Schrag (1999) under-reacts to s_{n+1} if it contradicts current beliefs, and correctly updates if it confirms current beliefs.

A pre-screener may over- or under-react to both confirmatory and contradictory news relative to the endowed Bayesian, depending on how the signal affects the first-stage trust ($\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$) in the combined evidence from the signal source ($\{\mathbf{s}_n, s_{n+1}\}$). We discuss the intuition briefly here and save details for the Appendix and the proofs.

Proposition 5 (Over- and under-reaction to new information) *Let $(\omega_0^A, \omega_0^H) = (1/2, \hat{\omega})$ for any $\hat{\omega} \in (0,1)$. Let $\mathbf{s}_n$ be a sequence of n observed signals, let s_{n+1} be the $(n+1)$th observed signal, and let ω_n^b equal the pre-screener's joint posterior after the sequence $\mathbf{s}_n$. WLOG, let the number of a's be greater than or equal to the number of b's in $\{\mathbf{s}_n, s_{n+1}\}$.*

1. *Relative to Endowed Bayesian:*

   (a) *$P^b[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] > P^u[\theta = A|prior = \omega_n^b, \{s_{n+1}\}]$ if $\{\mathbf{s}_n, s_{n+1}\}$ has strictly more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$,*

   (b) *$P^b[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] < P^u[\theta = A|prior = \omega_n^b, \{s_{n+1}\}]$ if $\{\mathbf{s}_n, s_{n+1}\}$ has strictly more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n)$,*

   (c) *$P^b[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] = P^u[\theta = A|prior = \omega_n^b, \{s_{n+1}\}]$ if $\{\mathbf{s}_n, s_{n+1}\}$ has an equal number of a's and b's or $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) = \kappa_H(\mathbf{s}_n)$.*

   Furthermore, $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) \geq \kappa_H(\mathbf{s}_n)$ if and only if $P^u(c = H|\{\mathbf{s}_n, s_{n+1}\}) \geq \omega_0^H$, with equality holding if and only if $P^u(c = H|\{\mathbf{s}_n, s_{n+1}\}) = \omega_0^H$.

2. *$P^b[c, \theta|\{\mathbf{s}_n, s_{n+1}\}] = P^b[c, \theta|prior = \omega_n^b, \{s_{n+1}\}]$ if and only if $P^b[c|\mathbf{s}_n] = \omega_0^c$.*

For signals that confirm beliefs ($s_{n+1} = a$), agents will over-react if first-stage trust is high (Part 1a) and under-react if first-stage trust is low (Part 1b), relative to the endowed Bayesian. The over-reaction is broadly consistent with confirmation bias and over-confidence, while the under-reaction is the opposite. The intuition for the under-reaction is that, even though the signal is confirmatory and the pre-screener and endowed Bayesian begin from the same beliefs, $\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$ may be too low relative to $\kappa_H(\{\mathbf{s}_n\})$ (the effective prior over the informativeness of $\{\mathbf{s}_n, s_{n+1}\}$ for the endowed Bayesian), weighing down the pre-screener's perceived informativeness of the total evidence $\{\mathbf{s}_n, s_{n+1}\}$, which supports A.

For signals that contradict beliefs ($s_{n+1} = b$), pre-screeners will under-react when first-stage trust is high (Part 1a), and over-react when first-stage trust is low (Part 1b). The under-reaction here is consistent with how a Rabin and Schrag (1999) agent behaves, while the over-reaction is the opposite. We label this over-reaction the *undercutting effect*. In this case, contradictory information undercuts the first-stage belief $\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$ and excessively undermines the history of evidence $\{\mathbf{s}_n, s_{n+1}\}$ (which supports A) in the second step, relative to the endowed Bayesian. Thus, contradictory signals can lead pre-screeners to wonder, "Can I trust anything they said?", and over-react. Experimentally, De Filippis et al. (2017) find evidence that individuals over-react to contradictory signals, citing a similar mechanism. Section 3 shows how prices of an asset can crash due to this effect if bad cash flow news causes agents to doubt the credibility of a source who previously reported good news.

Part 2 of Proposition 5 shows that the effect of a new signal s_{n+1} on a pre-screener's beliefs cannot be summarized simply by its effect on ω_n^b. This is because the pre-screener re-evaluates all the evidence $\{\mathbf{s}_n, s_{n+1}\}$ in light of the new first-stage belief $\kappa(\{\mathbf{s}_n, s_{n+1}\})$. In contrast, a Bayesian updates identically irrespective of whether she is endowed with a belief or observes a history of signals consistent with that belief: $P^u[\theta = A|\{\mathbf{s}_n, s_{n+1}\}] = P^u[\theta = A|prior = \omega_n^u, \{s_{n+1}\}]$, where ω_n^u equals the Bayesian posterior generated by $\mathbf{s}_n$.

In the Internet Appendix (Proposition 8) we characterize two sufficient conditions for which pre-screeners exhibit over- or underreaction, regardless of signal order. Underreaction can occur even when new information confirms beliefs if the proportion of a's is similar to that of b's, irrespective of signal order, because pre-screeners distrust the source too much.

3 Speculative trade, bubbles, and crashes

We characterize how pre-screening generates excessive speculative trade, price bubbles, and crashes compared to Bayesians with heterogeneous priors about credibility.

3.1 Speculative trade

We adopt a trading environment analogous to Harris and Raviv (1993). There are two groups of risk-neutral pre-screeners, X and Y, who trade, at dates $t = 1, 2, \ldots, T$, shares of a risky asset that make a single random payment of R immediately after the end of date T. If the state is A, then the payoff is $R = 1$. If the state is B, then the payoff is $R = 0$. There are a fixed number of shares available normalized to 1 with no short-sales. There is a risk-free asset whose return is zero. As in Harris and Raviv (1993), we assume Y has sufficient market power each period to make a take-it-or-leave-it offer to X, which focuses the analysis to comparisons of the two groups' beliefs. Traders "agree to disagree" about their beliefs.

At each date $t = 1, 2, \ldots T$, both groups X and Y first observe a common public signal $s_t \in \{a, b\}$, after which they update their beliefs and can trade at price p_t. Each signal s_t is independent and identically distributed conditional on the true payoff, and comes from a single source. Let the cumulative public signal path as of period t be denoted $\mathbf{s}_t = (s_1, \ldots, s_t)$ where $n_{a,t}$ and $n_{b,t}$ are the number of a's and b's in $\mathbf{s}_t$, respectively.

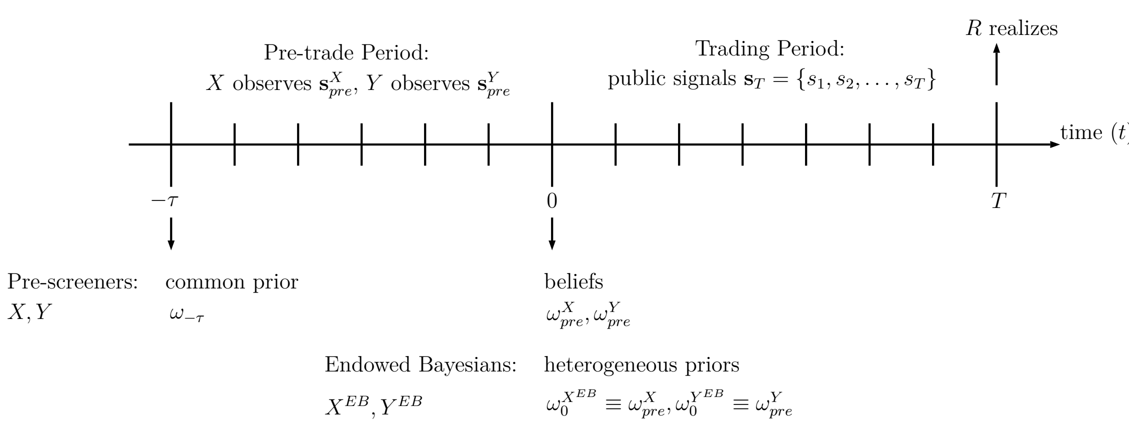


Figure 2: Timeline of trading game.

We ask: How much do X and Y trade, and how does this compare to trade between Bayesians? As a starting point, note that if all traders had common priors just before trade opens at $t = 1$, there would be trivially no trade for any $\mathbf{s}_t$ regardless of whether they are Bayesians or pre-screeners, as all agents observe the same signals in the same order during the trading period.

To make things interesting, we assume that pre-screeners observe signals during a set of "burn-in" periods before trade opens. Assume that there is a set of τ pre-trade periods starting in period $(-\tau+1) < 0$ and ending at the end of period 0 during which pre-screeners see signals before trade opens. X and Y share common priors $\omega_{-\tau}$ before they see any signals. During the pre-trade periods, X and Y separately observe one signal per period cumulating in signal path $\mathbf{s}^X_{pre}$ and $\mathbf{s}^Y_{pre}$, which generates beliefs ω^X_{pre} and ω^Y_{pre} at the end of period 0, before trade opens in period 1. We assume $\mathbf{s}^X_{pre}$ is a permutation of $\mathbf{s}^Y_{pre}$ (same information content but different orders). For example, X might observe $\mathbf{s}^X_{pre} = (a, a, b, b)$ while Y observes $\mathbf{s}^Y_{pre} = (a, b, a, b)$ in the pre-trade periods for $\tau = 4$. Figure 2 illustrates.

The assumption that $\mathbf{s}^X_{pre}$ is a permutation of $\mathbf{s}^Y_{pre}$ along with common priors $\omega_{-\tau}$ ensures that any trade between pre-screeners is due to differently-experienced first impressions of credibility early in their life prior to time 1, not differences in objective information or priors. To simplify, we focus on the case where: 1) the common prior is that each state is equally likely, and 2) pre-period signals do not change this belief but result in disagreement about credibility. Formally, suppose: 1) the common prior is $\omega_{-\tau} = \omega^\theta_{-\tau}\omega^c_{-\tau}$ with $\omega^\theta_{-\tau} = 1/2$, and 2) that $n^j_{a,pre} = n^j_{b,pre}$ for $j \in \{X, Y\}$, where $n^j_{a,pre}$ is the number of a signals observed in $\mathbf{s}^j_{pre}$. This makes beliefs just before trade start have the property that $P^b(H|\mathbf{s}^X_{pre}) \neq P^b(H|\mathbf{s}^Y_{pre})$, but $P^b(A|\mathbf{s}^X_{pre}) = P^b(A|\mathbf{s}^Y_{pre}) = 1/2$. Intuitively, the pre-screener understands that all signals come from the same source, and realizes that $n_{a,pre} = n_{b,pre}$ is equivalent to having no new information about the state, even if she has incorrect beliefs about credibility.[5]

To make apples-to-apples comparisons, we compare trade in two cases: 1) trade between X and Y as pre-screeners, and 2) trade between X^{EB} and Y^{EB}, two "Endowed Bayesians"

[5]This intuition holds more generally: Given any prior on the state, observing $n_a = n_b = k$ signals in any order will return the pre-screener back to this belief, just as it does with a Bayesian (see Lemma 2 in the proof of Proposition 6).

who start time 1 with heterogeneous priors equal to ω_{pre}^X and ω_{pre}^Y, the beliefs of the pre-screeners at the end of the pre-trade period. The key result is that there is more trade between X and Y than X^{EB} and Y^{EB}, and that this extra trade is "speculative trade" as defined by Harrison and Kreps (1978). We provide the formal statement and sketch the reasoning below, saving full details for the Appendix proof.

Proposition 6 (Speculative trade) *Suppose two groups of pre-screeners X and Y have observed pre-period signal paths $\mathbf{s}_{pre}^X$ and $\mathbf{s}_{pre}^Y$, where $n_{a,pre}^j = n_{b,pre}^j \geq 2$ for signal paths $j \in \{X, Y\}$, and then observe public signal path $\mathbf{s}_t$. Suppose two groups of Bayesians X^{EB} and Y^{EB} are endowed with priors that coincide with the pre-screeners' posterior beliefs after the pre-period, $\omega_0^{X^{EB}} = \omega_{pre}^X$ and $\omega_0^{Y^{EB}} = \omega_{pre}^Y$, and then observe $\mathbf{s}_t$.*

The price in the game with Endowed Bayesians is $p_t = E_t^{X^{EB}}(R)$, and the price in the game with pre-screeners is $p_t = E_t^X(R)$. Whenever groups X^{EB} and Y^{EB} trade, groups X and Y trade. Moreover, there exist paths where X and Y trade when X^{EB} and Y^{EB} do not:

1. *Trade cannot occur without disagreement: If $\mathbf{s}_{pre}^X = \mathbf{s}_{pre}^Y$, then neither pre-screeners nor endowed Bayesians trade.*

2. *If $P^b(H|\mathbf{s}_{pre}^X) \neq P^b(H|\mathbf{s}_{pre}^Y)$, then the* only *threshold at which X^{EB} and Y^{EB} trade is $n_{a,t} = n_{b,t}$. Pre-screeners X and Y also trade at the threshold $n_{a,t} = n_{b,t}$.*

3. *(Speculative trade, part a) If $P^b(H|\mathbf{s}_{pre}^X) > P^b(H|\mathbf{s}_{pre}^Y)$, then pre-screeners also trade when the following necessary conditions are satisfied:*

   (a) State A is objectively more likely: $n_{a,t} > n_{b,t}$.

   (b) Given $\mathbf{s}_{pre}^X$ and $\mathbf{s}_t$, group X over-reacts to confirming news: A necessary condition is $P^u(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t, s_{t+1} = a\}) > \omega_{-\tau}^H$. A ***sufficient*** *condition is that group X under-reacts to disconfirming news: $P^u(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t, s_{t+1} = b\}) \geq \omega_{-\tau}^H$.*

   When the necessary and sufficient conditions are satisfied given $\mathbf{s}_{pre}^X$ and $\mathbf{s}_t$, then there exists at least one signal path $\mathbf{s}_{pre}^Y$ such that Y holds the asset instead of X.

4. *(Speculative trade, part b) If* $P^b(H|\mathbf{s}^X_{pre}) < P^b(H|\mathbf{s}^Y_{pre})$*, then pre-screeners also trade when the following necessary conditions are satisfied:*

   (a) *State B is objectively more likely:* $n_{a,t} < n_{b,t}$.

   (b) *Given* $\mathbf{s}^X_{pre}$ *and* $\mathbf{s}_t$*, group* X *over-reacts to disconfirming news: A necessary condition is* $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = a\}) < \omega^H_{-\tau}$*. A* ***sufficient*** *condition is that group* X *under-reacts to confirming news:* $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = b\}) \leq \omega^H_{-\tau}$.

   When the necessary and sufficient conditions are satisfied given $\mathbf{s}^X_{pre}$ *and* $\mathbf{s}_t$*, then there exists at least one signal path* $\mathbf{s}^Y_{pre}$ *such that* Y *holds the asset instead of* X*.*

3.1.1 Bayesian benchmark

Endowed Bayesians $\{X^{EB}, Y^{EB}\}$ begin $t = 1$ with priors $\omega_0^{X^{EB}} = \omega^X_{pre}$ and $\omega_0^{Y^{EB}} = \omega^Y_{pre}$, respectively. If these priors were the same, they would not trade (Proposition 6 Part 1). Since instead we have $P^b(H|\mathbf{s}^X_{pre}) \neq P^b(H|\mathbf{s}^Y_{pre})$ but $P^b(A|\mathbf{s}^X_{pre}) = P^b(A|\mathbf{s}^Y_{pre}) = 1/2$, agents have heterogeneous priors about the credibility of the signal source but otherwise agree about the (marginal) probability of the state.

The environment is very close to that in Harris and Raviv (1993), where Bayesians have common priors on the state but have *fixed* heterogeneous beliefs about the credibility of the signal source. Traders learn about credibility in our setup, but as in Harris and Raviv (1993), trade only occurs when beliefs about the fundamental value of R "cross", which occurs at any period t at the threshold $n_{a,t} = n_{b,t}$ (Proposition 6 Part 2). We now discuss why.

As in Harris and Raviv (1993, Lemma 2), the asset's price is $p_t = E_t^{X^{EB}}(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})$, or group X^{EB}'s reservation price, as Y^{EB} makes a take-it-or-leave-it offer to X^{EB} (Lemma 3 in the Appendix). Group Y^{EB} buys from X^{EB} in period t if $WTP_t^{Y^{EB}} \geq p_t$, where $WTP_t^{Y^{EB}}$ is Y^{EB}'s willingness to pay at time t (we assume Y^{EB} holds the asset if equality holds, wlog). This willingness to pay is $WTP_t^{Y^{EB}} = \max\{E_t^{Y^{EB}}(p_{t+1}), E_t^{Y^{EB}}(R)\}$, where $E_t^{Y^{EB}}(p_{t+1})$ is Y^{EB}'s belief about the next period's price, and $E_t^{Y^{EB}}(R)$ is Y^{EB}'s belief about fundamentals.

Consider $E_t^{Y^{EB}}(p_{t+1})$. Since p_{t+1} will be X^{EB}'s reservation price after observing s_{t+1}, Y^{EB} needs to forecast what X^{EB} will believe next period about R. To do this, Y^{EB} needs to compute how she thinks X^{EB}'s beliefs will evolve according to X^{EB}'s learning rule, under Y^{EB}'s beliefs about the distribution of s_{t+1}: $E_t^{Y^{EB}}(p_{t+1}) = E_t^{Y^{EB}}(E_{t+1}^{X^{EB}}(R))$. Intuitively, Y^{EB} "agrees to disagree" with X^{EB} about credibility, with Y^{EB} believing she has the correct model and that X^{EB} has the wrong model, and forecasts X^{EB}'s "wrong" belief (from the standpoint of Y^{EB}) based on Y^{EB}'s "correct" belief about what signal will realize next.

If Y^{EB} believes that fundamental value is lower than X^{EB} does at time t, then: (1) $E_t^{Y^{EB}}(R) < p_t$, and (2) $E_t^{Y^{EB}}(p_{t+1}) < p_t$, as Y^{EB} thinks the next signal is less likely to be good news relative to X^{EB} so that X^{EB}'s beliefs will drift downward. Therefore, $WTP_t^{Y^{EB}} < p_t$ and Y^{EB} does not hold the asset. Once s_{t+1} realizes, Y^{EB} will buy from X^{EB} if the signal leads them to agree ($n_{a,t+1} = n_{b,t+1}$), as then $E_{t+1}^{Y^{EB}}(R) \geq p_{t+1}$ and $E_{t+1}^{Y^{EB}}(p_{t+2}) \geq p_{t+1}$. Analogous reasoning applies if Y^{EB} thinks fundamental value is higher than X^{EB} at t. Overall, trade only occurs when beliefs about fundamental value cross.

3.1.2 Pre-screening

Given any pre-period signals ($\mathbf{s}_{pre}^X$,$\mathbf{s}_{pre}^Y$) and public signal path $\mathbf{s}_t$, pre-screeners $\{X, Y\}$ trade any time EB's $\{X^{EB}, Y^{EB}\}$ trade, but also trade when EB's do not.

We continue to assume that agents "agree to disagree" about learning. In Harris and Raviv (1993), it is common knowledge that each agent understands all other agents' learning models and can anticipate other agents' beliefs, but views her own model as correct and rational and others' models as necessarily incorrect and irrational. The closest analog in our setting is that X and Y view themselves individually as rational and Bayesian, but understands and anticipates the other group's pre-screened beliefs. As we discuss later, this assumption is also internally consistent with hindsight bias and external experimental evidence.

If X and Y observe the same pre-period signals in the same order, there is no trade since all agents share exactly the same beliefs during the trading period (Proposition 6 Part 1). To make things interesting, suppose pre-period signals are permutations of each other as described above. Pre-screeners trade whenever their beliefs about fundamental value cross ($n_{a,t} = n_{b,t}$), which is also when EB's trade (Part 2).

However, pre-screeners engage in extra "speculative behavior" defined by Harrison and Kreps (1978): "an investor may buy the stock now so as to sell it later for more than he thinks it is actually worth, thereby reaping capital gains."[6] For example, Y might choose to hold the asset even when she thinks the fundamental value is lower than X does. Even if Y thinks the next signal is less likely to be good cash flow news than X, Y may also think that X will over-react to such news, leading her to forecast an upward drift in the price.

Proposition 6 Parts 3 and 4 show that speculative trade can occur in two cases. In Part 3, Y believes the source is less credible than X does after the pre-trade signals, so Y's belief about fundamental value will be lower than X's if they observe the same good cash flow news after trade opens. However, Y may still buy the asset speculatively: $E_t^Y(p_{t+1}) > E_t^X(R) = p_t$ even though $E_t^Y(R) < E_t^X(R) = p_t$, so long as disagreement between X and Y is not too large and there are enough a's relative to b's. The reason is that Y (correctly) believes X will over-react to further good cash flow news (using Proposition 5 Part 1a), and Y is not too skeptical about prospects of such further news, leading Y to expect an upward drift in the price. However, X believes they are Bayesians in the future and fail to see this upward drift. Part 4 is an analogous case, when Y believes the source is more credible than X does after the pre-trade signals and bad cash flow news arrives after trade opens.

Overall, the implication of Proposition 6 is that the extent of speculative trade in the game with pre-screeners, and therefore the amount of excess trading volume beyond what occurs in the game with Endowed Bayesians, depends on the extent of disagreement about credibility and how disagreement originates. In the game with pre-screeners, initial disagreement about credibility originates from differing first impressions of credibility from signals prior to trade, $\omega_{pre}^X \neq \omega_{pre}^Y$. Even though we endow identical disagreement in the game with Bayesians, $\omega_0^{jEB} = \omega_{pre}^j$ for $j \in \{X, Y\}$, and even though all traders see a common signal path when trade is open, outcomes differ across the two games because of how new signals interact with pre-screeners' first impressions of credibility.

The assumption that each group "agrees to disagree" by viewing themselves as individu-

[6] In contrast, Harris and Raviv (1993) use the term speculation to describe the fact that their agents trade because they are betting against others' different models.

ally rational but understanding and anticipating other groups' pre-screened beliefs is strong. We make three observations. First, the idea that each agent understands all other agents' models but views their own (objectively wrong) model as correct is intrinsic to the idea that agents in fact disagree with other agents' models.[7] Second, the assumption that individuals view themselves as rational is internally consistent with hindsight bias. Having observed any signal path $\mathbf{s}_t$, the pre-screener thinks she is a Bayesian whose prior was always $\kappa_c(\mathbf{s}_t)$, as in Equation 3. Because she treats $\kappa_c(\mathbf{s}_t)$ as her prior, naturally she also thinks that $\kappa_c(\mathbf{s}_t)$ will remain her prior in the future and thus fails to understand that she pre-screens. Third, this assumption is also consistent with external experimental evidence in psychology (Pronin et al., 2002; Ehrlinger et al., 2005) and economics (Danz et al., 2017; Fedyk, 2017) about the "bias blind spot", the idea that individuals recognize cognitive and motivational biases much more in others than in themselves.[8]

3.2 Bubbles and crashes

Let p_t^b be the price when the pre-screeners trade with each other, and p_t^{EB} be the price when Endowed Bayesians trade with each other. Proposition 7 considers when over- and undervaluation occur, and whether trading generates bubbles and crashes. Since the definition of a bubble can be subjective, we tie our hands and ask whether our model produces the six features of bubbles outlined by Barberis (2018, p.88): (i) existence of initial good news about asset cash flows, (ii) a sharp rise in asset prices followed by a reversal, (iii) abnormal volume as the price of the asset rises, (iv) highly extrapolative expectations about returns during the episode, (v) some sophisticated investors increasing asset exposure during the price rise, and (vi) media reports that the asset is overvalued during the price rise.

We take (i) as given and see to what extent trading between pre-screeners and trading

[7] To correctly calculate group X's beliefs, agent Y must also be aware of the signal path $\mathbf{s}_{pre}^X$ that X observed in the pre-period. This is also strong but in keeping with the common knowledge assumption.

[8] West et al. (2012) find that the degree of cognitive sophistication does not seem to attenuate the bias blind spot. Danz et al. (2017) show that individuals exhibit the bias blind spot with respect to information projection - while an agent naively believes that others are privy to her private information, she also correctly anticipates that others misperceive that their private information is more widely known than it actually is. Fedyk (2017) shows that individuals exhibit the bias blind spot with respect to self-control.

between Endowed Bayesians can generate features (ii)-(vi). Figure 1 provides a specific numerical example, which we discuss in narrative form to build intuition. We consider the case where trading begins with X trusting the signal source more than Y does but where X and Y agree about the state based on signals in the pre-trade period. During the trading period, public signals arrive with a string of good cash flow news as required by feature (i) (a string of 10 a's), followed by bad cash flow news (a string of 10 b's).

The game begins with Y holding the asset in period 0. In period 1, good cash flow news arrives and X thinks $R = 1$ is more likely than Y does, leading Y to sell the asset to X, in both the pre-screening and EB frameworks. Good cash flow news arrives through period 10, and through this point, prices rise sharply and trade occurs in the pre-screening game, both of which are abnormal relative to the EB game (features ii and iii). The trade in period 3 reflects group Y increasing asset exposure (feature v): Y speculatively "rides the bubble" and buys the asset from X even though the price is higher than Y's belief about fundamental value (Proposition 6 Part 3). As more good cash flow news arrives, prices rise substantially beyond the price in the EB game. The reason is that agents develop too much trust in the source and become too over-optimistic about cash flows.

Bad cash flow news begins to arrive after period 10. Prices in the pre-screening game remain high and are resistant to falling relative to the EB game, analogous to feature (vi). This is because X thinks the history of good cash flow news provided by the source is extremely credible and under-reacts to the bad cash flow news, behavior that is akin to confirmation bias (Proposition 5 Part 1a). As enough bad cash flow news comes in, X begins to doubt whether they believe anything the source reported before, due to the contradiction with the previously-reported good cash flow news (the "undercutting effect" of Proposition 5 Part 1b). Anticipating this possibility, Y sells the asset to X in period 13. Bad news continues to arrive, and prices steeply decline in periods 16-18 as X's belief about source credibility collapses. In period 20, X and Y share the same beliefs about the state, and Y buys the asset back from X.

In sum, the paths depicted in Figure 1 in the pre-screening game reasonably constitute a bubble and crash. The narrative fits all but feature (iv) of the Barberis (2018) criteria for a bubble. Prices rise more steeply than in the EB game, and additional speculative trade occurs. Furthermore, our model can produce steep price declines resembling a crash. By way of comparison, the Internet Appendix shows that confirmation bias does not produce such sudden price declines, because agents always under-react to contradictory information.

Proposition 7 generalizes these results beyond the specific signals and parameters considered in the Figure. Part 1 shows over-valuation (under-valuation) occurs when consistent (inconsistent) good news about cash flows leads the pre-screener to trust (distrust) the source so much that she over-reacts (under-reacts) to additional good news. Part 2 shows that this can accelerate into bubbles and crashes: when overvaluation occurs, rises and subsequent falls in prices are steeper when traders are pre-screeners rather than Bayesian. Over-reaction to good cash flow news generates the bubble; as in Figure 1, prices initially under-react to bad cash flow news before beliefs over-react and prices crash. Part 3 shows that when such a pricing bubble occurs, there exists at least one pre-path $\mathbf{s}_{pre}^{Y}$ such that speculative trade occurs during the bubble (agents "ride the bubble").

Proposition 7 (Bubbles and crashes) *Suppose two pre-screeners X and Y have observed pre-period signal paths $\mathbf{s}_{pre}^{X}$ and $\mathbf{s}_{pre}^{Y}$, where $n_{a,pre}^{j} = n_{b,pre}^{j} \geq 2$ for signal paths $j \in \{X, Y\}$, and then public signal path $\mathbf{s}_t$. Suppose two Bayesians X^{EB} and Y^{EB} are endowed with priors that coincide with the pre-screeners' posterior beliefs after the pre-period, $\omega_0^{X^{EB}} = \omega_{pre}^{X}$ and $\omega_0^{Y^{EB}} = \omega_{pre}^{Y}$, and then observe $\mathbf{s}_t$.*

Let p_t^b be the price when the pre-screeners trade with each other, and p_t^{EB} be the price when Endowed Bayesians trade with each other.

1. *(Over- and under-valuation) Wlog, suppose $n_{a,t} > n_{b,t}$. Under-valuation ($p_t^b < p_t^{EB}$) occurs if and only if $P^b(H|\{\mathbf{s}_{pre}^{X}, \mathbf{s}_t\}) < P^u(H|prior = \omega_{pre}^{X}, \mathbf{s}_t)$. Over-valuation ($p_t^b > p_t^{EB}$) occurs if and only if $P^b(H|\{\mathbf{s}_{pre}^{X}, \mathbf{s}_t\}) > P^u(H|prior = \omega_{pre}^{X}, \mathbf{s}_t)$.*

2. *(Bubbles and crashes) Consider a path $\mathbf{s}_T$ such that $n_{a,t} > n_{b,t}$ for $t \in (0, T)$ with $n_{a,T} = n_{b,T}$. If there exists $\hat{t} \in (0, T)$ such that $p_{\hat{t}}^b > p_{\hat{t}}^{EB} > 1/2$ where $p_{\hat{t}}^k \equiv \max p_t^k$ for $k \in \{b, EB\}$, the average price change of p_t^b must be strictly greater than the average price change of p_t^{EB} for $t \in [0, \hat{t}]$ (bubbles) and $t \in [\hat{t}, T]$ (crashes). Moreover, the*

pre-screener exhibits initial under-reaction relative to the endowed Bayesian after the peak: $|p^b_{\hat{t}+1} - p^b_{\hat{t}}| < |p^{EB}_{\hat{t}+1} - p^{EB}_{\hat{t}}|$.

3. *("Riding the bubble") Given any signal path* $\mathbf{s}_T$ *such that* $p^b_{\hat{t}} > p^{EB}_{\hat{t}} > 1/2$*, there exists at least one signal path* $\mathbf{s}^Y_{pre}$ *such that speculative trade between pre-screeners occurs (e.g.,* Y *holds the asset at* $t = \hat{t}$*, at least).*

4 Extensions and discussion

4.1 Can new information sources resolve disagreement?

Our model makes contrasting predictions about how signals from existing versus new sources resolve disagreements. Lemma 2 in the proof of Proposition 6 shows that, when signals come from a single source, the pre-screener's marginal beliefs about the state will equal the Bayesian's and disagreement will be "resolved" if the source reports the same number of a and b signals, $n_a = n_b = k$. This is because the pre-screener understands that all signals originate from the same source just like a Bayesian.

Instead of receiving all signals from one source, suppose an agent receives signals from a second source, so there are now three sources of uncertainty: the state, and the credibilities of sources 1 and 2. Pre-screening extends naturally from one to multiple sources. First, a pre-screener updates on the joint belief about source credibilities, using Bayes' Rule. Second, she uses this updated first-stage belief to form a joint posterior belief on the state and credibilities. In the Internet Appendix, we ask: assuming source 1 and source 2 are ex-ante identical, if source 1 first reports $k > 1$ signals of a, and then source 2 reports k signals of b, does source 2 resolve disagreement between a pre-screener and a Bayesian?

Proposition 9 in the Internet Appendix shows that the answer is, surprisingly, no. A Bayesian with $\omega^A_0 = 1/2$ infers that the sources cannot both be highly reliable. Despite both sources delivering internally consistent messages, their signals contradict each other, and the Bayesian understands that there is insufficient evidence to deduce which source is wrong. As a result, she concludes that neither state is more likely than the other.

However, the positive impression from the first source's consistency inflates the pre-screener's trust in the early source and deflates trust in the later source relative to the Bayesian inference. Like the Bayesian, the pre-screener concludes that the sources cannot both be highly reliable, but incorrectly concludes that the first source is more credible than the second and therefore differentially weights information in favor of the first. This asymmetry means that the second source cannot completely unravel the first source's signals. Furthermore, this asymmetry persists in the limit: Information that should lead to more uncertainty about credibilities and no change in beliefs about the state "backfires" and instead leads the pre-screener to be *more sure of and more wrong in* her beliefs along both dimensions when observing opposing information from different sources sequentially.

Proposition 9 suggests that signals that should objectively deflate price bubbles may fail to do unless they come from a source whose previous signals supported prices. Empirical evidence in other contexts, particularly political science, suggests that rumors are stubbornly resistant to debunking by outsiders (Berinsky, 2012). Attempts at correcting false beliefs can backfire and harden beliefs (Nyhan and Reifler, 2010), and are more successful when they include retractions from the original source (Simonsohn, 2011; Levine and Valle, 1975).

4.2 What if sources are slanted?

So far, we have considered credibility to be the source's underlying ability to determine the state, which we also call *reliability*. However, there are several settings where signals from a source may also contain a tilt, or *slant*, towards a given state. For example, analysts, financial advisors, or other experts may be too sanguine about future cash flows. The agent may not know the direction or magnitude of slant, and may try to infer it from the reports.

One can generalize credibility to be a triplet $c = (q_c, \alpha_c, \gamma_c) \in \mathcal{C} \subset [1/2, 1) \times [0, 1) \times [0, 1)$ that captures both reliability and slant. An expert who has a slant towards A has type $c = (q_c, \alpha_c, 0)$ and reports a when the true state is A with probability $P(s_t = a|c, A) = q_c + (1 - q_c)\alpha_c$. Either the private signal is a (with probability q_c), or the private signal is b and she flips it to a (with conditional probability α_c and total probability $(1 - q_c)\alpha_c$). The probability that she reports b when the true state is B is $P(s_t = b|c, B) = q_c(1 - \alpha_c)$.

Analogously, a B-slanted expert with type $c = (q_c, 0, \gamma_c)$ has $P(s_t = a|c, A) = q_c(1 - \gamma_c)$ and $P(s_t = b|c, B) = q_c + (1 - q_c)\gamma_c$. Experts can only be slanted one way: $\alpha_c \times \gamma_c = 0$.[9]

Proposition 10 in the Internet Appendix shows that, unlike in the case of unknown reliability alone, the presence of slant can lead the pre-screener's beliefs to be sufficiently wrong that she becomes certain of the state after observing a set of signals from a single expert, even when the evidence should not objectively change beliefs about the state from priors. To illustrate, consider the particularly stark case in which a single expert sends k identical a signals, followed by k identical b signals. Suppose that there are two possible expert types ($c \in \{v, w\}$), where reliability is identical and the degree of slant is symmetric ($\alpha = \gamma \geq 0$): $v = (q, \alpha, 0)$, $w = (q, 0, \gamma)$. Assume that each state and each type is equally likely ex-ante: $\omega_0^A = 1/2$ and $\omega_0^c = 1/2$.

The Bayesian realizes that there is no new information about either the state or the expert's slant. In contrast, the pre-screener can strongly believe that these signals reveal the state. If a pre-screener becomes convinced too early that a source is A-slanted, she will interpret subsequent b's as more indicative of state B than warranted, and vice versa. This implies that if pre-screener J first observes an a streak and then observes a b streak, while a pre-screener M observes the b streak before the a streak, they will have opposing beliefs about the state, even if both see the same number of a's and b's. Proposition 11 in the Internet Appendix shows that insights from Section 4.1 also hold in the presence of slant.

4.3 Comparison with other frameworks

Hong and Stein (2007) organize explanations for disagreement into gradual information flow, limited attention, and heterogeneous priors. Xiong (2013) adds overconfidence and distorted information transmission. Our work emphasizes erroneous learning about the credibility of sources, which we argue captures important features of real-world disagreement.[10]

Gradual information flow and inattention. A key feature of the gradual information

[9] In the scale example of Section 1.3, similar insights hold when the scale is possibly slanted; e.g., $\mathcal{C} = \{v = (0.9, 0, 0), w = (0.9, 0, 0.8)\}$.

[10] We refer the reader to Section 2.3 for a comparison of pre-screening with confirmation bias.

flow model of Hong and Stein (1999) is that investors pay attention to only a subset of signals, analogous to the broader literature on limited attention (Hirshleifer and Teoh, 2003; Peng and Xiong, 2006; Hirshleifer, 2015). Recent work on inattention includes Schwartzstein (2014), who considers a setting where agents learn to selectively pay attention to variables after evaluating their predictive ability. Kominers et al. (2016) assume that agents trade off attention costs and belief accuracy, and screen out signals with low decision value.

In contrast, a key feature of our model is that agents disagree about the credibility of signals *that they all see*. Recent empirical evidence suggests that this is an important feature of real-world disagreement. Cookson and Niessner (2016) provide evidence that differences in signal interpretation is an important component of overall disagreement about firm stock prices. Kandel and Zilberfarb (1999), Lahiri and Sheng (2008) and Patton and Timmermann (2010) provide evidence that differences in information sets do not explain disagreement among macroeconomic forecasters, and instead emphasize the importance of differences in how forecasters interpret information. During the dot-com boom, internet stocks were a key focus of media attention (Bhattacharya et al., 2009).

Heterogeneous priors, learning, and overconfidence. Harris and Raviv (1993), Kandel and Pearson (1995) and Scheinkman and Xiong (2003) employ heterogeneous learning to explain speculative trading and asset price bubbles. In these models, agents are Bayesian but are certain in their exogenously-endowed belief of which likelihood function generated the data. Acemoglu et al. (2016), along with Sethi and Yildiz (2016) and Suen (2004), study Bayesian agents who learn about uncertain credibility with heterogeneous priors and find that beliefs may not converge in the long-run.

Our framework does not rely on the exogenous endowment of heterogeneous priors or fixed beliefs to generate disagreement. Section 3 outlines how this generates distinct predictions in a trading game. More broadly, our model provides an explanation for why some agents trust signals more than others even when they start with common priors and observe common objective information. Even when agents begin with common priors, experimental evidence suggests that individuals may disagree when Bayesians should agree (Andreoni and Mylovanov, 2012). This approach complements the heterogeneous priors approach (Morris,

1995) by providing a theory for the origins of such heterogeneous priors.

Strategic motives and persuasion. A large literature documents the importance of strategic motives among financial analysts, advisors, and financial media (Lin and McNichols, 1998; Hong and Kubik, 2003; Tetlock, 2014). Gentzkow and Shapiro (2006) and Mullainathan and Shleifer (2005) emphasize that the media may slant news to build a reputation or to cater to consumers' preferences for beliefs. Several of these models assume that Bayesian receivers begin with heterogeneous priors or preferences. Intuitively, without such heterogeneity, media slant generates biased beliefs, but not disagreement. We raise the possibility that erroneous learning about credibility leads to demand distortions that complement the strategic motive of information suppliers in generating disagreement.

Bounded rationality. A distinguished set of models seeks to explain behavior using bounded rationality, the idea that agents follow simple heuristics due to cognitive limitations (Simon, 1957; Gigerenzer and Selten, 2002; Selten, 2002). In contrast, our approach is in the spirit of work that seeks to model systematic conceptual errors that lead to deviations from Bayes' Rule, even though they may involve more literal computations in the model. Enke and Zimmermann (forthcoming) and Enke (2017) emphasize the importance of conceptual errors rather than computational mistakes. Examples of other work where non-Bayesian agents employ large calculations include Brunnermeier and Parker (2005), where agents calculate optimal beliefs, and Rabin (2002), where agents mis-perceive draws from an urn that occur with replacement as occurring without replacement, requiring agents to keep track of what has been drawn.[11]

Koçak (2017) also studies uncertainty about source reliability and the state, but assumes that agents not only have a form of hindsight bias, but also do not understand correlations. After seeing each signal, they form separate marginal posteriors on credibility and the state sequentially, before combining them in a posterior that assumes the two are independent. Economically, this is as if the agent forgets the source of information each period even

[11] As further examples, agents in Bénabou and Tirole (2002) forget bad news, but also think about whether any information lost affects the value of new information. Agents in Kominers et al. (2016) process the decision value of signals before updating. Fryer et al. (2016) consider a case where agents optimally consider how to interpret ambiguous signals, trading off short-run and long-run informational value.

though she is attempting to learn the source's credibility. This assumption of "correlation neglect" (Ortoleva and Snowberg, 2015; DeMarzo et al., 2003) drives several predictions. For example, an agent in Koçak (2017) who observes $n_a = n_b$ signals thinks A is more likely if the a signals come first, because the correlation between credibility and the state is reset to zero each period. A pre-screener does not commit this error because she understands correlations. Instead, our core focus is on the idea that agents double-dip the data when credibility is uncertain, and we show that this idea generates disagreement, trading, and price bubbles, without the additional assumption of correlation neglect.

5 Conclusion

The key predictions of pre-screening are that: 1) Differing first impressions about credibility generate disagreement; 2) Disagreement about states of the world and credibility are endogenously correlated, and 3) Pre-screeners can over- and under-react depending on how signals interact with beliefs about credibility. In a trading game, pre-screening can generate price bubbles and crashes along with speculative trades, with traders "riding the bubble" along the way, even in an environment where Bayesians with heterogeneous priors would not do so (Harris and Raviv, 1993). New information sources may not resolve disagreement when they should, and agents can become certain of incorrect states if sources are slanted.

Our work opens a new perspective on the origins of disagreement that emphasizes how erroneous learning about the credibility of information sources endogenously explains joint disagreement about credibility and an unknown state of the world. This may play a role in explaining why rational learning may not best describe several of today's highest-pitched disagreements where opposing sides vehemently disagree about both the subject and the credibility of such sources. For example, in the debate about climate change, individuals compare each other's beliefs to a "religious mantra" (Bell, 2011) and an "intellectual stance…uncomfortably close to Hitler's" (Snyder, 2015). Future research on how errors in learning about the credibility of information sources may help explain polarizing disagreements spanning topics in finance, climate change, economics, medicine, and politics.

References

Abreu, Dilip and Markus K. Brunnermeier, "Bubbles and Crashes," *Econometrica*, 2003, *71* (1), 173–204.

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz, "Fragility of Asymptotic Agreement under Bayesian Learning," *Theoretical Economics*, 2016, *11*, 187–227.

Andreoni, James and Tymofiy Mylovanov, "Diverging Opinions," *American Economic Journal: Microeconomics*, 2012, *4* (1), 209–232.

Barberis, Nicholas, "Psychology-Based Models of Asset Prices and Trading Volume," in B. Douglas Bernheim, Stefano DellaVigna, and David Liabson, eds., *Handbook of Behavioral Economics - Foundations and Applications 1, Volume 1*, North Holland, 2018.

Bell, Larry, "Climate Change As Religion: The Gospel According To Gore," *Forbes Magazine*, April 26 2011. Available online: https://www.forbes.com/sites/larrybell/2011/04/26/climate-change-as-religion-the-gospel-according-to-gore [Last accessed: May 2017].

Bénabou, Roland and Jean Tirole, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

Berinsky, Adam, "Rumors, Truths, and Reality: A Study of Political Misinformation," 2012. Massachusetts Institute of Technology.

Bhattacharya, Utpal, Neal Galpin, Rina Ray, and Xiaoyun Yu, "The Role of the Media in the Internet IPO Bubble," *Journal of Financial and Quantitative Analysis*, 2009, *44* (3), 657–682.

Biais, Bruno and Martin Weber, "Hindsight Bias, Risk Perception, and Investment Performance," *Management Science*, 2009, *55* (6), 1018–1029.

Blank, Hartmut and Steffen Nestler, "Cognitive Process Models of Hindsight Bias," *Social Cognition*, 2007, *25* (1), 132–146.

Brunnermeier, Markus K. and Jonathan A. Parker, "Optimal expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

_ **and Stefan Nagel**, "Hedge Funds and the Technology Bubble," *Journal of Finance*, 2005, *59* (5), 2013–2040.

Camerer, Colin, George Loewenstein, and Martin Weber, "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 1989, *97* (5), 1232–1254.

Campbell, John Y. and Robert J. Shiller, "Valuation ratios and the long-run stock market outlook," *Journal of Portfolio Management*, 1998, *24* (2), 11–26.

_ **and** _ , "Valuation ratios and the long-run stock market outlook — an update," 2001. NBER Working Paper No. 8221.

Carlin, Bradley P. and Thomas A. Louis, "Empirical Bayes: Past, Present and Future," *Journal of the American Statistical Association*, 2000, *95* (452), 1286–1289.

Cheng, Ing-Haw and Alice Hsiaw, "Distrust in Experts and the Origins of Disagreement," 2017.

Clement, Michael B., "Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?," *Journal of Accounting and Economics*, 1999, *27* (3), 285–303.

Cookson, J. Anthony and Marina Niessner, "Why don't we agree? Evidence from a social network of investors," 2016.

Danz, David, Kristof Madarász, and Stephanie W. Wang, "The Biases of Others: Anticipating Informational Projection in an Agency Setting," 2017.

DellaVigna, Stefano and Devin Pope, "Predicting Experimental Results: Who Knows What?," *Journal of Political Economy*, forthcoming.

DeMarzo, Peter M., Dimitri Vayanos, and Zwiebel, "Persuasion Bias, Social Influence, and Unidimensional Opinions," *The Quarterly Journal of Economics*, 2003, *118* (3), 909–968.

Ehrlinger, Joyce, Thomas Gilovich, and Lee Ross, "Peering Into the Bias Blind Spot: People's Assessments of Bias in Themselves and Others," *Personality and Social Psychology Bulletin*, 2005, *31* (5), 680–692.

Enke, Benjamin, "What You See Is All There Is," 2017.

— **and Florian Zimmermann**, "Correlation Neglect in Belief Formation," *Review of Economic Studies*, forthcoming.

Fama, Eugene F. and Kenneth R. French, "Luck versus Skill in the Cross-Section of Mutual Fund Returns," *Journal of Finance*, 2010, *65* (5), 1915–1947.

Fedyk, Anastassia, "Asymmetric Naïveté: Beliefs about Self-Control," 2017.

Filippis, Roberta De, Antonio Guarino, Philippe Jehiel, and Toru Kitagawa, "Updating Ambiguous Beliefs in a Social Learning Experiment," 2017.

Fischhoff, Baruch, "Hindsight $\neq$ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty," *Journal of Experimental Psychology*, 1975, *1* (3), 288–299.

— **and Ruth Beyth**, ""I Knew It Would Happen": Remembered Probabilities of Once-Future Things," *Organizational Behavior and Human Performance*, 1975, *1*, 1–16.

Fryer, Roland G., Philipp Harms, and Matthew O. Jackson, "Updating Beliefs when Evidence is Open to Interpretation: Implications for Bias and Polarization," 2016.

Gentzkow, Matthew and Jesse M. Shapiro, "Media Bias and Reputation," *Journal of Political Economy*, 2006, *114* (2), 280–316.

Gervais, Simon and Terrance Odean, "Learning to be Overconfident," *Review of Financial Studies*, 2001, *14*, 1–27.

Gigerenzer, Gerd and Reinhard Selten, "Rethinking Rationality," in Gerd Gigerenzer and Reinhard Selten, eds., *Bounded Rationality: The Adaptive Toolbox*, MIT Press, 2002, pp. 1–12.

Griffin, Dale and Amos Tversky, "The weighing of evidence and the determinants of confidence," *Cognitive Psychology*, July 1992, *24* (3), 411–435.

Harris, Milton and Artur Raviv, "Differences of Opinion Make a Horse Race," *Review of Financial Studies*, 1993, *6* (3), 473–506.

Harrison, J. Michael and David M. Kreps, "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *The Quarterly Journal of Economics*, 1978, *92* (2), 323–336.

Hawkins, Scott A. and Reid Hastie, "Hindsight: Biased Judgments of Past Events After the Outcomes Are Known," *Psychological Bulletin*, 1990, *107* (3), 311–327.

Hertwig, Ralph, Carola Fanselow, and Ulrich Hoffrage, "Hindsight bias: How knowledge and heuristics affect our reconstruction of the past," *Memory*, 2003, *11* (4/5), 357–377.

Hirshleifer, David, "Behavioral Finance," *Annual Review of Financial Economics*, 2015, *7*, 133–159.

_ and Siew Hong Teoh, "Limited Attention, Information Disclosure and Financial Report," *Journal of Accounting and Economics*, 2003, *36* (1–3), 337–386.

Hoffrage, Ulrich and Ralph Hertwig, "Hindsight Bias: A Price Worth Paying for Fast and Frugal Memory," in Gerd Gigerenzer, Peter M. Todd, and The ABC Research Group, eds., *Simple heuristics that make us smart*, Oxford University Press, 1999, pp. 191–208.

_ , _ , and Gerd Gigerenzer, "Hindsight Bias: A By-Product of Knowledge Updating?," *Journal of Experimental Psychology*, 2000, *26* (3), 566–581.

Hong, Harrison and Jeffrey Kubik, "Analyzing the analysts: Career concerns and biased earnings forecasts," *Journal of Finance*, 2003, *58*, 313–351.

_ and Jeremy C. Stein, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *Journal of Finance*, 1999, *54* (6), 2143–2184.

_ and Jeremy Stein, "Disagreement and the Stock Market," *Journal of Economic Perspectives*, 2007, *21* (2), 109–128.

_ , José A. Scheinkman, and Wei Xiong, "Advisors and asset prices: a model of the origins of bubbles," *Journal of Financial Economics*, 2008, *89* (2), 268–287.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, "Time-Varying Fund Manager Skill," *Journal of Finance*, 2014, *69* (4), 1455–1484.

Kandel, Eugene and Ben-Zion Zilberfarb, "Differential Interpration of Information in Inflation Forecasts," *The Review of Economics and Statistics*, 1999, *81* (2), 217–226.

_ and Neil Pearson, "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy*, 1995, *103* (4), 831–872.

Koçak, Korhan, "Sequential Updating," February 2017.

Kominers, Scott Duke, Xiaosheng Mu, and Alexander Peysakhovich, "Paying (for) Attention: The Impact of Information Processing Costs on Bayesian Inference," 2016.

Kyle, Albert S., Anna A. Obizhaeva, and Yajun Wang, "Smooth trading with overconfidence and market power," *Review of Economic Studies*, 2017, *85* (1), 611–662.

Lahiri, Kajal and Xuguang Sheng, "Evolution of Forecast Disagreement in a Bayesian learning model," *Journal of Econometrics*, 2008, *144*, 325–340.

Levine, John M. and Ronald S. Valle, "The Convert as a Credible Communicator," *Social Behavior and Personality*, 1975, *3* (1), 81–90.

Lewis, Michael, "In defense of the boom," *New York Times Magazine*, October 27 2002.

Lin, Hsiou-wei and Maureen F. McNichols, "Underwriting relationships, analysts' earnings forecasts and investment recommendations," *Journal of Accounting and Economics*, 1998, *25* (1), 101–127.

Lindley, Dennis Victor, "Compound Decisions and Empirical Bayes: Discussion," *Journal of the Royal Statistical Society*, 1969, *31* (3), 397–425.

Lord, Charles G., Lee Ross, and Mark R. Lepper, "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence," *Journal of Personality and Social Psychology*, 1979, *37* (11), 2098–2109.

Madarász, Kristóf, "Information projection: Model and applications," *Review of Economic Studies*, 2012, *79* (3), 961–985.

Malmendier, Ulrike and Devin M. Shanthikumar, "Are small investors naive about incentives?," *Journal of Financial Economics*, 2007, *85*, 457–489.

— **and** — , "Do security analysts speak in two tongues?," *Review of Financial Studies*, 2014, *27* (5), 1287–1322.

McNichols, Maureen F. and Patricia C. O'Brien, "Self-Selection and Analyst Coverage," *Journal of Accounting Research*, 1997, *35*, 167–199.

Morgan, John and Phillip C. Stocken, "An analysis of stock recommendations," *RAND Journal of Economics*, 2003, *34* (1), 183–203.

Morris, Stephen, "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 1995, *11*, 227–253.

Mullainathan, Sendhil and Andrei Shleifer, "The Market for News," *The American Economic Review*, 2005, *95* (4), 1031–1053.

Nyhan, Brendan and Jason Reifler, "When Corrections Fail: The Persistence of Political Misperceptions," *Political Behavior*, 2010, *32* (2), 303–330.

Ortoleva, Pietro and Erik Snowberg, "Overconfidence in Political Behavior," *American Economic Review*, 2015, *105* (2), 504–535.

Patton, Andrew J. and Allan Timmermann, "Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion," *Journal of Monetary Economics*, 2010, *57* (7), 803–820.

Peng, Lin and Wei Xiong, "Investor attention, Overconfidence and Category Learning," *Journal of Financial Economics*, 2006, *80* (3), 563–602.

Pronin, Emily, Daniel Y. Lin, and Lee Ross, "The Bias Blind Spot: Perceptions of Bias in Self Versus Others," *Personality and Social Psychology Bulletin*, 2002, *28* (3), 369–381.

Rabin, Matthew, "Inference by Believers in the Law of Small Numbers," *Quarterly Journal of Economics*, 2002, *117* (3), 775–816.

_ and Joel L. Schrag, "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 1999, *114* (1), 37–82.

Scheinkman, José A. and Wei Xiong, "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 2003, *111* (6), 1183–1220.

Schwartzstein, Joshua, "Selective Attention and Learning," *Journal of the European Economic Association*, 2014, *12* (6), 1423–1452.

Selten, Reinhard, "What is bounded rationality?," in Gerd Gigerenzer and Reinhard Selten, eds., *Bounded Rationality: The Adaptive Toolbox*, MIT Press, 2002, pp. 13–36.

Sethi, Rajiv and Muhamet Yildiz, "Communication with unknown perspectives," *Econometrica*, 2016, *84* (6), 2029–2069.

Simon, Herbert A., *Models of man*, New York, New York: Wiley, 1957.

Simonsohn, Uri, "Lessons from an "Oops" at *Consumer Reports*: Consumers Follow Experts and Ignore Invalid Information," *Journal of Marketing Research*, February 2011, *48*, 1–12.

Snyder, Timothy, "The Next Genocide," *New York Times*, September 13 2015. Available online: https://www.nytimes.com/2015/09/13/opinion/sunday/the-next-genocide.html [Last accessed: May 2017].

Subramanyam, K. R., "Uncertain precision and price reactions to information," *The Accounting Review*, 1996, *71* (2), 207–219.

Suen, Wing, "The self-perpetuation of biased beliefs," *The Economic Journal*, 2004, *114*, 377–396.

Tetlock, Paul, "Information Transmission in Finance," *Annual Review of Financial Economics*, 2014, *6*, 365–384.

Tirole, Jean, "On the Possibility of Speculation under Rational Expectations," *Econometrica*, 1982, *50* (5), 1163–1181.

West, Richard F., Russell J. Meserve, and Keith E. Stanovich, "Cognitive Sophistication Does Not Attenuate the Bias Blind Spot," *Journal of Personality and Social Psychology*, 2012, *103* (3), 506–519.

Xiong, Wei, "Bubbles, Crises, and Heterogeneous Beliefs," in Jean-Pierre Fouque and Joseph A. Langsam, eds., *Handbook on Systemic Risk*, Cambridge University Press, 2013, pp. 663–713.

Appendix

Proof of Proposition 1

1. No average disagreement

   Define $D(\mathbf{s}_n) = P^b(\theta = A|\mathbf{s}_n) - P^u(\theta = A|\mathbf{s}_n)$ as the ex-post realized disagreement after any signal path. The proposition is that $E_0[D(\mathbf{s}_n)] = 0$, where the expectation E_0 is taken by the econometrician over the common prior of states and reliability, which we assume reflects the true ex-ante distribution of (θ, c). Note that the common prior on states and reliability generate a common distribution on the probability of any given signal path.

   Divide the set of all possible signal paths $\{\mathbf{s}_n\}$ into two groups: one group $\{\mathbf{g}_n\}$ where the first signal is a and another group $\{\mathbf{h}_n\}$ where the first signal is b. Because there are two states, there are the same number of signal paths in each group, and the union of these two groups equals $\{\mathbf{s}_n\}$.

   It is clear that taking any signal path $\mathbf{g}_n$ and flipping all the a's to b and b's to a defines a one-to-one and onto mapping F of $\{\mathbf{g}_n\}$ into $\{\mathbf{h}_n\}$. This mapping has two properties:

   (a) $P(\mathbf{g}_n|q, \theta) = P(F(\mathbf{g}_n)|q, -\theta)\ \forall(c, \theta)$, and

   (b) $P^b(\theta = A|F(\mathbf{g}_n)) - P(\theta = A|F(\mathbf{g}_n)) = -\left(P^b(\theta = A|\mathbf{g}_n) - P(\theta = A|\mathbf{g}_n)\right)\ \forall \mathbf{g_n}$,

   where $-\theta$ is the opposite state as θ. The first property says that the probability of the flipped signal sequence is the same as the original signal sequence, once the true state is flipped. The second property can be re-written as $D(F(\mathbf{g}_n)) = -D(\mathbf{g}_n)$ and says that disagreement under the flipped signal path equals the opposite disagreement under the original signal path. Intuitively, these properties follow because, starting from a neutral prior about the state which is independent from credibility, the model is symmetric in A and B irrespective of the true source type.

   More precisely, the first property follows because:

$$P(\mathbf{g}_n|c, A) = q_c^{n_a^g}(1 - q_c)^{n_b^g} = q_c^{n_b^h}(1 - q_c)^{n_a^h} = P(F(\mathbf{g}_n))|c, B),$$

   where n_θ^g, n_θ^h represent the number of times a signal indicating state θ appears in signal sequence $\mathbf{g}_n$ and $F(\mathbf{g}_n)$, respectively, and $n_a^g = n_b^h, n_b^g = n_a^h$ by construction. Similarly, $P(\mathbf{g}_n|c, B) = P(F(\mathbf{g}_n)|c, A)$.

   To prove the second property, note that, for the Bayesian, $P(\theta = A|\mathbf{g}_n) = P(\theta = B|F(\mathbf{g}_n)) = 1 - P(\theta = A|F(\mathbf{g}_n))$. The first equality follows from applying the first property and $\omega_0^A = \omega_0^B = 0.5$ to Equation 4, noting that a Bayesian has constant $\beta_c(F(\mathbf{g}_n))$.

Now consider the pre-screener. Given any sequence $\mathbf{g}_n$, let g_i and h_i be the i-th elements of $\mathbf{g}_n$ and $F(\mathbf{g}_n)$, respectively. Clearly, h_i is the flip of g_i, and $P(g_i|c,\theta) = P(h_i|c,-\theta)$, as both equal q_c if $g_i = \theta$ and $1 - q_c$ if $g_i = -\theta$. Therefore, $\sum_\theta \left(\prod_{i=1}^m P(g_i|c,\theta)\right)\omega_0^\theta = \sum_\theta \left(\prod_{i=1}^m P(h_i|c,\theta)\right)\omega_0^\theta$ for any m due to the summation over both values of θ. Applying this to Equation 5, $\beta_c(\mathbf{g}_n) = \beta_c(F(\mathbf{g}_n))$. From Equation 4, $P^b(\theta = A|\mathbf{g}_n) = P^b(\theta = B|F(\mathbf{g}_n)) = 1 - P^b(\theta = A|F(\mathbf{g}_n))$, and the second property follows.

We claim that $E_0[D(\mathbf{s}_n)|c = \bar{c}] = 0$ for any $\bar{c} \in \{L, H\}$. To be clear, this conditional expectation is taken over the econometrician's information set, but the true c remains unknown to the Bayesian and pre-screener. The proposition then follows due to the tower property of conditional expectations.

Let $\bar{c}$ be given. Observe that:

$$E[D(\mathbf{s}_n)|c = \bar{c}] = \omega_0^A \left(\sum_{\{\mathbf{g}_n\}} P(\mathbf{g}_n|\bar{c}, A)D(\mathbf{g}_n) + \sum_{\{\mathbf{h}_n\}} P(\mathbf{h}_n|\bar{c}, A)D(\mathbf{h}_n) \right) + \omega_0^B \left(\sum_{\{\mathbf{g}_n\}} P(\mathbf{g}_n|\bar{c}, B)D(\mathbf{g}_n) + \sum_{\{\mathbf{h}_n\}} P(\mathbf{h}_n|\bar{c}, B)D(\mathbf{h}_n) \right).$$

The two properties, along the fact that F is one-to-one and onto, imply:

$$\sum_{\{\mathbf{h}_n\}} P(\mathbf{h}_n|\bar{c}, A)D(\mathbf{h}_n) = -\sum_{\{\mathbf{g}_n\}} P(\mathbf{g}_n|\bar{c}, B)D(\mathbf{g}_n)$$
$$\sum_{\{\mathbf{h}_n\}} P(\mathbf{h}_n|\bar{c}, B)D(\mathbf{h}_n) = -\sum_{\{\mathbf{g}_n\}} P(\mathbf{g}_n|\bar{c}, A)D(\mathbf{g}_n).$$

With $\omega_0^A = \omega_0^B$, the claim follows.

2. Expected squared disagreement

   (a) The corollary $Var_0[D(\mathbf{s}_n)] > 0$ follows because $D(F(\mathbf{g}_n))^2 = D(\mathbf{g}_n)^2$.

   (b) Let n_a^s be the number of a signals and n_b^s be the number of b signals in sequence s. Consider any two sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ with identical information content ($n_a^x = n_a^y$ and $n_b^x = n_b^y$). Let β_c^s correspond to sequence $s \in \{\mathbf{x}_n, \mathbf{y}_n\}$ and $c \in \{L, H\}$. Let $\mathbf{s}_n^J = \mathbf{x}_n$ and $\mathbf{s}_n^M = \mathbf{y}_n$. Without loss of generality, let $n_a > n_b$.

   It is sufficient to show that $P^b(\theta = A|\mathbf{x}_n) \neq P^b(\theta = A|\mathbf{y}_n)$ for at least two sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ with identical information content ($n_a^x = n_a^y = n_a$ and

$n_b^x = n_b^y = n_b$). Consider two sequences such that $n_a \geq n_b$, where the first $j = n_a - n_b$ signals are the same and $n-2 \geq j \geq 1$, the two sequences differ in the $j+1$ and $j+2$ signals, and then all subsequent signals are identical (i.e., terms $j+3$ through n). Let $x_{j+1} = a$, $x_{j+2} = b$, $y_{j+1} = b$, and $y_{j+2} = a$. Suppose the first j terms contain k a's and $j-k$ b's, where $k \geq j-k$. As shown in the proof of Proposition 2, $P^b(c = H|\mathbf{x}_n) > P^b(c = H|\mathbf{y}_n)$ when $k > j-k$. As shown in the preceding proof of correlated disagreement between two prescreeners, this implies that $P^b(\theta = A|\mathbf{x}_n) > P^b(\theta = A|\mathbf{y}_n)$. Thus, $P^b(\theta = A|\mathbf{x}_n) \neq P^b(\theta = A|\mathbf{y}_n)$ for at least two sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ with identical information content ($n_a^x = n_a^y = n_a$ and $n_b^x = n_b^y = n_b$).

Proof of Proposition 2

Rearrange Equations 2 and 3 to obtain:

$$P^b(c, \theta|\mathbf{s}_n) = \frac{\beta_c(\mathbf{s}_n)\left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta \omega_0^c}{\sum_c \beta_c(\mathbf{s}_n) \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta \omega_0^c}, \tag{4}$$

where $\beta_c(\mathbf{s}_n)$ is defined as:

$$\begin{aligned}\beta_c(\mathbf{s}_n) &\equiv \left(\textstyle\sum_\theta P(s_1|c,\theta)\omega_0^\theta\right) \times \left(\textstyle\sum_\theta P(s_1|c,\theta)P(s_2|c,\theta)\omega_0^\theta\right) \times \ldots \times \left(\textstyle\sum_\theta P(s_1|c,\theta)P(s_2|c,\theta)\ldots P(s_n|c,\theta)\omega_0^\theta\right) \\ &= \prod_{m=1}^n \left(\sum_\theta \left(\prod_{t=1}^m P(s_t|c,\theta)\right)\omega_0^\theta\right), \end{aligned} \tag{5}$$

and where $\beta_c(\emptyset) \equiv 1$.

Let n_a^s be the number of a signals and n_b^s be the number of b signals in sequence s. Consider any two sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ with identical information content ($n_a^x = n_a^y = n_a$) and $n_b^x = n_b^y = n_b$). Let β_c^s correspond to sequence $s \in \{\mathbf{x}_n, \mathbf{y}_n\}$ and $c \in \{L, H\}$. Without loss of generality, let $n_a > n_b$.

By direct comparison of the posteriors on source reliability, a necessary and sufficient condition for sequence x to generate more trust than sequence y (i.e., $P^b(c = H|\mathbf{x}_n) > P^b(c = H|\mathbf{y}_n)$) is $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y > 0$, where β_c^s corresponds to sequence $s \in \{\mathbf{x}_n, \mathbf{y}_n\}$ and $q \in \{L, H\}$. Consider two sequences such that $n_a \geq n_b$, where the first $j = n_a - n_b$ signals are the same and $n-2 \geq j \geq 1$, the two sequences differ in the $j+1$ and $j+2$ signals, and then all subsequent signals are identical (i.e., terms $j+3$ through n). Let $x_{j+1} = a$, $x_{j+2} = b$, $y_{j+1} = b$, and $y_{j+2} = a$. (For example, sequence 1 could be *aababaa* and sequence 2 could be *aabbaaa* - here $j = 3$, $n_a = 2$, $n_b = 1$.) Then $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y > 0$ whenever $n_a > n_b$ and $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y = 0$ whenever $n_a = n_b$. To see this, note that, given the general expression for β_c^s, all of the terms are identical for β_c^x and β_c^y except term $j+1$. Suppose the first j terms contain k a's and $j-k$ b's, where $k \geq j-k$. This implies that when $\omega_0^\theta = 1/2$, then

$\beta_H^x \beta_L^y - \beta_L^x \beta_H^y \geq 0$ if

$$\left(q_H^{k+1}(1-q_H)^{j-k} + (1-q_H)^{k+1} q_H^{j-k}\right)\left(q_L^{k}(1-q_L)^{j-k+1} + (1-q_L)^{k} q_L^{j-k+1}\right) -$$
$$\left(q_L^{k+1}(1-q_L)^{j-k} + (1-q_L)^{k+1} q_L^{j-k}\right)\left(q_H^{k}(1-q_H)^{j-k+1} + (1-q_H)^{k} q_H^{j-k+1}\right) \geq 0$$
$$(q_H - q_L)\left((q_H q_L)^{2k-j} - ((1-q_H)(1-q_L))^{2k-j}\right) + (q_H + q_L - 1)\left((q_H(1-q_L))^{2k-j} - (q_L(1-q_H))^{2k-j}\right) \geq 0.$$

We can verify that both terms are positive when $k > j-k$ and zero when $k = j-k$. Thus, $P^b(c = H|\mathbf{x}_n) > P^b(c = H|\mathbf{y}_n)$ when $k > j-k$. Using this result, we can iteratively apply it to order sequences of fixed composition in decreasing trust by starting with the sequence with the least reversals (all a's followed by all b's), and iteratively switching the first b and last a to generate sequences where the first b moves forward. E.g., *aaaabb* generates more trust than *aaabab*, which generates more trust than *aabaab* which generates more trust than *abaaab*. Then, *aaabba* generates more trust than *aababa* than *abaaba*, where *aaabab* generates more trust than *aaabba* and *abaaab* generates more trust than *abaaba*. We can keep doing this (and applying the result that $P^b(c = H|\mathbf{x}_n) > P^b(c = H|\mathbf{y}_n)$ when $k > j-k$) to establish that *aaaabb* generates the most trust and *ababaa* generates the least trust.

Proof of Proposition 3

1. Positive first impressions

   Suppose the agent observes $n_a \geq 1$ consecutive a signals, followed by m pairs of (b, a) signals: $\mathbf{s}_n = (a, a, a, \ldots, b, a, b, a)$. This sequence generates[12]:

$$\beta_c(\mathbf{s}_n) = \left(\tfrac{1}{2}\right)^{n_a+m}[q_c(1-q_c)]^{m(m+1)}\left([q_c^{n_a-1} + (1-q_c)^{n_a-1}][q_c^{n_a} + (1-q_c)^{n_a}]\right)^m \left(\prod_{i=1}^{n_a}(q_c^i + (1-q_c)^i)\right) \tag{6}$$

$$\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} = \left(\tfrac{1}{2}\right)^{n_a+m}[q_c(1-q_c)]^{m(m+1)}\left([q_c^{n_a-1} + (1-q_c)^{n_a-1}][q_c^{n_a} + (1-q_c)^{n_a}]\right)^{m-1} \left(\prod_{i=1}^{n_a}(q_c^i(1-q_c)^i)\right)$$
$$(mq_c(1-q_c)\left((n_a-1)(q_c^{n_a-2} - (1-q_c)^{n_a-2})(q_c^{n_a} + (1-q_c)^{n_a}) + n_a(q_c^{n_a-1} + (1-q_c)^{n_a-1})(q_c^{n_a-1} - (1-q_c)^{n_a-1}\right.$$
$$\left.+(q_c^{n_a-1} + (1-q_c)^{n_a-1})(q_c^{n_a} + (1-q_c)^{n_a})\left(m(m+1)(1-2q_c) + q_c(1-q_c)\sum_{i=1}^{n_a}\frac{i(q_c^{i-1} - (1-q_c)^{i-1})}{q_c^i + (1-q_c)^i}\right)\right). \tag{7}$$

[12]We use the property that a product of multiple factors is given by $\frac{d}{dx}\left(\prod_{i=1}^k f_i(x)\right) = \left(\prod_{i=1}^k f_i(x)\right)\left(\sum_{i=1}^k \frac{f_i'(x)}{f_i(x)}\right)$.

Note that Equation (7) can be re-written as

$$\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c} = \left(\tfrac{1}{2}\right)^{n_a+m} m[q_c(1-q_c)]^{m(m+1)} \left([q_c^{n_a-1}+(1-q_c)^{n_a-1}][q_c^{n_a}+(1-q_c)^{n_a}]\right)^{m-1} \left(\prod_{i=1}^{n_a}(q_c^i(1-q_c)^i)\right)(Z),$$

where $\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c}$ is negative whenever Z is negative and $q_c \in (\frac{1}{2}, 1)$, and

$$Z = q_c(1-q_c)\left((n_a-1)(q_c^{n_a-2}-(1-q_c)^{n_a-2})(q_c^{n_a}+(1-q_c)^{n_a}) + n_a(q_c^{n_a-1}+(1-q_c)^{n_a-1})(q_c^{n_a-1}-(1-q_c)^{n_a-1})\right)$$
$$+ (q_c^{n_a-1}+(1-q_c)^{n_a-1})(q_c^{n_a}+(1-q_c)^{n_a})\left((m+1)(1-2q_c) + \left(\frac{1}{m}\right) q_c(1-q_c)\sum_{i=1}^{n_a}\frac{i(q_c^{i-1}-(1-q_c)^{i-1})}{q_c^i+(1-q_c)^i}\right).$$

For given n_a, Z is more than linearly decreasing in m. Thus, there exists $\hat{m}$, defined by $Z(\hat{m}) = 0$, such that $\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c} < 0$ for all $q_c \in (\frac{1}{2}, 1)$ when $m > \hat{m}$. Thus for any given n_a, there exists $\hat{m}$ such that when $m > \hat{m}$, the pre-screener under-trusts and is pessimistic about the most likely state for any (q_L, q_H).

2. Negative first impressions

Suppose the agent observes $n_b \geq 1$ pairs of (a, b) signals, followed by $m \geq 1$ consecutive a signals, where $m \geq 1$: $\mathbf{s}_n = (a, b, a, b, \ldots, a, a, a)$. This sequence generates:

$$\beta_c(\mathbf{s}_n) = \left(\frac{1}{2}\right)^{n_b} (q_c(1-q_c))^{n_b^2} \left(\prod_{i=1}^{m}\frac{1}{2}\left(q_c^{i+n_b}(1-q_c)^{n_b} + q_c^{n_b}(1-q_c)^{i+n_b}\right)\right) \tag{8}$$

$$= \left(\frac{1}{2}\right)^{n_b+m} (q_c(1-q_c))^{n_b(n_b+m)} \left(\prod_{i=1}^{m}(q_c^i+(1-q_c)^i)\right). \tag{9}$$

$$\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c} = \left(\frac{1}{2}\right)^{n_b+m} (q_c(1-q_c))^{n_b(n_b+m)-1} \left(\prod_{i=1}^{m}(q_c^i+(1-q_c)^i)\right)$$
$$\left(n_b(n_b+m)(1-2q_c) + q_c(1-q_c)\sum_{i=1}^{m}\left(\frac{i(q_c^{i-1}-(1-q_c)^{i-1})}{q_c^i+(1-q_c)^i}\right)\right) \tag{10}$$

Evaluating Equation (10) when $m = 1$ (i.e., $n_a = n_b+1$ where $n_b \geq 1$), $\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c}|_{m=1} < 0$. Thus, by Proposition 4, the pre-screener under-trusts and is pessimistic about the mostly likely state, A, when she observes a sequence $\mathbf{s}_n = (a, b, a, b, \ldots, a, b, a)$ where $n_a = n_b + 1$. Further, evaluating Equation (10) when $m = 2$ (i.e., $n_a = n_b + 2$ where $n_b \geq 1$), $\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c}|_{m=1} < 0$. Thus, the pre-screener still under-trusts and is pessimistic about the most likely state, A, when $\mathbf{s}_n = (a, b, a, b, \ldots, a, a)$ where $n_a = n_b + 2$ for all $n_b \geq 1$. Further, evaluating Equation 10 when $m = 3$ (i.e., $n_a = n_b + 3$ where $n_b \geq 1$), $\frac{\partial\beta_c(\mathbf{s}_n)}{\partial q_c}|_{m=3} \leq 0$ with equality at $q_c = \frac{1}{2}$ only if $n_b = 1$. Since the third term of Equation

(10) is decreasing in n_b for all $q_c \in (\frac{1}{2}, 1]$, then $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c}|_{m=3} < 0$ for all $n_b > 1$. Thus, the pre-screener still under-trusts and is pessimistic about the most likely state, A, when $\mathbf{s}_n = (a, b, a, b, \ldots, a, a)$ where $n_a = n_b + 3$ for all $n_b \geq 1$. Therefore, there exists some $m^* > 3$ such that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} < 0$ for all $m < m^*$, which implies that the pre-screener will under-trust for $m < m^*$. Moreover, since the third term of Equation (10) is decreasing in n_b for all $q_c \in (\frac{1}{2}, 1]$, then m^* is increasing in n_b.

From Equations (9) and (10), we can see that $\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) = 0$ when $q_c \in \{\frac{1}{2}, 1\}$, $\beta_c(\mathbf{s}_n) > 0$ when $q_c = \frac{1}{2}$, and $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1$. Since $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1$, $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c}|_{q_c=1} = 0$, and $\beta_c(\mathbf{s}_n) \geq 0$ for any $q_c \in [0, 1]$, then there exists some threshold $\hat{q} < 1$ such that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} < 0$ and $\beta_c(\mathbf{s}_n) < \beta_c(\mathbf{s}_n)\Big|_{q_c=\frac{1}{2}}$ for all $q_c > \hat{q}$. Therefore, $\beta_L(\mathbf{s}_n) > \beta_H(\mathbf{s}_n)$ so the pre-screener under-trusts and is pessimistic about the most likely state if $\hat{q} \leq q_L < q_H$.

Moreover, using the fact that $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1/2$, then

$$\begin{aligned}
\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}} &= \beta_c(\mathbf{s}_n)\left(-8n_b(n_b+m) + \sum_{i=1}^{m} 4i(i-1)\right) \\
&= \beta_c(\mathbf{s}_n)\left(-8n_b(n_b+m) + \frac{4}{3}m(m-1)(m+1)\right). \qquad (11)
\end{aligned}$$

Thus there exists some threshold $\frac{1}{2} < \check{q} < 1$ such that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} > 0$ for all $q_c < \check{q}$ when $\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}} > 0$. Note that $\check{q} > \frac{1}{2}$ for any given n_b if m is sufficiently large that $\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}} > 0$. Thus the pre-screener also under-trusts and is pessimistic about the most likely state if $q_L \leq \check{q}$ and $q_H > \hat{q}$ where $\check{q} \geq \frac{1}{2}$. Note that we have already shown directly that the pre-screener under-trusts and is pessimistic for $m = 1, 2, 3$ regardless of q_L, q_H, and n_b.

Proof of Proposition 4

1. **Lemma 1** *For all $\omega_0^\theta \in (0, 1)$ and $\omega_0^c \in (0, 1)$, $\kappa_H(\mathbf{s}_n) < w_0^H$ if and only if $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$. Likewise, $\kappa_H(\mathbf{s}_n) > w_0^H$ if and only if $\beta_H(\mathbf{s}_n) > \beta_L(\mathbf{s}_n)$. $\kappa_H(\mathbf{s}_n) = w_0^H$ if and only if $\beta_H(\mathbf{s}_n) = \beta_L(\mathbf{s}_n)$.*

   Proof. For any given sequence of signals $\mathbf{s}_n = (s_1, s_2, \ldots, s_n)$, $\kappa_c(\mathbf{s}_n)$ can be re-written as

$$\begin{aligned}\kappa_c(\mathbf{s}_n) &= \frac{\beta_c(\mathbf{s}^{n-1})\omega_0^c \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta}{\sum_c \beta_c(\mathbf{s}^{n-1})\sum_\theta \left(\prod_{t=1}^n P(s_n|c,\theta)\right)\omega_0^\theta \omega_0^c} \\ &= \frac{\left(\prod_{m=1}^{n-1}\left(\sum_\theta \left(\prod_{t=1}^m P(s_t|c,\theta)\right)\omega_0^\theta\right)\right)\omega_0^c \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta}{\sum_c \left(\prod_{m=1}^{n-1}\left(\sum_\theta \left(\prod_{t=1}^m P(s_t|c,\theta)\right)\omega_0^\theta\right)\right)\sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta \omega_0^c} \\ &= \frac{\beta_c(\mathbf{s}_n)\omega_0^c}{\sum_c \beta_c(\mathbf{s}_n)\omega_0^c}.\end{aligned}$$

Thus, the statement is shown for $\mathbf{s}_n = (s_1, s_2, \ldots, s_n)$.[13] ■

From Equation (4), the pre-screener's posterior that the source is high reliability is lower than the Bayesian's if and only if $\kappa_H(\mathbf{s}_n) < w_0^H$. Lemma 1 shows that this is only the case if and only if $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$. Thus, $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$ if and only if $P^b(c = H|\mathbf{s}_n) < P^u(c = H|\mathbf{s}_n)$.

Consider $P^b(\theta = A|\mathbf{s}_n) < P^u(\theta = A|\mathbf{s}_n)$:

$$\begin{gathered}P^b(\theta = A|\mathbf{s}_n) < P^u(\theta = A|\mathbf{s}_n) \\ \frac{\omega_0^A \sum_c \beta_c(\mathbf{s}_n)\left(\prod_{t=1}^n P(s_t|c,A)\right)\omega_0^c}{\sum_c \beta_c(\mathbf{s}_n)\sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta\omega_0^c} < \frac{\omega_0^A \sum_c \left(\prod_{t=1}^n P(s_t|c,A)\right)\omega_0^c}{\sum_c \sum_\theta \left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^\theta\omega_0^c},\end{gathered}$$

which is true if and only if

$$\begin{aligned}0 &< \omega_0^A(1-\omega_0^A)\omega_0^H(1-\omega_0^H)(\beta_L(\mathbf{s}_n)-\beta_H(\mathbf{s}_n))\left(\left(\prod_{t=1}^n P(s_t|H,A)\right)\left(\prod_{t=1}^n P(s_t|L,B)\right) - \left(\prod_{t=1}^n P(s_t|H,B)\right)\left(\prod_{t=1}^n P(s_t|L,A)\right)\right) \\ 0 &< \omega_0^A(1-\omega_0^A)\omega_0^H(1-\omega_0^H)(\beta_L(\mathbf{s}_n)-\beta_H(\mathbf{s}_n))\left(q_H^{na}(1-q_H)^{nb}q_L^{nb}(1-q_L)^{na} - q_H^{nb}(1-q_H)^{na}q_L^{na}(1-q_L)_b^n\right) \\ 0 &< \omega_0^A(1-\omega_0^A)\omega_0^H(1-\omega_0^H)(\beta_L(\mathbf{s}_n)-\beta_H(\mathbf{s}_n))\left(q_H(1-q_H)q_L(1-q_L)\right)^{nb}\left((q_H(1-q_L))^{na-nb} - ((1-q_H)q_L)^{na-nb}\right),\end{aligned}$$

which is true when $n_a > n_b$ since $q_H > q_L$. Clearly, $P^u(A|\mathbf{s}_n) > \omega_0^A$ only if $n_a > n_b$, so A is (objectively) more likely state than originally believed. Note that if $n_a = n_b$, then $P^b(A|\mathbf{s}_n) = Pr^u(A|\mathbf{s}_n)$ regardless of the pre-screener's beliefs on the source's reliability. Thus, for any $n_a > n_b$ set of signals and for all $\omega_0^\theta \in (0,1)$, under-trust in source reliability implies pessimism in beliefs about the more likely state: If $P^b(H|\mathbf{s}_n) < P^u(H|\mathbf{s}_n)$, then $P^b(A|\mathbf{s}_n) < P^u(A|\mathbf{s}_n)$. Likewise, $P^b(A|\mathbf{s}_n) < P^u(A|\mathbf{s}_n)$ if and only if $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$ when $n_a > n_b$, which implies that $P^b(H|\mathbf{s}_n) < P^u(H|\mathbf{s}_n)$. Reversing the inequalities yields that overtrust in source reliability implies optimism in beliefs about the more likely state, and vice versa.

[13] If the signals are observed simultaneously (e.g., in period 1), then the above argument applies analogously, where $\beta_c(\mathbf{s}^{t-1}) = \beta_c(\emptyset) = 1$ instead. Thus, $\kappa_H(\mathbf{s}_n) < w_0^H$ implies $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$ and vice versa. Likewise when the unreliability reverses or when the reliability holds.

2. Let n_a^s be the number of a signals and n_b^s be the number of b signals in sequence s. Consider any two sequences $\mathbf{x}_n$ and $\mathbf{y}_n$ with identical information content ($n_a^x = n_a^y$ and $n_b^x = n_b^y$). Let β_c^s correspond to sequence $s \in \{\mathbf{x}_n, \mathbf{y}_n\}$ and $c \in \{L, H\}$. Let $\mathbf{s}_n^J = \mathbf{x}_n$ and $\mathbf{s}_n^M = \mathbf{y}_n$. Without loss of generality, let $n_a > n_b$.

   By direct comparison of the posteriors on source credibility, a necessary and sufficient condition for $P^b(H|\mathbf{x}_n) > P^b(H|\mathbf{y}_n)$ is $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y > 0$. By direct comparison of the posteriors on the most likely state (which is A because $n_a > n_b$), a necessary and sufficient condition for $P^b(A|\mathbf{x}_n) > P^b(A|\mathbf{y}_n)$ is $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y > 0$. Since the same condition $\beta_H^x \beta_L^y - \beta_L^x \beta_H^y > 0$ is required for both $P^b(H|\mathbf{x}_n) > P^b(H|\mathbf{y}_n)$ and $P^b(A|\mathbf{x}_n) > P^b(A|\mathbf{y}_n)$, then disagreement between pre-screeners is correlated. That is, $P^b(H|\mathbf{x}_n) > P^b(H|\mathbf{y}_n)$ if and only if $P^b(A|\mathbf{x}_n) > P^b(A|\mathbf{y}_n)$. Clearly reversing all the inequalities applies as well.

Proof of Proposition 5

1. Let $\mathbf{s}_n$ be a sequence of n observed signals with n_a a's and n_b b's, let s_{n+1} be the $(n+1)$th observed signal, and let ω_n^b equal the pre-screener's joint posterior after the sequence $\mathbf{s}_n$.

   First, note that each joint belief on the state and credibility for the prior ω_n^b, denoted $\omega_n^{c\theta}$, is given by

$$\omega_n^{c\theta} \equiv P^b(c, \theta|\{\mathbf{s}_n\}) = \frac{\left(\prod_{t=1}^n P(s_t|c,\theta)\right) \omega_0^\theta \omega_0^c \beta_c(\mathbf{s}_n)}{\sum_\theta \sum_c \left(\prod_{t=1}^n P(s_t|c,\theta)\right) \omega_0^\theta \omega_0^c \beta_c(\mathbf{s}_n)}, \tag{12}$$

   where

$$\beta_c(\mathbf{s}_n) = \prod_{m=1}^n \left(\sum_\theta \left(\prod_{t=1}^m P(s_t|c,\theta) \right) \omega_0^\theta \right). \tag{13}$$

   Thus, the Bayesian's posterior belief given the biased prior is

$$P^u(\theta = A|\text{prior} = \omega_n^b, \{s_{n+1}\}) = \frac{\omega_0^A \sum_c \left(\prod_{t=1}^{n+1} P(s_t|q,A)\right) \omega_0^c \beta_c(\mathbf{s}_n)}{\sum_\theta \omega_0^\theta \sum_c P(s_{n+1}|c,\theta) \left(\prod_{t=1}^n P(s_t|c,\theta)\right) \omega_0^c \beta_c(\mathbf{s}_n)}.$$

   In contrast, the pre-screener's posterior belief after observing $\{\mathbf{s}_n, s_{n+1}\}$ is

$$P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) = \frac{\omega_0^A \sum_c \left(\prod_{t=1}^{n+1} P(s_t|q,A)\right) \omega_0^c \beta_c(\{\mathbf{s}_n, s_{n+1}\})}{\sum_\theta \omega_0^\theta \sum_c \left(\prod_{t=1}^{n+1} P(s_t|c,\theta)\right) \omega_0^c \beta_c(\{\mathbf{s}_n, s_{n+1}\})},$$

where

$$\beta_c(\{\mathbf{s}_n, s_{n+1}\}) = \prod_{m=1}^{n+1}\left(\sum_\theta\left(\prod_{t=1}^{m} P(s_t|c,\theta)\right)\omega_0^\theta\right) = \beta_c(\mathbf{s}_n)\left(\sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|c,\theta)\right)\omega_0^\theta\right). \tag{14}$$

Substituting all of the preceding information into $P^b(\omega = A|\{\mathbf{s}_n, s_{n+1}\}) > P^u(\omega = A|prior = \omega_n^b, \{s_{n+1}\})$, the inequality is only satisfied if

$$\omega_0^A(1-\omega_0^A)\omega_0^H(1-\omega_0^H)\beta_L(\mathbf{s}_n)\beta_H(\mathbf{s}_n)\underbrace{\left(\left(\sum_\theta\left(\prod_{t=1}^{n+1}P(s_t|H,\theta)\right)\omega_0^\theta\right) - \left(\sum_\theta\left(\prod_{t=1}^{n+1}P(s_t|L,\theta)\right)\omega_0^\theta\right)\right)}_{X}$$
$$\underbrace{\left(\left(\prod_{t=1}^{n+1}P(s_t|H,A)\right)\left(\prod_{t=1}^{n+1}P(s_t|L,B)\right) - \left(\prod_{t=1}^{n+1}P(s_t|H,B)\right)\left(\prod_{t=1}^{n+1}P(s_t|L,A)\right)\right)}_{Y} > 0, \tag{15}$$

Without loss of generality, suppose that $n_a \geq n_b$.

If $s_{n+1} = a$, then $\{\mathbf{s}_n, s_{n+1}\}$ has $n_a + 1$ a's and n_b b's. Then the term Y is given by

$$\begin{aligned}&q_H^{n_a+1}(1-q_H)^{n_b}(1-q_L)^{n_a+1}(q_L)^{n_b} - (1-q_H)^{n_a+1}(q_H)^{n_b}(q_L)^{n_a+1}(1-q_L)^{n_b}\\ &= [q_Hq_L(1-q_H)(1-q_L)]^{n_b}[(q_H(1-q_L))^{n_a-n_b+1} - (q_L(1-q_H))^{n_a-n_b+1}]\end{aligned}$$

Thus if $s_{n+1} = a$, then

$$Y(s_{n+1} = a)\begin{cases} > 0 & \text{if } n_a + 1 > n_b\\ = 0 & \text{if } n_a + 1 = n_b\\ < 0 & \text{if } n_a + 1 < n_b.\end{cases}$$

.

If $s_{n+1} = b$, then $\{\mathbf{s}_n, s_{n+1}\}$ has n_a a's and $n_b + 1$ b's. Then the term Y is given by

$$\begin{aligned}&q_H^{n_a}(1-q_H)^{n_b+1}(1-q_L)^{n_a}(q_L)^{n_b+1} - (1-q_H)^{n_a}(q_H)^{n_b+1}(q_L)^{n_a}(1-q_L)^{n_b+1}\\ &= [q_Hq_L(1-q_H)(1-q_L)]^{n_b}[(q_H(1-q_L))^{n_a-n_b}q_L(1-q_H) - ((1-q_H)q_L)^{n_a-n_b}q_H(1-q_L)]\end{aligned}$$

Thus if $s_{n+1} = b$, then

$$Y(s_{n+1} = b)\begin{cases} > 0 & \text{if } n_a > n_b + 1\\ = 0 & \text{if } n_a = n_b + 1\\ < 0 & \text{if } n_a < n_b + 1.\end{cases}$$

Thus, Y is positive if $\{\mathbf{s}_n, s_{n+1}\}$ has more a's than b's, Y is negative if $\{\mathbf{s}_n, s_{n+1}\}$ has more b's than a's, and Y is zero if $\{\mathbf{s}_n, s_{n+1}\}$ has an equal number of a's and b's.

Moreover, note that $\kappa_c(\mathbf{s}_n) \equiv \frac{\beta_c(\mathbf{s}_n)\omega_0^c}{\sum_c \beta_c(\mathbf{s}_n)\omega_0^c}$. Then Equation (14) implies that $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$ if and only if
$\omega_0^H(1-\omega_0^H)\beta_H(\mathbf{s}_n)\beta_L(\mathbf{s}_n)\left(\left(\sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|H,\theta)\right)\omega_0^\theta\right) - \left(\sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|L,\theta)\right)\omega_0^\theta\right)\right) > 0$, which is the requirement that $X > 0$.

In other words,

$$X \begin{cases} > 0 & \text{if and only if } \kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n) \\ = 0 & \text{if and only if } \kappa_H(\{\mathbf{s}_n, s_{n+1}\}) = \kappa_H(\mathbf{s}_n) \\ < 0 & \text{if and only if } \kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n). \end{cases}$$

From above, we can see that the sign of Equation (15) depends on the sign of XY. Putting everything together, then

$P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) = P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if either (1) $\{\mathbf{s}_n, s_{n+1}\}$ has an equal number of a's and b's or (2) $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) = \kappa_H(\mathbf{s}_n)$,

$P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) > P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if (3) $\{\mathbf{s}_n, s_{n+1}\}$ has more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$ or (4) $\{\mathbf{s}_n, s_{n+1}\}$ has more b's than a's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n)$,

$P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) < P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if (5) $\{\mathbf{s}_n, s_{n+1}\}$ has more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n)$ or (6) $\{\mathbf{s}_n, s_{n+1}\}$ has more b's than a's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$.

Note that the statement $P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) > P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if $\{\mathbf{s}_n, s_{n+1}\}$ has more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$ is equivalent to the statement $P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) < P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if $\{\mathbf{s}_n, s_{n+1}\}$ has more b's than a's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n)$. Likewise, the statement $P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) > P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if $\{\mathbf{s}_n, s_{n+1}\}$ has more b's than a's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n)$ is equivalent to the statement $P^b(\theta = A|\{\mathbf{s}_n, s_{n+1}\}) < P^u(\theta = A|prior = \omega_n^b, \{s_{n+1}\})$ if $\{\mathbf{s}_n, s_{n+1}\}$ has more a's than b's and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n)$. Therefore, we can state the proposition assuming that the number of a's be greater than or equal to the number of b's in $\{\mathbf{s}_n, s_{n+1}\}$ without loss of generality.

Moreover, note that $Pr^u(H|\{\mathbf{s}_n, s_{n+1}\}) = \frac{\omega_0^H \sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|c,\theta)\right)\omega_0^\theta}{\sum_c \omega_0^c \sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|c,\theta)\right)\omega_0^\theta}$. From this definition, we know that $Pr^u(H|\{\mathbf{s}_n, s_{n+1}\}) > \omega_0^H$ if and only if
$\left(\left(\sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|H,\theta)\right)\omega_0^\theta\right) - \left(\sum_\theta\left(\prod_{t=1}^{n+1} P(s_t|L,\theta)\right)\omega_0^\theta\right)\right) > 0$, which is the require-

ment that $X > 0$. Thus,

$$\begin{aligned}
\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\mathbf{s}_n) &\text{ if and only if } Pr^u(q = H|\{\mathbf{s}_n, s_{n+1}\}) > \omega_0^H \\
\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) = \kappa_H(\mathbf{s}_n) &\text{ if and only if } Pr^u(q = H|\{\mathbf{s}_n, s_{n+1}\}) = \omega_0^H \\
\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\mathbf{s}_n) &\text{ if and only if } \; Pr^u(q = H|\{\mathbf{s}_n, s_{n+1}\}) < \omega_0^H.
\end{aligned}$$

For a more detailed explanation of the intuition for these results, please refer to the Internet Appendix.

2. First, note that $P^b[c, \theta|\{\mathbf{s}_n, s_{n+1}\}]$ is equal to

$$P^b[c, \theta|\{\mathbf{s}_n, s_{n+1}\}] = \frac{\beta_c(\{\mathbf{s}_n, s_{n+1}\}) \left(\prod_{t=1}^{n+1} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c}{\sum_c \beta_c(\{\mathbf{s}_n, s_{n+1}\}) \sum_\theta \left(\prod_{t=1}^{n+1} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c} \tag{16}$$

where $\beta_c(\{\mathbf{s}_n, s_{n+1}\})$ is described by Equation (14). Second, applying the generalized pre-screening described in the Internet Appendix, $P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}]$ is equal to

$$P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}] = \frac{\beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) P(s_{t+1}|c, \theta) \omega_n^{c\theta}}{\sum_c \sum_\theta \beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) P(s_{t+1}|c, \theta) \omega_n^{c\theta}}, \tag{17}$$

where $\beta_{c\theta}(s_{n+1}) = \sum_\theta P(s_{n+1}|c, \theta)\omega_n^{c\theta}$ and $\omega_n^{c\theta}$ is described by Equation (12) and $\beta_c(\mathbf{s}_n)$ is described by Equation (13). Substituting this into $P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}]$ yields:

$$\begin{aligned}
P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}] &= \frac{\beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) P(s_{t+1}|c, \theta) \beta_c(\mathbf{s}_n) \left(\prod_{t=1}^{n} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c}{\sum_c \sum_\theta \beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) P(s_{t+1}|c, \theta) \beta_c(\mathbf{s}_n) \left(\prod_{t=1}^{n} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c} \\
&= \frac{\beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) \beta_c(\mathbf{s}_n) \left(\prod_{t=1}^{n+1} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c}{\sum_c \sum_\theta \beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) \beta_c(\mathbf{s}_n) \left(\prod_{t=1}^{n+1} P(s_t|c, \theta)\right) \omega_0^\theta \omega_0^c},
\end{aligned} \tag{18}$$

where

$$\beta_{c\theta}(s_{n+1}) \left(\frac{1}{\sum_\theta \omega_n^{c\theta}}\right) = \beta_c(\mathbf{s}_n) \left(\sum_\theta \left(\prod_{t=1}^{n+1} P(s_t|c, \theta)\right) \omega_0^\theta\right) \left(\frac{\omega_0^c}{\sum_\theta \omega_n^{c\theta}}\right).$$

Equation (14) implies that Equation (18) equals Equation (16) if and only if $\omega_0^c = \sum_\theta \omega_n^{c\theta}$. Since $\omega_n^{c\theta} \equiv P^b(c, \theta|\{\mathbf{s}_n\})$, then $P^b[c, \theta|\{\mathbf{s}_n, s_{n+1}\}] \neq P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}]$ if $P^b[c|\mathbf{s}_n] \neq \omega_0^c$ and $P^b[c, \theta|\{\mathbf{s}_n, s_{n+1}\}] = P^b[c, \theta|\text{prior} = \omega_n^b, \{s_{n+1}\}]$ if $P^b[c|\mathbf{s}_n] = \omega_0^c$.

Proof of Proposition 6

Lemma 2 *Let* $(\omega^A_{-\tau}, \omega^H_{-\tau}) = (\hat{\theta}, \hat{\omega})$ *for any* $\hat{\theta} \in (0,1)$ *and* $\hat{\omega} \in (0,1)$. *After observing* $n_{a,t} = k > 1$ *and* $n_{b,t} = k$ *(in any order), the disagreement about the state is zero.*

Proof. Follows directly from Proof of Proposition 4 by setting $n_{a,t} = n_{b,t}$. ■

Lemma 3 *If group* Y *offers the price and group* X *takes the price, then in any period* t *the price of the asset is* $p_t = E^X_t(R)$. *If group* Y^{EB} *offers the price and group* X^{EB} *takes the price, then in any period* t *the price of the asset is* $p_t = E^{X^{EB}}_t(R)$.

Proof. We can determine the price in any period $t \in \{1, 2, \dots T\}$ by backwards induction, as in the proof of Lemma 2 of Harris and Raviv (1993). In period T, Y buys from X if and only if $E^Y_T(R) \geq p_T$ where p_T is the time-T price. Since Y has all the bargaining power, she offers to buy at $p_T = E^X_T(R)$, which X accepts (and receives zero expected utility). Conversely, Y sells to X if and only if $E^Y_T(R) < p_T$, and sells at $p_T = E^X_T(R)$. In period $T-1$, consider what price Y would offer if she wants to trade (buy or sell). X is willing to sell if $p_{T-1} \geq E^X_{T-1}[p_T]$ or $p_{T-1} \geq E^X_{T-1}(R)$. By LIE *because* X *thinks she is Bayesian*, these two conditions are identical: $E^X_{T-1}[p_T] = E^X_{T-1}[E^X_T(R)] = E^X_{T-1}(R)$. If Y wants to buy, she can force X down to their reservation price (lowest willingness to sell) since Y has all the bargaining power: $p_{T-1} = E^X_{T-1}(R)$. Conversely, X is willing to buy if $p_{T-1} \leq E^X_{T-1}(R)$. If Y wants to sell, she sells at the highest willingness to pay, and $p_{T-1} = E^X_{T-1}(R)$ again. And so on for all preceding periods. Thus, in any given period t, we have $p_t = E^X_t(R)$.

Exactly the same argument applies for trade between X^{EB} and Y^{EB}. Thus, in any given period t, we have $p_t = E^{X^{EB}}_t(R)$. ■

Lemma 4 *Suppose that pre-screener 1 observes signal path* $\mathbf{s}^X_{pre}$ *in the pre-period. In contrast, suppose pre-screener 2 observes signal path* $\mathbf{s}^Y_{pre}$ *in the pre-period. Then both pre-screeners observes public signal path* $Z = \mathbf{s}_t$ *in the trading period. Let* $n^X_{a,pre} = n^Y_{a,pre}$ *and* $n^X_{b,pre} = n^Y_{b,pre}$ *where* $n^j_{a,pre} + n^j_{b,pre} = n^j_{pre}$ *for* $j \in \{X, Y\}$, *so signal paths* $\mathbf{s}^X_{pre}$ *and* $\mathbf{s}^Y_{pre}$ *contain the same information content. Let* $c \in \{v, w\}$.

If $P^b(v|\mathbf{s}^X_{pre}) > P^b(v|\mathbf{s}^Y_{pre})$, *then* $P^b(v|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(v|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$.

Proof. From the proof of Proposition 2, we already know that the necessary and sufficient condition for $P^b(v|\mathbf{s}^X_{pre}) > P^b(v|\mathbf{s}^Y_{pre})$ is that $\beta^X_v \beta^Y_w - \beta^X_w \beta^Y_v > 0$.[14] Suppose that this holds.

Analogously, we can only have $P^b(v|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(v|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ if $\beta^{\{X,Z\}}_v \beta^{\{Y,Z\}}_w - \beta^{\{X,Z\}}_w \beta^{\{Y,Z\}}_v > 0$. Note that

$$\beta^{\{X,Z\}}_c = \beta^X_c b_c$$
$$\beta^{\{Y,Z\}}_c = \beta^Y_c b_c,$$

[14] Note that this property does not require any restrictions on the c type-space or on the number a's and b's in the pre-period signals, only that X and Y have the same information content.

where

$$b_c = \prod_{m=1}^{t} \left(\sum_{\theta} \left(\prod_{i=-\tau+1}^{0} P(s_i|c,\theta) \right) \left(\prod_{i=1}^{m} P(s_i|c,\theta) \right) \omega_{-\tau}^{\theta} \right)$$

because $\prod_{i=-\tau+1}^{0} P(s_i|c,\theta)$ is the same for signal paths $\mathbf{s}_{pre}^{X}$ and $\mathbf{s}_{pre}^{Y}$ since they have the same information content. Thus,

$$\begin{aligned}\beta_v^{\{X,Z\}}\beta_w^{\{Y,Z\}} - \beta_w^{\{X,Z\}}\beta_v^{\{Y,Z\}} &= \beta_v^X b_v \beta_w^Y b_w - \beta_w^X b_w \beta_v^Y b_v \\ &= b_v b_w (\beta_v^X \beta_w^Y - \beta_w^X \beta_v^Y) > 0,\end{aligned}$$

since $\beta_v^X \beta_w^Y - \beta_w^X \beta_v^Y > 0$. ■

Lemma 5 *Suppose a pre-screener or Bayesian with prior $\omega_0 = \omega_0^{\theta}\omega_0^{c}$ observes $n_{a,t} = n_{b,t}$ signals. Then $P^b(\theta, c|\mathbf{s}_t) = P^b(\theta|\mathbf{s}_t)P^b(c|\mathbf{s}_t)$ and $P^u(\theta, c|\mathbf{s}_t) = P^u(\theta|\mathbf{s}_t)P^b(c|\mathbf{s}_t)\ \ \forall\ \theta, c$.*

Proof. This is easily verified by direct calculation. ■

First, suppose groups X and Y observe $\mathbf{s}_{pre}^{X}$ and $\mathbf{s}_{pre}^{Y}$ such that $P^b(H|\mathbf{s}_{pre}^{X}) > P^b(H|\mathbf{s}_{pre}^{Y})$.

Lemma 6 *Suppose two Bayesians X^{EB} and Y^{EB} are endowed with priors that coincide with the pre-screeners' posterior beliefs after the pre-period: $\omega_0^{X^{EB}} = \omega_{pre}^{X}$ and $\omega_0^{Y^{EB}} = \omega_{pre}^{Y}$. Then the* only *threshold at which X^{EB} and Y^{EB} trade is $n_{a,t} = n_{b,t}$: X^{EB} holds the asset when $n_{a,t} > n_{b,t}$ and Y^{EB} holds the asset when $n_{a,t} \leq n_{b,t}$.*

Proof. By Lemma 5, each Bayesian is endowed with independent priors on the state and credibility, where $\omega_0^{j^{EB},A} = 1/2$ and $\omega_0^{j^{EB},H} = P^b(c|\mathbf{s}_{pre}^{j})$ where $j \in \{X, Y\}$. Clearly, the analogous argument and conclusion of Lemma 4 apply to two endowed Bayesians: If $\omega_0^{X^{EB},H} > \omega_0^{Y^{EB},H}$, then $P^u(H|prior = \omega_0^{X^{EB}}, \mathbf{s}_t\}) > P^u(H|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t\})$. Thus, X^{EB} always trusts the source more than Y^{EB} does. By Proposition 4, this means that $P^u(H|prior = \omega_0^{X^{EB}}, \mathbf{s}_t) > P^u(H|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t)$ if and only if $P^u(A|prior = \omega_0^{X^{EB}}, \mathbf{s}_t) > P^u(A|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t)$ when $n_{a,t} > n_{b,t}$. Likewise, $P^u(H|prior = \omega_0^{X_{EB}}, \mathbf{s}_t) > P^u(H|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t)$ if and only if $P^u(A|prior = \omega_0^{X^{EB}}, \mathbf{s}_t) < P^u(A|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t)$ when $n_{a,t} < n_{b,t}$. When $n_{a,t} = n_{b,t}$, then by direct calculation $P^u(A|prior = \omega_0^{X^{EB}}, \mathbf{s}_t) = P^u(A|prior = \omega_0^{Y^{EB}}, \mathbf{s}_t) = 1/2$.

Thus, $E_t^{X^{EB}}(R) > E_t^{Y^{EB}}(R)$ when $n_{a,t} > n_{b,t}$, $E_t^{X^{EB}}(R) < E_t^{Y^{EB}}(R)$ when $n_{a,t} < n_{b,t}$, and $E_t^{X^{EB}}(R) = E_t^{Y^{EB}}(R)$ when $n_{a,t} = n_{b,t}$. Combining this with the law of iterated expectations also implies that $E_t^{X^{EB}}(R) > E_t^{Y^{EB}}[E_{t+1}^{X^{EB}}(R)] > E_t^{Y^{EB}}(R)$ when $n_{a,t} > n_{b,t}$, $E_t^{X^{EB}}(R) < E_t^{Y^{EB}}[E_{t+1}^{X^{EB}}(R)] < E_t^{Y^{EB}}(R)$ when $n_{a,t} < n_{b,t}$, and $E_t^{X^{EB}}(R) = E_t^{Y^{EB}}[E_{t+1}^{X^{EB}}(R)] = E_t^{Y^{EB}}(R)$ when $n_{a,t} = n_{b,t}$. Thus, the only threshold at which X^{EB} and Y^{EB} trade in the trading period is $n_{a,t} = n_{b,t}$ (i.e., when the two sides "switch sides" at $n_{a,t} = n_{b,t}$). ■

Consider pre-screeners X and Y. By Lemma 4, since $P^b(H|\mathbf{s}^X_{pre}) > P^b(H|\mathbf{s}^Y_{pre})$, then $P^b(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(H|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$. By Proposition 4, this means that $P^b(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(H|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ if and only if $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(A|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ when $n_{a,t} > n_{b,t}$. Likewise, $P^b(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^u(H|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ if and only if $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) < P^b(A|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ when $n_{a,t} < n_{b,t}$. When $n_{a,t} = n_{b,t}$, then by direct calculation $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) = P^b(A|\{\mathbf{s}^Y, \mathbf{s}_t\}) = 1/2$.

Thus, $E^X_t(R) > E^Y_t(R)$ when $n_{a,t} > n_{b,t}$, $E^X_t(R) < E^Y_t(R)$ when $n_{a,t} < n_{b,t}$, and $E^X_t(R) = E^Y_t(R)$ when $n_{a,t} = n_{b,t}$.

To determine trade between pre-screeners, we also need to compare $E^Y_t(E^X_{t+1}(R))$ to $E^X_t(R)$.

First, we will show that pre-screeners always trade whenever their endowed Bayesian counterparts do.

Lemma 7 *Suppose two pre-screeners have independent priors $\omega_{-\tau} = \omega^\theta_{-\tau}\omega^c_{-\tau}$ where $\omega^\theta_{-\tau} = 1/2$, and they observe pre-period signal paths $j \in \{X, Y\}$, respectively, where $n^j_{a,pre} = n^j_{b,pre} \geq 2$. Then X and Y always trade at the threshold $n_{a,t} = n_{b,t}$: X holds the asset when $n_{a,t} - n_{b,t} = 1$ and Y holds the asset when $n_{b,t} - n_{a,t} \in \{0, 1\}$.*

Proof. Note that:

$$E^X_t(E^X_{t+1}(R)) - E^X_t(R) = P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})[P^b(A|\{\mathbf{s}^X_{pre}, s_{t+1} = a\}) - P^u(A|prior = \omega^X_{pre}, s_{t+1} = a)]$$
$$+[1 - P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})][P^b(A|\{\mathbf{s}^X_{pre}, s_{t+1} = b\}) - P^u(A|prior = \omega^X_{pre}, s_{t+1} = b)].$$

Recall that $n^j_{a,pre} = n^j_{b,pre}$ for signal paths $j \in \{X, Y\}$. Let $d = |n^{all}_a - n^{all}_b| = |n_{a,t+1} - n_{b,t+1}|$, where $n^{all}_a = n_{a,pre} + n_{a,t+1}$ is the number of a's in $\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}$ and $n^{all}_b = n_{b,pre} + n_{b,t+1}$ is the number of b's in $\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}$.

We can easily show that there exists a unique threshold $\overline{d} > 2$ such that $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}) < \omega^H_{-\tau}$ when $d < \overline{d}$. By direct calculation when $\omega^A_{-\tau} = 1/2$:

$$P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}) = \frac{\omega^H_{-\tau}(q_H^{n^{all}_a}(1-q_H)^{n^{all}_b} + (1-q_H)^{n^{all}_a} q_H^{n_b})}{\omega^H_{-\tau}(q_H^{n^{all}_a}(1-q_H)^{n^{all}_b} + (1-q_H)^{n^{all}_a} q_H^{n^{all}_b}) + (1-\omega^H_{-\tau})(q_L^{n^{all}_a}(1-q_L)^{n^{all}_b} + (1-q_L)^{n^{all}_a} q_L^{n^{all}_b})}$$

WLOG suppose that $n^{all}_a \geq n^{all}_b \geq 2$ (if $n_{b,t+1} \geq n_{a,t+1} \geq 2$, then we can perform the same exercise by switching $n_{a,t+1}$ and $n_{b,t+1}$).[15] Then $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}) < \omega^H_{-\tau}$ if and only if

$$(q_H(1-q_H))^{n^{all}_b}(q^d_H + (1-q_H)^d) < (q_L(1-q_L))^{n^{all}_b}(q^d_L + (1-q_L)^d)$$
$$1 < \left(\frac{q_L(1-q_L)}{q_H(1-q_H)}\right)^{n^{all}_b} \left(\frac{q^d_L + (1-q_L)^d}{q^d_H + (1-q_H)^d}\right),$$

[15] Recall that we are restricting this proof to $n^{all}_b \geq 2$ because two pre-screeners will have exactly the same beliefs, and therefore no trade ever occurs, if the pre-paths X and Y have $n^j_{a,pre} = n^j_{b,pre} \leq 1$. Thus the relevant case to consider trading is $n_{b,pre} \geq 2$.

which we can verify is satisfied when $d \in \{0,1,2\}$. Moreover, the right-hand side is decreasing in d. Thus there exists a unique threshold $\overline{d} > 2$ such that $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}) < \omega^H_{-\tau}$ when $d < \overline{d}$.

Suppose $n^{all}_b \geq 2$ and $n^{all}_a - n^{all}_b \in \{0,1\}$. Then $d \in \{0,1,2\}$, so $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1}\}) < \omega^H_{-\tau}$. By Proposition 5, $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = a\}) - P^u(A|prior = \omega^X_{pre}, s_{t+1} = a) \leq 0$ and $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = a\}) - P^u(A|prior = \omega^X_{pre}, s_{t+1} = a) \leq 0$. This implies that $E^X_t(E^X_{t+1}(R)) - E^X_t(R) \leq 0$, with strict equality only for $n^{all}_a = n^{all}_b$. Therefore, $E^X_t(R) \geq E^X_t(E^X_{t+1}(R)) \geq E^Y_t(E^X_{t+1}(R))$ and $E^X_t(R) \geq E^Y_t(R)$ with strict equality only for $n^{all}_a = n^{all}_b$. Thus, $E^X_t(R) > \max\{E^Y_t(E^X_{t+1}(R)), E^Y_t(R)\}$ when $n^{all}_a - n^{all}_b = 1$ and X holds the asset. Applying the same argument symmetrically, $E^X_t(R) < \max\{E^Y_t(E^X_{t+1}(R)), E^Y_t(R)\}$ when $n^{all}_b - n^{all}_a = 1$ and Y holds the asset. Since $n_{a,pre} = n_{b,pre}$, this implies that X and Y always trade at the threshold $n_{a,t} = n_{b,t}$: X holds the asset when $n_{a,t} - n_{b,t} = 1$ and Y holds the asset when $n_{b,t} - n_{a,t} \in \{0,1\}$. ■

To show that speculative trade can occur when agents are pre-screeners, suppose that a signal path $\mathbf{s}_t$ with $n_{a,t} > n_{b,t}$. We have shown in Lemma 6 that endowed Bayesians will never trade in this case, and X will hold the asset. Moreover, we know that $p_t = E^X_t(R) > E^Y_t(R)$ when $n_{a,t} > n_{b,t}$. Thus, the pre-screeners will only trade in period t if there exists some point at which Y buys the asset from X because $E^Y_t(E^X_{t+1}(R)) > E^X_t(R)$. Note that we can re-write this as:

$$\begin{aligned}
E^Y_t(E^X_{t+1}(R)) - E^X_t(R) = {} & P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})E^X_{t+1}(R|s_{t+1} = a) \\
& + (1 - P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\}))E^X_{t+1}(R|s_{t+1} = b) \\
& - P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})E_{t+1}(R|prior = \omega^X_{pre}, s_{t+1} = a) \\
& - (1 - P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}))E_{t+1}(R|prior = \omega^X_{pre}, s_{t+1} = b) \\
& = P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})\left(E^X_{t+1}(R|s_{t+1} = a) - E_{t+1}(R|prior = \omega^X_{pre}, s_{t+1} = a)\right) \\
& +(1 - P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}))\left(E^X_{t+1}(R|s_{t+1} = b) - E_{t+1}(R|prior = \omega^X_{pre}, s_{t+1} = b)\right) \\
& + \left(P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) - P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})\right)\left(E^X_{t+1}(R|s_{t+1} = b) - E^X_{t+1}(R|s_{t+1} = a)\right).
\end{aligned} \tag{19}$$

By Proposition 5, Equation (19) is negative if we have signals such that $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = a\}) \leq \omega^H_{-\tau}$. Thus, a necessary condition for Equation (19) to be positive is that we have signals such that $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = a\}) > \omega^H_{-\tau}$.

To demonstrate existence of the speculative trade, fix the parameters $q_c \in [1/2, 1)$, prior $\omega_{-\tau}$, and signal paths $\mathbf{s}^X_{pre}$ and $\mathbf{s}_t$ such that $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = b\}) > \omega^H_{-\tau}$, so the first two terms of Equation (19) are strictly positive. Since $n_a > n_b$ and $E^X_t(R) > E^Y_t(R)$, then the third term of Equation (19) is strictly negative because $P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\} > P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ whenever $E^X_t(R) > E^Y_t(R)$. Note that the third term of Equation

(19) is strictly increasing in $P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$ and equals zero if $\mathbf{s}^Y_{pre} = \mathbf{s}^X_{pre}$. We will show that there exists some signal path $\bar{\mathbf{s}}^Y_{pre}$ such that Equation (19) is satisfied whenever $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = b\}) > \omega^H_{-\tau}$, so $P^u(H|\{\mathbf{s}^X_{pre}, \mathbf{s}_t, s_{t+1} = b\}) \geq \omega^H_{-\tau}$ is a sufficient condition for extra trade.

To find the sequence(s) $\mathbf{s}^Y_{pre}$ to satisfy Equation (19) when $n^Y_{a,pre} = n^X_{a,pre}$ and $n^Y_{b,pre} = n^X_{b,pre} > 1$: Let $\mathbf{s}^Y_{pre}$ be identical to $\mathbf{s}^X_{pre}$ in all positions except for the last reversal pair (a, b), if it exists, in the sequence $\mathbf{s}^X_{pre}$. Call the positions of this pair $j + 1$ and $j + 2$ (so this means that $s^X_{j+1} = a$ and $s^X_{j+2} = b$). If $n^j_{a,pre} > n^j_{b,pre}$ (i.e., more a's have been observed than b's in the subsequence $\mathbf{s}^X_j$, which is the first j signals of $\mathbf{s}^X_{pre}$), then replace this (a, b) with (b, a) so that $s^Y_{j+1} = b$ and $s^Y_{j+2} = a$. If $n^j_{a,pre} \leq n^j_{b,pre}$ or the pair (a, b) does not exist in $\mathbf{s}^X_{pre}$, then instead find the last reversal pair (b, a) in the sequence $\mathbf{s}^X_{pre}$. Call the positions of this pair $k + 1$ and $k + 2$ (so this means that $s^X_{k+1} = b$ and $s^X_{k+2} = a$).If $n^k_{b,pre} > n^k_{a,pre}$ (i.e., more b's have been observed than a's in the subsequence $\mathbf{s}^X_k$, which is the first k signals of $\mathbf{s}^X_{pre}$), then replace this (b, a) with (a, b) so that $s^Y_{k+1} = a$ and $s^X_{k+2} = b$. If $n^k_b \leq n^k_a$, then continue by finding the second-to-last reversal pair (a, b) and applying this procedure, and so on. By the argument made in the proof of Proposition 2, this constructed sequence generates the greatest degree of trust such that we still have $P^b(H|\mathbf{s}^Y_{pre}) < P^b(H|\mathbf{s}^X_{pre})$, and therefore by Lemma 4 and Proposition 4 it generates the greatest belief in A such that we still have $E^Y_t(R|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\})) < E^X_t(R|\{\mathbf{s}^X, \mathbf{s}_t\})$. By Proposition 2, we can construct such a sequence $\mathbf{s}^Y_{pre}$ as long as $\mathbf{s}^X_{pre}$ is not the sequence that generates the minimal degree of trust. This is already satisfied by assumption that $P^b(H|\mathbf{s}^X_{pre}) > P^b(H|\mathbf{s}^Y_{pre})$. We can continue constructing sequences that lead to decreasing degrees of trust by iterating in this procedure for each constructed $\mathbf{s}^Y_{pre}$.

Then there exists some signal path $\bar{\mathbf{s}}^Y_{pre}$ such that $E^X_t(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > E^Y_t(R|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}))$ and $E^Y_t(E^X_{t+1}(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}) - E^X_t(R) = 0$. This implies that for any signal path $\mathbf{s}^Y_{pre}$ that results in beliefs such that $P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\}) > P^b(s_{t+1} = a|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\})$, we thus have $E^X_t(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > E^Y_t(R|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}))$ and $E^Y_t(E^X_{t+1}(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}) - E^X_t(R) > 0$. Thus, Y will buy the asset from X. Likewise, for any signal path $\mathbf{s}^Y_{pre}$ that results in beliefs such that $P^b(s_{t+1} = a|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > P^b(s_{t+1} = a|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}) > P^b(s_{t+1} = a|\{\mathbf{s}^Y_{pre}, \mathbf{s}_t\})$, we have $E^X_t(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\}) > E^Y_t(R|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}))$ and $E^Y_t(E^X_{t+1}(R|\{\mathbf{s}^X_{pre}, \mathbf{s}_t\})|\{\bar{\mathbf{s}}^Y_{pre}, \mathbf{s}_t\}) - E^X_t(R) < 0$. Thus, Y will not buy the asset from X.

Suppose that $n_{a,t} < n_{b,t}$. As we have already shown, $p_t = E^X_t(R) < E^Y_t(R)$ when $n_{a,t} < n_{b,t}$. Thus, $E^X_t(R) < \max\{E^Y_t(E^X_{t+1}(R)), E^Y_t(R)\}$ when $n_{a,t} < n_{b,t}$ and no trade occurs because Y always holds the asset. Thus, $n_{a,t} > n_{b,t}$ is a necessary condition for speculative trade between pre-screeners to occur.

Suppose that $\mathbf{s}^X_{pre} = \mathbf{s}^Y_{pre}$. Since Y buys the asset whenever $\max\{E^Y_t(R), E^Y_t(E^X_{t+1}(R))\} \geq E^X_t(R)$, then Y always buys the asset if $\mathbf{s}^X_{pre} = \mathbf{s}^Y_{pre}$. Therefore, when $\mathbf{s}^X_{pre} = \mathbf{s}^Y_{pre}$, there is no trade between X and Y, nor between X^{EB} and Y^{EB}.

Second, suppose groups X and Y observe $\mathbf{s}^X_{pre}$ and $\mathbf{s}^Y_{pre}$ such that $P^b(H|\mathbf{s}^X_{pre}) < P^b(H|\mathbf{s}^Y_{pre})$.

We can apply the exact same analysis as in the preceding case of $P^b(H|\mathbf{s}_{pre}^X) > P^b(H|\mathbf{s}_{pre}^Y)$ to show that the mirror image holds when $P^b(H|\mathbf{s}_{pre}^X) < P^b(H|\mathbf{s}_{pre}^Y)$. For brevity, we do not repeat it in detail. In particular, it is easy to show that Y will always hold the asset when $n_{a,t} > n_{b,t}$, and X will buy it at $n_{a,t} = n_{b,t}$. The key portion is that extra trade can only occur if $n_{b,t} < n_{a,t}$ and $E_t^Y(E_{t+1}^X(R)) - E_t^X(R) > 0$ as in Equation (19). Again by Proposition 5, Equation (19) is negative if we have signals such that $P^u(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t, s_{t+1} = a\}) \geq \omega_{-\tau}^H$. Thus, a necessary condition for Equation (19) to be positive is that we have signals such that $P^u(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t, s_{t+1} = a\}) < \omega_{-\tau}^H$. Likewise, the sufficient condition is that $P^u(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t, s_{t+1} = b\}) \leq \omega_{-\tau}^H$. When the necessary and sufficient conditions are satisfied given $\mathbf{s}_{pre}^X$ and $\mathbf{s}_t$, then there exists at least one signal path $\mathbf{s}_{pre}^Y$ such that Y holds the asset.

Proof of Proposition 7

Since the pre-screener and endowed Bayesian have the same beliefs at the start of the trading period, $p_0^b = p_0^{EB} = 1/2$. As can be seen from the proof of Proposition 4, $p_T^b = p_T^{EB} = 1/2$ when $n_{a,T} = n_{b,T}$. Also $n_{a,t} > n_{b,t}$ implies that $p_t^b > 1/2$ and $p_t^{EB} > 1/2$.

1. First, we demonstrate existence of $p_t^b > p_t^{EB} > 1/2$. Note that

$$\begin{aligned} p_t^b &= P^b(A|\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) \\ &= \frac{\omega_{-\tau}^A \sum_c \left(\prod_{i=-\tau+1}^0 P(s_i|q,A)\right)\left(\prod_{i=1}^t P(s_i|q,A)\right)\omega_{-\tau}^c \beta_c(\{\mathbf{s}_{pre}^X, \mathbf{s}_t\})}{\sum_\theta \omega_{-\tau}^\theta \sum_c \left(\prod_{i=-\tau+1}^0 P(s_i|q,A)\right)\left(\prod_{i=1}^t P(s_i|c,\theta)\right)\omega_{-\tau}^c \beta_c(\{\mathbf{s}_{pre}^X, \mathbf{s}_t\})} \\ p_t^{EB} &= P^u(A|prior = \omega_{pre}^X, \mathbf{s}_t) \\ &= \frac{\omega_{-\tau}^A \sum_c \left(\prod_{i=-\tau+1}^0 P(s_i|q,A)\right)\left(\prod_{i=1}^t P(s_i|q,A)\right)\omega_{-\tau}^c \beta_c(\mathbf{s}_{pre}^X)}{\sum_\theta \omega_{-\tau}^\theta \sum_c \left(\prod_{i=-\tau+1}^0 P(s_i|q,A)\right)\left(\prod_{i=1}^t P(s_i|c,\theta)\right)\omega_{-\tau}^c \beta_c(\mathbf{s}_{pre}^X)} \end{aligned}$$

where

$$\begin{aligned} \beta_c(\mathbf{s}_{pre}^X) &= \prod_{m=-\tau}^{0}\left(\sum_\theta\left(\prod_i^m P(s_i|c,\theta)\right)\omega_{-\tau}^\theta\right) \\ \beta_c(\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) &= \beta_c(\mathbf{s}_{pre}^X)\prod_{m=1}^{n}\left(\sum_\theta\left(\prod_{i=-\tau+1}^{0} P(s_i|c,\theta)\right)\left(\prod_{i=1}^{m} P(s_i|c,\theta)\right)\omega_{-\tau}^\theta\right). \end{aligned}$$

By direct comparison, $p_t^b > p_t^{EB}$ if and only if $\omega_{-\tau}^H(1-\omega_{-\tau}^H)\beta_H(\mathbf{s}_{pre}^X)\beta_L(\mathbf{s}_{pre}^X)FG > 0$,

where

$$F = \left(\prod_{m=1}^{t}\left(\sum_{\theta}\left(\prod_{t=-\tau}^{0} P(s_i|H,\theta)\right)\left(\prod_{i=1}^{m} P(s_i|H,\theta)\right)\omega_{-\tau}^{\theta}\right)\right) - \left(\prod_{m=1}^{t}\left(\sum_{\theta}\left(\prod_{t=-\tau}^{0} P(s_i|L,\theta)\right)\left(\prod_{i=1}^{m} P(s_i|L,\theta)\right)\omega_{-\tau}^{\theta}\right)\right)$$

$$G = \left(\prod_{i=-\tau+1}^{0} P(s_i|H,A)\right)\left(\prod_{i=1}^{t} P(s_i|H,A)\right)\left(\prod_{i=-\tau+1}^{0} P(s_i|L,B)\right)\left(\prod_{i=1}^{t} P(s_i|L,B)\right)$$
$$- \left(\prod_{i=-\tau+1}^{0} P(s_i|H,B)\right)\left(\prod_{i=1}^{t} P(s_i|H,B)\right)\left(\prod_{i=-\tau+1}^{0} P(s_i|L,A)\right)\left(\prod_{i=1}^{n} P(s_i|L,A)\right),$$

where we have already shown in the Proof of Proposition 5 that $G > 0$ when $n_{a,t} > n_{b,t}$, since $n_{a,pre}^X = n_{b,pre}^X$. Thus, $p_t^b > p_t^{EB} > 0$ if and only if $F > 0$. Moreover, we can easily show that $P^b(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) > P^u(H|prior = \omega_{pre}^X, \mathbf{s}_t)$ if and only if $F > 0$. Thus, $p_t^b > p_t^{EB} > 1/2$ if and only if $P^b(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) > P^u(H|prior = \omega_{pre}^X, \mathbf{s}_t)$. By the argument in Proposition 2, there exists at least one path $\mathbf{s}_t$ that generates $P^b(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) > P^u(H|prior = \omega_{pre}^X, \mathbf{s}_t)$. For example, if $\mathbf{s}_t$ contains $n_{a,t} > 1$ signals, we can show that there exists a unique threshold $\overline{n}_{a,t}(n_{b,t})$ such that $F > 0$ if $n_{a,t} > \overline{n}_{a,t}(n_{b,t})$ and $F \leq 0$ if $n_{a,t} \leq \overline{n}_{a,t}(n_{b,t})$. By direct computation, $F < 0$ when $n_{a,t} = n_{b,t}$, $\frac{\partial F}{\partial n_{a,t}} > 0$, and $\lim_{n_{a,t}\to\infty} F = \infty$. Thus, there exists a unique threshold $\overline{n}_{a,t}$ such that $p_t^b > p_t^{EB}$ for all $\overline{n}_{a,t} < n_{a,t} \leq n_{a,T}$ and $p_t^b \leq p_t^{EB}$ for all $0 \leq n_a \leq \overline{n}_a$.

Clearly, we can reverse the inequalities to show $p_t^{EB} > p_t^b > 1/2$ if and only if $P^b(H|\{\mathbf{s}_{pre}^X, \mathbf{s}_t\}) < P^u(H|prior = \omega_{pre}^X, \mathbf{s}_t)$.

2. Suppose we have $\mathbf{s}_T$ such that $p_{\hat{t}}^b > p_{\hat{t}}^{EB} > 1/2$. Since $p_0^b = p_0^{EB} = p_T^b = p_T^{EB} = 1/2$ and $p_{\hat{t}}^b > p_{\hat{t}}^{EB}$, then the average price change of p_t^b must be strictly greater than the average price change of p_t^{EB} for $t \in [0, \hat{t}]$ and $t \in [\hat{t}, T]$.

   By Proposition 5, $p_1^b < p_1^{EB}$ because $P^u(H|\{\mathbf{s}_{pre}^X, s_1 = a\}) < \omega_{-\tau}^H$ for $n_{a,pre}^X = n_{b,pre}^X \geq 1$. This implies that we can only have $p_{\hat{t}}^b > p_{\hat{t}}^{EB} \geq 0$ if there are sufficiently many a's that the pre-screener over-reacts (under-reacts) to confirming (disconfirming) signals (i.e., more a signals). By Proposition 5, if at least one such $p_{t'}^b > p_{t'}^{EB}$ exists, then it must be that $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{t'}) > \omega_{-\tau}^H$ for at least one of these t' dates. Suppose there exists some t'' such that $p_{t''}^b > p_{t''}^{EB}$ but $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{t'}) \leq \omega_{-\tau}^H$. Since $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{t'}) \leq \omega_{-\tau}^H$, then there must necessarily be a lower proportion of a's than b's observed at t'' than at t'. Thus it cannot be that $p_{t''}^b = \max p_t^b$ and we must have that $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{\hat{t}}) > \omega_{-\tau}^H$ whenever $p_{\hat{t}}^b > p_{\hat{t}}^{EB}$.

   Suppose $p_{\hat{t}}^b = p_{\hat{t}}^{EB}$ and $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{\hat{t}}) > \omega_{-\tau}^H$ but $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b) < \omega_{-\tau}^H$. Since $p_{\hat{t}}^b \equiv \max p_t^b$, this implies $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{\hat{t}}, \mathbf{s}_t) < \omega_0^H$ for all signal paths $\mathbf{s}_t$ with $t \in (\hat{t}, T]$ because they are below the peak. By Proposition 7, this implies that the pre-screener under-reacts to each a and over-reacts to each b in $t \in (\hat{t}, T]$. But this means that for any t such that $n_{a,t} = n_{b,t}$, $p_t^b < p_t^{EB}$, which cannot be true. Thus, $P^u(H|\mathbf{s}_{pre}^X, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} =$

$b) \geq \omega^H_{-\tau}$ when $p^b_{\hat{t}} = p^{EB}_{\hat{t}}$. If $p^b_{\hat{t}} > p^{EB}_{\hat{t}}$, then the Bayesian posterior belief in reliability must be even higher than when $p^b_{\hat{t}} = p^{EB}_{\hat{t}}$. Thus $P^u(H|\mathbf{s}^X, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b) \geq \omega^H_{-\tau}$ when $p^b_{\hat{t}} \geq p^{EB}_{\hat{t}}$.

Since $p_{\hat{t}} \equiv \max p_t$, then $s_{\hat{t}+1} = b$ and $p^{EB}_{\hat{t}+1} - p^{EB}_{\hat{t}} < 0$ and $p^b_{\hat{t}+1} - p^b_{\hat{t}} < 0$. Moreover, because have we have shown that $P^u(H|\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b) \geq \omega^H_{-\tau}$, then $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b\}) > P^u(A|prior = \omega^b_{\hat{t}}, s_{\hat{t}+1} = b)$ by Proposition 7, where each joint belief for the prior $\omega^b_{\hat{t}}$, denoted by $\omega^{c\theta}_{\hat{t}}$, is the pre-screener's belief after observing signal path $\mathbf{s}^X_{pre}$ and the public path $\mathbf{s}_{\hat{t}}$:

$$\omega^{c\theta}_{\hat{t}} \equiv P^b(c, \theta|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}\}) = \frac{\omega^A_{-\tau} \sum_c \left(\prod^0_{i=-\tau+1} P(s_i|q, A)\right) \left(\prod^{\hat{t}}_{i=1} P(s_i|q, A)\right) \omega^c_{-\tau} \beta_c(\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}\})}{\sum_\theta \omega^\theta_{-\tau} \sum_c \left(\prod^0_{i=-\tau+1} P(s_i|q, A)\right) \left(\prod^{\hat{t}}_{i=1} P(s_i|c, \theta)\right) \omega^c_{-\tau} \beta_c(\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}\})} \tag{20}$$

Since $p^b_{\hat{t}} = P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}\}) = P^u(A|prior = \omega^b_{\hat{t}})$ and $p^{EB}_{\hat{t}} = P^u(A|prior = \omega^X_{pre}, \mathbf{s}_{\hat{t}})$ by definition, then $p^b_{\hat{t}} > p^{EB}_{\hat{t}}$ implies $P^u(A|prior = \omega^b_{\hat{t}}) > P^u(A|prior = \omega^X_{pre}, \mathbf{s}_{\hat{t}})$. Thus $0 > P^u(A|prior = \omega^b_{\hat{t}}, s_{\hat{t}+1} = b) - P^u(A|prior = \omega^b_{\hat{t}}) > P^u(A|prior = \omega^X_{pre}, \{\mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b\}) - P^u(A|prior = \omega^X_{pre}, \mathbf{s}_{\hat{t}})$. Combining this with the fact that $P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b\}) > P^u(A|prior = \omega^b_{\hat{t}}, s_{\hat{t}+1} = b)$ by Proposition 7, then $0 > P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b\}) - P^b(A|\{\mathbf{s}^X_{pre}, \mathbf{s}_{\hat{t}}\}) > P^u(A|prior = \omega^X_{pre}, \{\mathbf{s}_{\hat{t}}, s_{\hat{t}+1} = b\}) - P^u(A|prior = \omega^X_{pre}, \mathbf{s}_{\hat{t}})$. Thus, $|p^b_{\hat{t}+1} - p^b_{\hat{t}}| < |p^{EB}_{\hat{t}+1} - p^{EB}_{\hat{t}}|$.

3. If $p^b_{\hat{t}} > p^{EB}_{\hat{t}} > 1/2$, then we have already shown in the proof of Proposition 7 that the necessary and sufficient conditions given in Proposition 6 must hold. Thus, there exists at least one signal path $\mathbf{s}^Y_{pre}$ such that extra trade between pre-screeners occurs (e.g., Y holds the asset at $t = \hat{t}$, at least).

Internet Appendix

Trust in Signals and the Origins of Disagreement

A Generalized Pre-Screening

Let $\omega_0^{c\theta}$ be the prior belief on credibility c and state θ, where $\sum_c \sum_\theta \omega_0^{c\theta} = 1$.

When the prior beliefs about the credibility and state can potentially be correlated, we cannot apply the first-stage updated belief $\kappa_c(\mathbf{s}_n)$, which is a marginal belief on reliability, directly to the second stage in place of a prior belief on credibility because the joint priors on reliability and state are not independent. Therefore, the generalized pre-screening algorithm requires the second stage to apply Bayes' Rule to a belief whose marginal prior about credibility sums to $\kappa_c(\mathbf{s}_n)$. Thus, in the second stage, we assume that the agent applies the *weighted* first-stage updated belief $\kappa_c(\mathbf{s}_n)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)$ in place of the existing joint prior. If the prior beliefs about credibility and state are independent ($\omega_0^{c\theta} = \omega_0^c \omega_0^\theta$ for all c and θ), then Equations (21) and (22) reduce to Equations (2) and (3) respectively.

To illustrate the pre-screener's updating algorithm, suppose she observes two signals, one in each period. After observing the first signal (s_1), the pre-screener's updated belief about the source's credibility, $\kappa_c(s_1)$, is:

$$\kappa_c(s_1) = \frac{\sum_\theta P(s_1|c,\theta)\omega_0^{c\theta}}{\sum_c \sum_\theta P(s_1|c,\theta)\omega_0^{c\theta}}.$$

Using the weighted first-stage updated belief $\kappa_c(s_1)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)$ to form her joint posterior belief on the state and credibility, $P^b(c,\theta|s_1)$, yields her posterior beliefs after the first signal:

$$P^b(c,\theta|s_1) = \frac{P(s_1|c,\theta)\kappa_c(s_1)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)}{\sum_c \sum_\theta P(s_1|c,\theta)\kappa_c(s_1)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)}.$$

After observing the second signal (s_2), the pre-screener's updated belief about the source's credibility, $\kappa_c(s_1, s_2)$ is

$$\kappa_c(s_1, s_2) = \frac{\sum_\theta P(s_2|c,\theta)P^b(c,\theta|s_1)}{\sum_c \sum_\theta P(s_2|c,\theta)P^b(c,\theta|s_1)}.$$

Using the weighted first-stage updated belief $\kappa_c(s_1, s_2)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)$ to form her joint posterior belief on the state and credibility, $P^b(c,\theta|s_1,s_2)$, yields:

$$P^b(c,\theta|s_1,s_2) = \frac{P(s_2|c,\theta)P(s_1|c,\theta)\kappa_c(s_1,s_2)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)}{\sum_c\sum_\theta P(s_2|c,\theta)P(s_1|c,\theta)\kappa_c(s_1,s_2)\left(\frac{\omega_0^{c\theta}}{\sum_\theta \omega_0^{c\theta}}\right)}.$$

Iterating on the pre-screener's updating process allows us to characterize her posterior beliefs:

Applying the generalized pre-screening procedure described above to prior beliefs $\omega_0^{c\theta}$ yields:

$$\kappa_c(\mathbf{s}_n) = \frac{\left(\frac{\kappa_c(\mathbf{s}^{n-1})}{\sum_\theta \omega_0^{c\theta}}\right)\sum_\theta\left(\prod_{t=1}^n P(s_t|c,\theta)\omega_0^{c\theta}\right)}{\sum_c\left(\frac{\kappa_c(\mathbf{s}^{n-1})}{\sum_\theta \omega_0^{c\theta}}\right)\sum_\theta\left(\prod_{t=1}^n P(s_t|c,\theta)\omega_0^{c\theta}\right)}, \tag{21}$$

where $\kappa_c(\emptyset) = \sum_\theta \omega_0^{c\theta}$.

$$P^b(c,\theta|\mathbf{s}_n) = \frac{\left(\prod_{t=1}^n P(s_t|c,\theta)\right)\left(\frac{\kappa_c(\mathbf{s}_n)}{\sum_\theta \omega_0^{c\theta}}\right)\omega_0^{c\theta}}{\sum_c\sum_\theta\left(\prod_{t=1}^n P(s_t|c,\theta)\right)\left(\frac{\kappa_c(\mathbf{s}_n)}{\sum_\theta \omega_0^{c\theta}}\right)\omega_0^{c\theta}} \tag{22}$$

$$= \frac{\beta_{c\theta}(\mathbf{s}_n)\left(\frac{1}{\sum_\theta \omega_0^{c\theta}}\right)^n\left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^{c\theta}}{\sum_c\left(\frac{1}{\sum_\theta \omega_0^{c\theta}}\right)^n\sum_\theta \beta_{c\theta}(\mathbf{s}_n)\left(\prod_{t=1}^n P(s_t|c,\theta)\right)\omega_0^{c\theta}}. \tag{23}$$

where $\beta_{c\theta}(\mathbf{s}_n)$ is given by:

$$\beta_{c\theta}(\mathbf{s}_n) = \left(\sum_\theta P(s_1|c,\theta)\omega_0^{c\theta}\right)\times\left(\sum_\theta P(s_1|c,\theta)P(s_2|c,\theta)\omega_0^{c\theta}\right)\times\ldots\times\left(\sum_\theta P(s_1|c,\theta)P(s_2|c,\theta)\ldots P(s_n|c,\theta)\omega_0^{c\theta}\right)$$

$$= \prod_{m=1}^n\left(\sum_\theta\left(\prod_{t=1}^m P(s_t|c,\theta)\right)\omega_0^{c\theta}\right), \tag{24}$$

B Detailed Intuition for Proposition 5, Part 1

In more detail, consider first the endowed Bayesian's (EB's) beliefs. EB begins with ω_b^n and applies it to evaluate s_{n+1}. But because ω_b^n is generated by applying a prior of $\kappa_c(\mathbf{s}_n)$ to signals $\mathbf{s}_n$, and a Bayesian's beliefs are invariant to signal order, this is equivalent to applying a prior of $\kappa_c(\mathbf{s}_n)$ to $\{\mathbf{s}_n, s_{n+1}\}$. EB's resulting marginal posterior belief about reliability then equals $\kappa_c(\{\mathbf{s}_n, s_{n+1}\})$, the pre-screener's first-stage updated belief of credibility. However, in applying the second step, the pre-screener makes the mistake of using $\kappa_c(\{\mathbf{s}_n, s_{n+1}\})$ as a prior to to evaluate $\{\mathbf{s}_n, s_{n+1}\}$, whereas EB used $\kappa_c(\{\mathbf{s}_n\})$.

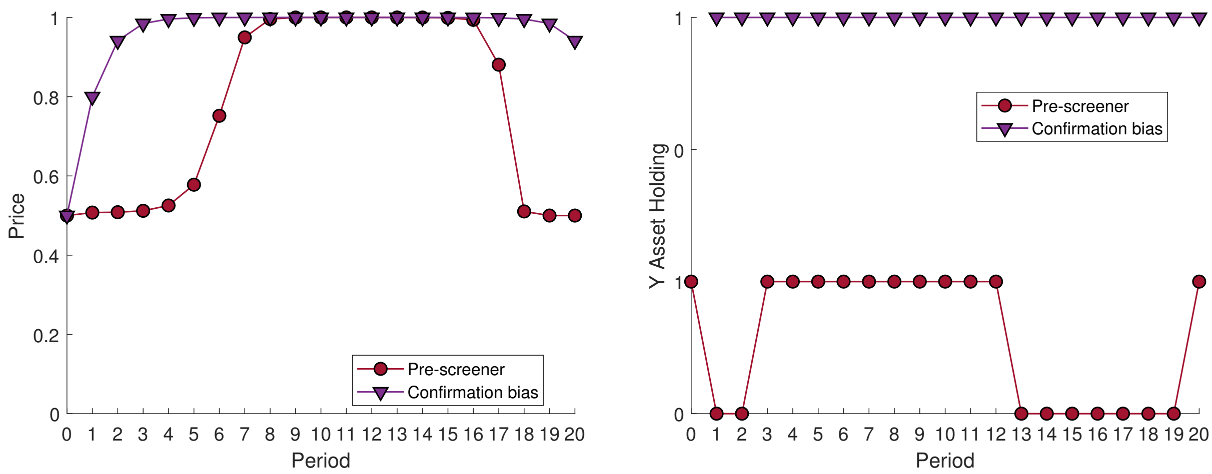
As a result, the pre-screener will think A is more likely than the endowed Bayesian if and only if the *new* signal s_{n+1} makes her strictly more confident that the source is high reliability relative to EB's prior over *all* signals $\{\mathbf{s}_n, s_{n+1}\}$ and the combined evidence $\{\mathbf{s}_n, s_{n+1}\}$ objectively strictly favors A. This is the condition that $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\{\mathbf{s}_n\})$ in Part 1(a). Similarly, Part 1(b) describes how the pre-screener will think A is less likely than the endowed Bayesian if and only if the new signal makes her strictly less confident that the source is high reliability relative to EB's prior over all signals $\{\mathbf{s}_n, s_{n+1}\}$, which is the condition that $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\{\mathbf{s}_n\})$. Part 1(c) reflects the edge case.

Importantly, Parts 1(a) and 1(b) can arise when new signals s_{n+1} are either confirmatory or contradictory. In Part 1(b), the opposite of confirmation bias arises: there is under-reaction to a confirmatory signal and over-reaction to a contradictory signal. Under-reaction arises because we can have $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\{\mathbf{s}_n\})$ even though the signal is confirmatory and thus $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \omega_b^{n,H}$—that is, even though the first-stage belief that the source is high reliability increases over the time-n posterior. Intuitively, $\omega_b^{n,H}$ arises from applying $\kappa_H(\{\mathbf{s}_n\})$ to $\mathbf{s}_n$, while $\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$ arises from applying $\kappa_H(\{\mathbf{s}_n\})$ to $\{\mathbf{s}_n, s_{n+1}\}$. With confirmatory news, $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \omega_b^{n,H}$ but may fall below $\kappa_H(\{\mathbf{s}_n\})$ if the history of evidence contains some bad news about reliability. To be precise, the pre-screener's over-inference at each step implies $\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$ contains $\kappa_H(\{\mathbf{s}_n\})$ (see Equation 2). Factoring out $\kappa_H(\{\mathbf{s}_n\})$ followed by algebraic manipulation shows that the effective difference between the two equals the difference between an *objective* Bayesian's belief that the source is high reliability, $P^u(H|\{\mathbf{s}_n, s_{n+1}\})$ and the time-0 prior common to the pre-screener and objective Bayesian, ω_0^H. Thus, even if s_{n+1} is confirmatory, the pre-screener under-reacts if the *combined evidence* $\{\mathbf{s}_n, s_{n+1}\}$ is objectively bad news for credibility.

Over-reaction to the contradictory signal is similar. The intuition is more straightforward in this case because $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \kappa_H(\{\mathbf{s}_n\})$ and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) < \omega_b^{n,H}$—the new signal s_{n+1} is deleterious for both. Re-arranging the discussion above yields similar intuitions for why Part 1(a) and $\kappa_H(\{\mathbf{s}_n, s_{n+1}\}) > \kappa_H(\{\mathbf{s}_n\})$ can arise when new signals are either confirmatory or contradictory. Even though the resulting observed behavior is akin to confirmation bias, the mechanism is fundamentally different and arises due to the over-interpretation of signals embedded in $\kappa_H(\{\mathbf{s}_n, s_{n+1}\})$.

C Trading with Confirmation Bias

Agents in Rabin and Schrag (1999) have fixed beliefs about credibility but probabilistically flip signals that contradict their beliefs about the state. Figure 3 illustrates why this framework has difficulty generating a crash even when confirmation bias agents ("CB agents") are certain that the source is credible. We endow two CB agents with common neutral priors about the state, $\omega_0^{X^{CB}} = \omega_0^{Y^{CB}} = 0.5$. These equal the pre-screeners' beliefs about the state at time 0, $P(A|\mathbf{s}_{pre}^X) = P(A|\mathbf{s}_{pre}^Y) = 0.5$. The CB agents are certain that the source is credible (i.e., both X^{CB} and Y^{CB} believe $c = H$ with probability 1) but flip signals that oppose their



(a) Prices

(b) Asset holdings

Figure 3: Trading game with confirmation bias. This figure plots outcomes from a trading game between two groups of traders, X and Y, described in Internet Appendix Section C, when they are either both pre-screeners or both have confirmation bias (CB). Prices are in Panel (a) and the asset holdings of group Y are in Panel (b). Realized signals in periods 1-10 are good cash flow news ('a'), while periods 11-20 have bad cash flow news ('b'). Trader beliefs at period 0 are equal to the beliefs that pre-screeners would have after observing '$aabb$' for X and '$abab$' for Y, starting from common priors. Parameters (defined in Section C) are $\{q_H, q_L, \omega^A_{-\tau}, \omega^H_{-\tau}, T, \tau\} = \{0.8, 0.5, 0.5, 0.3, 20, 4\}$. Confirmation bias traders begin period 0 with $\omega_0^{X^{CB}} = \omega_0^{Y^{CB}} = 0.5$, which equal the pre-screeners' beliefs about the state at time 0, and are certain that the source is type H. CB traders independently flip signals that oppose their beliefs about the state with probability 0.2.

current beliefs about the state with probability 0.2, as in Rabin and Schrag (1999). We then expose all agents to the same signals during the public trading period.

Prices rise more quickly in periods 1-10 between two confirmation bias agents relative to the game between two pre-screeners. This quick rise is by construction as the CB agents are certain the source is high quality, while pre-screeners are learning about credibility. However, there is no crash in periods 11-20 in the game between CB agents, even though they are *certain* the source is credible. This is because the CB agents are resistant too *all* of the bad cash flow news during those periods: State A is subjectively more likely, and so they flip each bad cash flow signal with a fixed probability. The only way for a crash to occur is for the cash flow R to realize after the end of period 20. In contrast, the pre-screeners are *initially* resistant to bad cash flow news (confirmation bias, Proposition 5 Part 1a), but after enough bad cash flow news, there is a subsequent, endogenous, sudden collapse in agents' beliefs about source credibility (the "opposite" of confirmation bias, Proposition 5 Part 1b), leading prices to fall quickly. Furthermore, because the CB agents begin the trading period with common priors about the state and see common public signals, they do not trade unless they randomly and exogenously flip different signals.

These insights are analogous to Rabin and Schrag (1999) Proposition 1: agents with confirmation bias can have beliefs that are so sticky that beliefs cannot change suddenly. In our example, even if more bad cash flow news arrived before cash flows realize (e.g., if we extended the game through period 21), the price would continue to exceed 0.5.

D Supplemental Propositions

Proposition 8 characterizes sufficient conditions under which pre-screeners exhibit confirmation bias or its opposite, regardless of the order in which they observed signals. Without loss of generality, assume that the source has sent more a signals than b signals ($n_a > n_b$).

Proposition 8 (Over- and under-trust without knowing signal order) *Let* $(\omega_0^A, \omega_0^H) = (1/2, \hat{\omega})$ *for any* $\hat{\omega} \in (0,1)$*, and let* $n_a > n_b$*. Whether pre-screeners exhibit confirmation bias or the opposite depends on the relative proportion of a's and b's and the distribution of beliefs about reliability:*

1. *There exists some* n_a^* *and* $\check{q} > \frac{1}{2}$ *such that the agent overtrusts the source and is optimistic that the state is A for* any *sequence with fixed* n_a, n_b *when* $n_b < n_a^* < n_a$ *and* $q_L < q_H \leq \check{q}$.
2. *There exists some* $\hat{n}_b$, $\underline{q} > \frac{1}{2}$ *and* $\overline{q} < 1$ *such that the agent under-trusts the source and is pessimistic that the state is A for* any *sequence with fixed* n_a, n_b *when* $0 \leq \hat{n}_b < n_b < n_a$ *and one of the following sufficient conditions is met: (a)* $\overline{q} \leq q_L < q_H$*, or (b)* $q_L \leq \underline{q}$ *and* $q_H > \overline{q}$.

When the proportion of a's is much greater than the proportion of b's (Part 1), a pre-screener exhibits confirmation bias in that she may always be optimistic about A relative to the Bayesian. This is true even for the sequence that generates the lowest possible trust in the source (fixing information content), so long as mixed signals do not sufficiently distinguish between high and low reliability (q_L and q_H sufficiently low). Here, the initial negative impression from mixed signals is relatively weak, so the ensuing consistency of many a's countervails the initial under-trust, creating optimism and overtrust.

When the proportion of a's is sufficiently similar to the proportion of b's (Part 2), this information content strongly suggests that the source is unreliable, as long as q_H is sufficiently high. In this case, the pre-screener may be under-trusting and pessimistic given *any* observed order, including the sequence that generates the highest degree of trust. This is because ensuing contrary signals are extremely informative of low reliability and therefore overcome even the most positive first impression, resulting in under-trust and pessimism. In contrast, on average an agent in Rabin and Schrag (1999) would exhibit optimism toward state A in both of the cases described in Proposition 8.

What resolves the persistent disagreement created by first impressions? Suppose the pre-screener begins with a neutral prior on the state, and observes k identical a signals, resulting in a positive first impression. After an additional k identical b signals from the same source, she will have the correct posterior on the state, though not necessarily on the source's reliability. This is shown in Lemma 2. Intuitively, the pre-screener understands that all signals originate from the same source, so she realizes that $n_a = n_b = k$ is equivalent to having no new information about the state, even if she is incorrect about the reliability.

A more realistic situation is one in which an agent receives additional signals from another source, or a "second opinion." Suppose the pre-screener now receives signals from two independently drawn sources, $j = 1, 2$, with credibilities c_j, where she encounters source 1 before source 2. Let s_{tj} be the signal sent in period t by source $j \in \{1, 2\}$, who sends a sequence of n_j signals, $\mathbf{s}^{n_j}$. Let $\mathbf{s}_{n_1,n_2}$ be the sequence of observed signals from both sources, where $\mathbf{s}_{n_1,n_2} = \{\mathbf{s}_{n_1}, \mathbf{s}_{n_2}\}$. Let $\mathbf{s}_{n_1,0}$ denote the sequence of signals from source 1 when agent has not yet seen signals from source 2.

The pre-screener now faces three uncertainties—the credibility of each of the two sources and the state of the world. Since source credibility is independent and identically distributed, $\omega_0^{c_j} = \omega_0^c$ for all j. Since source j's signal is independent of source k's signal, $P(s_{tj}|c_j, c_k, \theta) = P(s_{tj}|c_j, \theta)$ for all t where $j \neq k$.

The pre-screening procedure extends naturally from one to multiple sources. First, a pre-screener updates on the joint belief about source credibilities, denoted $\kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2})$, using Bayes' Rule. Second, she uses this updated first-stage belief $\kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2})$ to form a joint posterior beliefs on the state and credibilities. A Bayesian's final posterior beliefs depend purely on the information content of signals from each source, and depend neither on the order in which signals are received from a given source, nor on the order in which she encounters sources.

Iterating on the pre-screener's updating process allows us to characterize posterior beliefs when she receives any set of signals from both sources, $\mathbf{s}_{n_1,n_2}$. The pre-screener's beliefs after observing only source 1 are:

$$\kappa_{c_1c_2}(\mathbf{s}_{n_1,0}) = \frac{\kappa_{c_1c_2}(\mathbf{s}_{n_1-1,0}) \left(\sum_\theta \left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \omega_0^\theta\right)}{\sum_{c_1}\sum_{c_2} \kappa_{c_1c_2}(\mathbf{s}_{n_1-1,0}) \left(\sum_\theta \left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \omega_0^\theta\right)}, \tag{25}$$

where $\kappa_{c_1c_2}(\emptyset) = \omega_0^{c_1}\omega_0^{c_2}$, and:

$$P^b(c_1, \theta|\mathbf{s}_{n_1,0}) = \frac{\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \kappa_{c_1c_2}(\mathbf{s}_{n_1,0})\omega_0^\theta}{\sum_c\sum_\theta \left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \kappa_{c_1c_2}(\mathbf{s}_{n_1,0})\omega_0^\theta}. \tag{26}$$

After observing source 2, her beliefs are:

$$\kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2}) = \frac{\left(\sum_\theta \left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right) \left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \omega_0^\theta\right) \kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2-1})}{\sum_{c_2}\sum_{c_1} \left(\sum_\theta \left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right) \left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right) \omega_0^\theta\right) \kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2-1})}, \tag{27}$$

and:

$$P^b(c_1,c_2,\theta|\mathbf{s}_{n_1,n_2}) = \frac{\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2})\omega_0^\theta}{\sum_{c_2}\sum_{c_1}\sum_\theta \left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\kappa_{c_1c_2}(\mathbf{s}_{n_1,n_2})\omega_0^\theta}. \tag{28}$$

Suppose a pre-screener has a positive first impression about source 1 because source 1 reported $k > 1$ identical signals of a. Suppose source 2 reports k identical signals of b. Ex-ante in priors, source 1 and source 2 have the same credibility. Proposition 9 shows that source 2 does not resolve disagreement between a pre-screener and a Bayesian who observes the same signals. Proposition 9 in the Internet Appendix shows that the answer is, surprisingly, no.

The positive impression from the first source's consistency inflates the pre-screener's trust in the early source and deflates trust in the later source relative to the Bayesian inference. Like the Bayesian, the pre-screener concludes that the sources cannot both be highly reliable, but incorrectly concludes that the first source is more credible than the second and therefore differentially weights information in favor of the first. The reason for this backfire effect is that her interpretation of the second source's consistent signals are biased by the fact that the signals contradict the overtrusted first source. This asymmetry means that the second source cannot completely unravel the first source's signals.

Furthermore, this asymmetry persists in the limit: Information that should lead to more uncertainty about credibilities and no change in beliefs about the state instead leads the pre-screener to be *more sure of and more wrong in* her beliefs along both dimensions when observing opposing information from different sources sequentially.

Proposition 9 (Backfire effect) *Let* $(\omega_0^A, \omega_0^{H_1}, \omega_0^{H_2}) = (1/2, \hat{\omega}, \hat{\omega})$ *for any* $\hat{\omega} \in (0,1)$ *where sources 1 and 2 are independent. Let the agent observe* k a *signals from source 1, followed by* k b *signals from source 2:* $\mathbf{s}_{n_1} = (a,\ldots,a)$ *and* $\mathbf{s}_{n_2} = (b,\ldots,b)$ *where* $n_1 = n_2 = k$ *and* $k > 1$.

1. *The pre-screener believes that state* A *is more likely than* B*, and that the first source is more likely to be high-reliability than the second source:* $P^b(\theta = A|\mathbf{s}_{n_1,n_2}) > 1/2$ *and* $P^b(H_1|\mathbf{s}_{n_1,n_2}) > P^b(H_2|\mathbf{s}_{n_1,n_2})$.

2. *Persistence in the limit:* $\lim_{k\to\infty} P^b(\theta = A|\mathbf{s}_{n_1,n_2}) = 1$ *and* $\lim_{k\to\infty} P^b(H_1, L_2|\mathbf{s}_{n_1,n_2}) = 1$.

Proposition 10 shows that unlike in the case of unknown reliability alone, the presence of slant can lead the pre-screener's beliefs to be sufficiently wrong that she becomes certain of the state after observing a set of signals from a single expert, even though the evidence should not objectively change beliefs about the state from priors. To illustrate, consider the particularly stark case in which a single expert sends k identical a signals, followed by k

identical b signals. Suppose that there are two possible expert types ($c \in \{v, w\}$), where reliability is identical and the degree of slant is symmetric ($\alpha = \gamma \geq 0$): $v = (q, \alpha, 0)$, $w = (q, 0, \gamma)$. Assume that each state and each type is equally likely ex-ante: $\omega_0^A = 1/2$ and $\omega_0^c = 1/2$. Clearly, the Bayesian realizes that there is no new information about either the state or the expert's slant. In contrast, the pre-screener can strongly believe that these signals reveal the state:

Proposition 10 (Disagreement when slant is uncertain) *Let the agent observe path $\mathbf{s}_{2k}$, which is k a signals followed by k b signals from a single expert. Let $(\omega_0^A, \omega_0^c) = (1/2, 1/2)$ and $\omega_0^c = 1/2$, where $c \in \{v, w\}$. Let $v = (q, \alpha, 0)$, $w = (q, 0, \gamma)$ where $q \in (1/2, 1)$ and $\alpha = \gamma \in (0, 1)$. Then $P^u(A, w|\mathbf{s}_{2k}) = P^u(B, v|\mathbf{s}_{2k}) = 1/2 \ \forall k$, but*

1. *If $q \leq 3/4$, $\lim_{k\to\infty} P^b(B, v|\mathbf{s}_{2k}) = 1$;*

2. *If $q > 3/4$, $\exists$ a unique $\bar{\alpha}(q) \in (0, 1)$ such that:*

   (a) *$\forall\ \alpha > \bar{\alpha}(q), \lim_{k\to\infty} P^b(B, v|\mathbf{s}_{2k}) = 1$,*

   (b) *$\forall\ \alpha < \bar{\alpha}(q), \lim_{k\to\infty} P^b(A, w|\mathbf{s}_{2k}) = 1$,*

   (c) *If $\alpha = \bar{\alpha}(q)$, then $\lim_{k\to\infty}[P^b(A, w|\mathbf{s}_{2k}) + P^b(B, v|\mathbf{s}_{2k})] = 1$ and $P^b(B, v|\mathbf{s}_{2k}) > 1/2$.*

Key insights from Section 4.1 qualitatively hold even in the presence of expert slant in the relevant cases. For example, in reality, an agent who encounters experts with contrary reports to one another may be especially concerned about potential slant in addition to reliability. In Proposition 11, we show that uncertainty about reliability generates disagreement when contrary experts are encountered in different order, irrespective of uncertainty about slant.

Proposition 11 (Backfire effect when reliability and slant are uncertain) *Let the agent observe k a signals from expert 1, followed by k b signals from expert 2: $\mathbf{s}_{n_1} = (a, \ldots, a)$ and $\mathbf{s}_{n_2} = (b, \ldots, b)$ where $n_1 = n_2 = k$ and $k > 1$. Suppose there are four expert types ($c \in \{v, w, x, z\}$), where $\alpha = \gamma \geq 0$ and $q_H \geq q_L \geq 1/2$: $v = (q_H, \alpha, 0)$, $w = (q_L, \alpha, 0)$, $x = (q_H, 0, \gamma)$, $w = (q_L, 0, \gamma)$. Let $\omega_0^A = 1/2$ and $\omega_0^c = 1/4$.*

Then $P^u(A|\mathbf{s}_{n_1,n_2}) = 1/2 \ \forall k$, but

1. *$\lim_{k\to\infty} P^b(A|\mathbf{s}_{n_1,n_2}) = 1/2$ if $q_H = q_L$ and $\alpha = \gamma \geq 0$,*

2. *$\lim_{k\to\infty} P^b(A|\mathbf{s}_{n_1,n_2}) = 1$ if $q_H > q_L$ and $\alpha = \gamma \geq 0$,*

Intuitively, if reliability is fixed and known ($q_H = q_L$), a pre-screener thinks that an initial expert who reports all a's is slanted towards A, and that a second expert who reports all b's is slanted towards B. Even though the pre-screener overweights the belief that the two experts are slanted, they are symmetric and cancel, so that the pre-screener has the correct neutral posterior on the state.[16] However, if reliability is also unknown ($q_H \neq q_L$), the pre-screener deduces that the first expert is correct and that the true state is A. The reason is that the initial streak of a's from the first expert indicates both A-slant *and* high reliability. The pre-screener overweights this, leading her to deduce from the second expert's streak of b's that the second expert is both B-slanted and low reliability using similar reasoning to Proposition 9. Thus, even with uncertainty about expert slant, opposing information from sequential experts leads the pre-screener to be more sure of and more wrong in her beliefs.

Proof of Proposition 8

1. **Lemma 8** *Suppose the agent observes $n_a = n_b$ signals of* a*'s and* b*'s in alternating order:* $\mathbf{s}_n = (a, b, \ldots a, b)$ *where* $n_a = n_b = k$*. Then the pre-screener always undertrusts the signal source.*

   Proof. An alternating sequence of $n_a = n_b = k$ signals of a's and b's generates:

   $$\beta_c(\mathbf{s}_n) = \left(\frac{1}{2}\right)^k \left(\prod_{i=1}^{k-1} (q_c(1-q_c))^{2i}\right) (q_c(1-q_c))^k = \left(\frac{1}{2}\right)^k (q_c(1-q_c))^{k^2}.$$

   This implies that $\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) < 0$ for all q_c:

   $$\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) = \left(\frac{1}{2}\right)^k k^2 \left(q_c(1-q_c)\right)^{k^2-1} (1-2q_c),$$

   Since $(q_H(1-q_H))^{k^2} < (q_L(1-q_L))^{k^2}$ whenever $q_H > q_L \geq \frac{1}{2}$ or $\frac{1}{2} \geq q_L > 1 - q_H$, then $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$ for all $q_H > q_L \geq \frac{1}{2}$, which implies that the pre-screener's belief that the source is high reliability is underconfident relative to the Bayesian: $P^b(H|\mathbf{s}_n) < P^u(H|\mathbf{s}_n)$. ■

   Suppose the agent observes $n_a > n_b$ signals, where n_b a's and n_b b's alternate followed by the remaining $m \equiv n_a - n_b$ a's where $m \geq 1$: $\mathbf{s}_n = (a, b, a, b, \ldots, a, a, a)$. From Equations (9) and (10), we can see that $\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) = 0$ when $q_c \in \{\frac{1}{2}, 1\}$, $\beta_c(\mathbf{s}_n) > 0$ when $q_c = \frac{1}{2}$, and $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1$. Moreover, using the fact that $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1/2$, then

   $$\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\bigg|_{q_c=\frac{1}{2}} = \beta_c(\mathbf{s}_n)\left(-8n_b(n_b+m) + \textstyle\sum_{i=1}^m 4i(i-1)\right) = \beta_c(\mathbf{s}_n)\left(-8n_b(n_b+m) + \frac{4}{3}m(m-1)(m+1)\right).$$

   Thus for any given n_b, there exists some threshold $\frac{1}{2} < \check{q} < 1$ whenever $m > m^*$, where $-8n_b(n_b + m^*) + \frac{4}{3}m^*(m^*-1)(m^*+1) = 0$. Let $n_a^* = n_b + m^*$. Then for n_a, n_b

[16] This differs from the single expert case discussed in Proposition 10. There, the pre-screener had to determine the slant of a single expert who contradicted herself with a streak of a's and a streak of b's. Here, each expert is reporting an internally consistent streak.

where $0 \geq n_b < n_a^* < n_a$ and $q_L < q_H \leq \check{q}$, the agent overtrusts and is optimistic that the state is A. Since this is the sequence that generates the least trust by Proposition 2, then if it results in overtrust then all other sequences of such a combination must generate overtrust and optimism as well.

2. **Lemma 9** *After observing $n_a > 1$ and $n_b = 0$ signals in sequence or simultaneously, the pre-screener overtrusts and is overoptimistic about the reported state.*

   Proof. Without loss of generality, suppose the sequence is n_a a's: $\mathbf{s}_n = (a, a, \ldots, a)$ where $n_a = n$ and $n_b = 0$. Then $\beta_c(\mathbf{s}_n) = \prod_{i=1}^{n_a}(\frac{1}{2})(q_c^i + (1-q_c)^i)$. Considering each ith component of $\beta_c(\mathbf{s}_n)$, $q_H^i + (1-q_H)^i > q_L^i + (1-q_L)^i$ is positive for $i > 0$ when $q_H > q_L \geq \frac{1}{2}$ or when $\frac{1}{2} \geq q_L > 1 - q_H$, which implies that $\beta_H(s_{11} = a, s_{22} = a, \ldots s_{n_a,n_a} = a) > \beta_L(s_{11} = a, s_{22} = a, \ldots s_{n_a,n_a} = a)$. Thus, applying Proposition 4, the pre-screener overtrusts and is overoptimistic about the reported state when she observes $n_a > 1$ and $n_b = 0$ signals in sequence. Since the simultaneous case implies $\beta_c(\mathbf{s}_n) = q_c^{n_a} + (1-q_c)^{n_a}$, then this argument also shows the result when the pre-screener observes $n_a > 1$ signals simultaneously. ■

Lemma 10 *Consider a sequence of signals such that the first k observed signals are a, followed by k b signals: $\mathbf{s}_n = (a, a, \ldots, a, b, b, \ldots, b)$ where $n_a = n_b = k$. There exists some $\underline{q} > \frac{1}{2}$ and $\overline{q} < 1$ such that the pre-screener under-trusts if (1) $k \in \{1, 2, 3\}$, (2) if $\overline{q} \leq q_L < q_H$, or (3) if $q_L \leq \underline{q}$ and $q_H \geq \overline{q}$.*

Proof. WLOG, suppose the sequence is n_a a's, then n_b b's. Then

$$\beta_c(\mathbf{s}_n) = \left(\prod_{i=1}^{n_a}(\frac{1}{2})(q_c^i + (1-q_c)^i)\right)\left(\prod_{i=1}^{n_b}(\frac{1}{2})(q_c^{n_a}(1-q_c)^i + q_c^i(1-q_c)^{n_a})\right)$$

In particular, if the sequence is $n_a = n_b = k$, then:

$$\beta_c(\mathbf{s}_n) = (\frac{1}{2})^{2k}\prod_{i=1}^{k}\left(q_c^i + (1-q_c)^i\right)\left(q_c^k(1-q_c)^i + q_c^i(1-q_c)^k\right) \tag{29}$$

Characterizing $\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n))$ when $n_a = n_b = k$ yields

$$\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) = (\frac{1}{2})^{2k}\left(\prod_{i=1}^{k}\left(q_c^i + (1-q_c)^i\right)\left(q_c^k(1-q_c)^i + q_c^i(1-q_c)^k\right)\right)$$
$$\left(\sum_{i=1}^{k}\frac{i(q_c^{i-1} - (1-q_c)^{i-1})(q_c^k(1-q_c)^i + q_c^i(1-q_c)^k) + \left(q_c^i + (1-q_c)^i\right)\left(k\left(q_c^{k-1}(1-q_c)^i - q_c^i(1-q_c)^{k-1}\right) + i\left(q_c^{i-1}(1-q_c)^k - q_c^k(1-q_c)^{i-1}\right)\right)}{(q_c^i + (1-q_c)^i)\left(q_c^k(1-q_c)^i + q_c^i(1-q_c)^k\right)}\right) \tag{30}$$

From Equation (29) we can see that $\frac{\partial}{\partial q_c}(\beta_c(\mathbf{s}_n)) = 0$ when $q_c \in \{\frac{1}{2}, 1\}$, $\beta_c(\mathbf{s}_n) > 0$ when $q_c = \frac{1}{2}$, and $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1$. Moreover, using the fact that $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1/2$, then

$$\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}} = \beta_c(\mathbf{s}_n)\left(\sum_{i=1}^{k} 4(2i(i-1) + k(k-1) - 2ki)\right) = \beta_c(\mathbf{s}_n)\left(\frac{8}{3}k(-3k + k^2 - 1)\right),$$

so $\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}}$ is negative when $k < \frac{3+\sqrt{13}}{2} \approx 3.3028$ and positive when $k > \frac{3+\sqrt{13}}{2}$.

Since $\beta_c(\mathbf{s}_n) = 0$ when $q_c = 1$, $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c}|_{q_c=1} = 0$, and $\beta_c(\mathbf{s}_n) \geq 0$ for any $q_c \in [0,1]$, then there exists some threshold $\overline{q} < 1$ such that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} < 0$ and $\beta_c(\mathbf{s}_n) < \beta_c(\mathbf{s}_n)\Big|_{q_c=\frac{1}{2}}$ for all $q_c > \overline{q}$.

Since $\beta_c(\mathbf{s}_n) > 0$ when $q_c = \frac{1}{2}$, $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c}\Big|_{q_c=\frac{1}{2}} = 0$, and $\frac{\partial^2 \beta_c(\mathbf{s}_n)}{\partial q_c^2}\Big|_{q_c=\frac{1}{2}} > 0$ when $k > \frac{3+\sqrt{13}}{2}$, then there exists some threshold $\underline{q} > \frac{1}{2}$ such that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} > 0$ and $\beta_c(\mathbf{s}_n) > \beta_c(\mathbf{s}_n)\Big|_{q_c=\frac{1}{2}}$ for all $q_c < \underline{q}$ when $k > \frac{3+\sqrt{13}}{2}$. When $k \leq \frac{3+\sqrt{13}}{2}$, we can show by direct computation of $\beta_c(\mathbf{s}_n)$ that $\frac{\partial \beta_c(\mathbf{s}_n)}{\partial q_c} < 0$ for all $q_c \in (\frac{1}{2}, 1)$ when $k \in \{1, 2, 3\}$. This implies that the pre-screener under-trusts for all values of $q_L < q_H$ whenever $k \leq 3$, since $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$. When $k > 3$, there are two other sufficient conditions for the pre-screener to under-trust: (1) if $\overline{q} \leq q_L < q_H$, or (2) if $q_L \leq \underline{q}$ and $q_H > \overline{q}$ where $\underline{q} > \frac{1}{2}$ and $\overline{q} < 1$. If either of these sufficient conditions is met, then $\beta_H(\mathbf{s}_n) < \beta_L(\mathbf{s}_n)$ for $k > 3$. ■

Lemma 9 shows that the agent overtrusts and is overoptimistic about the reported state for a given $n_a > 1$ and $n_b = 0$. Clearly, the agent's degree of overtrust is monotonically decreasing as n_b increases. Lemma 10 shows that there exists some $\underline{q} > \frac{1}{2}$ and $\overline{q} < 1$ such that the pre-screener under-trusts if (1) $k \in \{1, 2, 3\}$, (2) if $\overline{q} \leq q_L < q_H$, or (3) if $q_L \leq \underline{q}$ and $q_H \geq \overline{q}$. By the intermediate value theorem, there exists some $\hat{n_b}$ such that the agent under-trusts when $\mathbf{s}_n = (a, a, \ldots, a, b, b, \ldots, b)$ where $0 \leq \hat{n_b} \leq n_b < n_a$. By Proposition 2, this is the sequence most likely to generate overtrust, so *all other sequences* of such a fixed combination (n_a, n_b) will also result in under-trust and pessimism about the mostly likely state. Thus, if one of the last two sufficient conditions for Lemma 10 is satisfied, then there exists some $\hat{n_b}$ such that the agent under-trusts when $\mathbf{s}_n = (a, a, \ldots, a, b, b, \ldots, b)$ where $0 \leq \hat{n_b} \leq n_b < n_a$.

Proof of Proposition 9

1. To show the results when agents receive signals from multiple sources, note that Equation (28) can also be re-written as

$$P^b(c_1,c_2,\theta|\mathbf{s}_{n_1,n_2}) = \frac{\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{c_1}\omega_0^{c_2}\omega_0^{\theta}\beta_{c_1}(\mathbf{s}_{n_1})\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2})}{\sum_{c_2}\sum_{c_1}\sum_{\theta}\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{c_1}\omega_0^{c_2}\omega_0^{\theta}\beta_{c_1}(\mathbf{s}_{n_1})\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2})}, \tag{31}$$

where the functions $\beta_{c_1}(\mathbf{s}_{n_1})$ and $\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2})$ reflect the path dependency of the pre-screener's beliefs and $\beta_{c_1}(\emptyset) = 1$:

$$\beta_{c_1}(\mathbf{s}_{n_1}) = \prod_{m=1}^{n_1}\left(\sum_{\theta}\left(\prod_{t=1}^{m} P(s_{t1}|c_1,\theta)\right)\omega_0^{\theta}\right)$$

$$\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2}) = \prod_{m=n_1+1}^{n_1+n_2}\left(\sum_{\theta}\left(\prod_{t=n_1+1}^{m} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{\theta}\right).$$

Consider a sequence of signals such that the agent observes k a signals from source 1, followed by k b signals from source 2: $\mathbf{s}_{n_1} = (a,\ldots,a)$ and $\mathbf{s}_{n_2} = (b,\ldots,b)$ where $n_1 = n_2 = k$.

To show this, note that the following properties hold when $\omega_0^{\theta} = 1/2$ and the two sources send either (1) an equal number k of opposing signals, or (2) an equal number of completely mixed signals: $\prod_{i=k+1}^{2k} P(s_{i2}|H_2,A) = \prod_{i=1}^{k} P(s_{i1}|H_1,B)$, $\prod_{i=k+1}^{2k} P(s_{i2}|L_2,A) = \prod_{i=1}^{k} P(s_{i1}|L_1,B)$, $\prod_{i=k+1}^{2k} P(s_{i2}|H_2,B) = \prod_{i=1}^{k} P(s_{i1}|H_1,A)$, and $\prod_{i=k+1}^{2k} P(s_{i2}|L_2,B) = \prod_{i=1}^{k} P(s_{i1}|L_1,A)$.

For all $\omega_0^{\theta} \in (0,1)$, then $P^b(\theta|\mathbf{s}_{n_1,n_2}) > 1/2$ only if

$$\omega_0^H(1-\omega_0^H)\left(\left(\prod_{i=k+1}^{2k} P(s_{i2}|H_2,A)\right)\left(\prod_{i=k+1}^{2k} P(s_{i2}|L_2,B)\right) - \left(\prod_{i=k+1}^{2k} P(s_{i2}|L_2,A)\right)\left(\prod_{i=k+1}^{2k} P(s_{i2}|H_2,B)\right)\right)$$
$$(\beta_{L_1}(\mathbf{s}_n)\beta_{H_2L_1}(\mathbf{s}_n) - \beta_{H_1}(\mathbf{s}_n)\beta_{L_2H_1}(\mathbf{s}_n)) > 0. \tag{32}$$

When the two sources send an equal number of opposing signals in sequence (and suppressing the arguments of $\beta_{c_1}(\mathbf{s}_n)$ and $\beta_{c_1c_2}(\mathbf{s}_n)$ for brevity of exposition), we also know

$\prod_{i=k+1}^{2k} P(s_{i2}|H_2,A) = \prod_{i=1}^{k} P(s_{i1}|H_1,B) = (1-q_H)^k$
$\prod_{i=k+1}^{2k} P(s_{i2}|L_2,A) = \prod_{i=1}^{k} P(s_{i1}|L_1,B) = (1-q_L)^k$
$\prod_{i=k+1}^{2k} P(s_{i2}|H_2,B) = \prod_{i=1}^{k} P(s_{i1}|H_1,A) = q_H^k$
$\prod_{i=k+1}^{2k} P(s_{i2}|L_2,B) = \prod_{i=1}^{k} P(s_{i1}|L_1,A) = q_L^k$

$\beta_{c_1} = \prod_{i=1}^{k}(\frac{1}{2})(q_{c_1}^i + (1-q_{c_1})^i)$, where we have previously shown that $\beta_{H_1} > \beta_{L_1}$

$\beta_{c_2c_1} = \prod_{i=1}^{k}(\frac{1}{2})\left((1-q_{c_2})^i q_{c_1}^k + q_{c_2}^i(1-q_{c_1})^k\right)$

Substituting all of these into the pre-screener's posterior on the state, $P^b(\theta|\mathbf{s}_{n_1,n_2}) > 1/2$ only if

$$\omega_0^H(1-\omega_0^H)\left((1-q_H)^k q_L^k - (1-q_L)^k q_H^k\right)(\beta_{L_1}\beta_{H_2L_1} - \beta_{H_1}\beta_{L_2H_1}) > 0. \quad (33)$$

The first term of Equation (33) is positive and the second term is clearly negative, since $q_H > q_L$. Note that $\beta_L < \beta_H$ for $n_a > 1$ and $\beta_L = \beta_H$ for $n_a = 1$. Comparing a given mth term of $\beta_{L_2H_1} - \beta_{H_2L_1}$ yields

$$\begin{aligned}&(\frac{1}{2})\left((1-q_L)^m q_H^k + q_L^m(1-q_H)^k - (1-q_H)^m q_L^k - q_H^m(1-q_L)^k\right)\\ &\quad = (\frac{1}{2})\left(q_H^m(1-q_L)^m(q_H^{k-m} - (1-q_L)^{k-m}) + q_L^m(1-q_H)^m((1-q_H)^{k-m} - q_L^{k-m})\right),\end{aligned}$$

which is zero if $k = m$ and positive if $m < k$. Thus, each mth term of $\beta_{L_2H_1}$ is strictly greater than the mth term of $\beta_{H_2L_1}$ for $m < k$ and is equal when $m = k$, implying that $\beta_{L_2H_1} > \beta_{H_2L_1}$ if $k > 1$ (and $\beta_{L_2H_1} = \beta_{H_2L_1}$ if $k = 1$) . This implies that the third term of Equation (33) is strictly negative when $k > 1$, so Equation (33) is satisfied. Thus, $P^b(\theta = A|\mathbf{s}_{n_1,n_2}) > 1/2$ when $\omega_0^A = 1/2$ and $k > 1$, and $P^b(\theta = A|\mathbf{s}_{n_1,n_2}) = 1/2$ when $\omega_0^A = 1/2$ and $k = 1$.

Substituting all of these into the pre-screener's posteriors on source credibilities, we have that $P^b(c_1|\mathbf{s}_{n_1,n_2}) > P^b(c_2|\mathbf{s}_{n_1,n_2})$ only if $\omega_0^H(1-\omega_0^H)\left((1-q_H)^k q_L^k - (1-q_L)^k q_H^k\right)(\beta_{L_1}\beta_{H_2L_1} - \beta_{H_1}\beta_{L_2H_1}) > 0$, which is exactly Equation (33) again. Thus, the pre-screener believes that the first source is more likely to be high reliability than the second source: $P^b(c_1|\mathbf{s}_{n_1,n_2}) > P^b(c_2|\mathbf{s}_{n_1,n_2})$.

2. The pre-screener's joint posteriors on sources' credibilities are of the form $P^b(H_1, H_2|s_{1j}, s_{2j}) = \frac{w}{w+x+y+z}$, $P^b(L_1, H_2|s_{1j}, s_{2j}) = \frac{x}{w+x+y+z}$, $P^b(H_1, L_2|s_{1j}, s_{2j}) = \frac{y}{w+x+y+z}$, and $P^b(H_1, H_2|s_{1j}, s_{2j}) = \frac{z}{w+x+y+z}$, where

$$w = 2(\omega_0^H)^2(1-q_H)^k q_H^{2k(k+1)}\left(\prod_{m=1}^{k}(1+(\frac{1-q_H}{q_H})^m)((\frac{1-q_H}{q_H})^m + (\frac{1-q_H}{q_H})^k)\right)$$

$$x = \omega_0^H(1-\omega_0^H)\left((1-q_H)^k q_L^k + q_H^k(1-q_L)^k\right) q_L^{\frac{k(k+1)}{2}} q_H^{\frac{k(k+1)}{2}} q_L^{k^2}\left(\prod_{m=1}^{k}(1+(\frac{1-q_L}{q_L})^m)((\frac{1-q_H}{q_H})^m + (\frac{1-q_L}{q_L})^k)\right)$$

$$y = \omega_0^H(1-\omega_0^H)\left((1-q_L)^k q_H^k + q_L^k(1-q_H)^k\right) q_H^{\frac{k(k+1)}{2}} q_L^{\frac{k(k+1)}{2}} q_H^{k^2}\left(\prod_{m=1}^{k}(1+(\frac{1-q_H}{q_H})^m)((\frac{1-q_L}{q_L})^m + (\frac{1-q_H}{q_H})^k)\right)$$

$$z = 2(1-\omega_0^H)^2(1-q_L)^k q_L^{2k(k+1)}\left(\prod_{m=1}^{k}(1+(\frac{1-q_L}{q_L})^m)((\frac{1-q_L}{q_L})^m + (\frac{1-q_L}{q_L})^k)\right).$$

Letting $k \to \infty$ and factoring, we can re-write the terms w, x, y, and z as

$$
\begin{aligned}
w &= 2(\omega_0^H)^2(1-q_H)^k q_H^{2k(k+1)}(\frac{1-q_H}{q_H})^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(2(\omega_0^H)^2(\frac{1-q_H}{1-q_L})^{k+\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)\right) \\
x &= \omega_0^H(1-\omega_0^H)\left((1-q_H)^k q_L^k + q_H^k(1-q_L)^k\right) q_L^{\frac{k(k+1)}{2}} q_H^{\frac{k(k+1)}{2}} q_L^{k^2}(\frac{1-q_H}{q_H})^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_L}{q_L})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(\omega_0^H(1-\omega_0^H)\left(1+(\tfrac{q_L(1-q_H)}{q_H(1-q_L)})^k\right)\left(\tfrac{1-q_H}{1-q_L}\right)^{\frac{k(k+1)}{2}}(\tfrac{q_L}{q_H})^{k^2+\frac{k(k+1)}{2}}\left(\prod_{m=1}^{\infty}(1+(\tfrac{1-q_L}{q_L})^m)\right)\right) \\
y &= \omega_0^H(1-\omega_0^H)\left((1-q_L)^k q_H^k + q_L^k(1-q_H)^k\right) q_H^{\frac{k(k+1)}{2}} q_L^{\frac{k(k+1)}{2}} q_H^{k^2}(\frac{1-q_L}{q_L})^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(\omega_0^H(1-\omega_0^H)\left(1+\left(\frac{q_L(1-q_H)}{q_H(1-q_L)}\right)^k\right)\left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)\right) \\
z &= 2(1-\omega_0^H)^2(1-q_L)^k q_L^{2k(k+1)}(\frac{1-q_L}{q_L})^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_L}{q_L})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(2(1-\omega_0^H)^2\left(\frac{q_L}{q_H}\right)^{\frac{3k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_L}{q_L})^m)\right)\right)
\end{aligned}
$$

Note that the term $q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}}$ drops out since it is in every term when calculating the joint posteriors. Also, note that a necessary and sufficient condition for $\prod_{m=1}^{\infty}(1+(\frac{1-p_q}{p_q})^m)$ to converge is that $\sum_{m=1}^{\infty}(\frac{1-p_q}{p_q})^m$ is absolutely convergent, which is clearly satisfied when $p_q > \frac{1}{2}$. Thus, when $1 > q_H > q_L > \frac{1}{2}$, $\lim_{k\to\infty} w = 0, \lim_{k\to\infty} x = 0, \lim_{k\to\infty} y = \omega_0^H(1-\omega_0^H)\left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right), \lim_{k\to\infty} z = 0$. This implies that when $1 > q_H > q_L > \frac{1}{2}$, $\lim_{k\to\infty} P^b(H_1, H_2|\mathbf{s}_{n_1,n_2}) = 0$, $\lim_{k\to\infty} P^b(L_1, H_2|\mathbf{s}_{n_1,n_2}) = 0$, $\lim_{k\to\infty} P^b(H_1, L_2|\mathbf{s}_{n_1,n_2}) = 1$, $\lim_{k\to\infty} P^b(L_1, L_2|\mathbf{s}_{n_1,n_2}) = 0$.

The result also holds when $q_L = \frac{1}{2}$. Letting $k \to \infty$ and $q_L = \frac{1}{2}$ and factoring, we can re-write the terms w, x, y, and z as

$$
\begin{aligned}
w &= 2(\omega_0^H)^2(1-q_H)^k q_H^{2k(k+1)}(\frac{1-q_H}{q_H})^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= 2(\omega_0^H)^2(1-q_H)^{k+\frac{k(k+1)}{2}} q_H^{\frac{3k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(2(\omega_0^H)^2(\frac{1-q_H}{1-q_L})^{k+\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)\right) \\
x &= \omega_0^H(1-\omega_0^H)(\frac{1}{2})^k \left((1-q_H)^k + q_H^k\right)(\frac{1}{2})^{k^2+\frac{k(k+1)}{2}} q_H^{\frac{k(k+1)}{2}}(2)^k \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= \omega_0^H(1-\omega_0^H)(\frac{1}{2})^{k^2+\frac{k(k+1)}{2}} \left((1-q_H)^k + q_H^k\right) q_H^{\frac{k(k+1)}{2}} \left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(\omega_0^H(1-\omega_0^H)\left(\tfrac{2}{(2q_H)^k}\right)^k (1+(\tfrac{1-q_H}{q_H})^k)\left(\textstyle\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)\right) \\
y &= \omega_0^H(1-\omega_0^H)(\tfrac{1}{2})^{k+\frac{k(k+1)}{2}} q_H^{k^2+\frac{k(k+1)}{2}}(q_H^k+(1-q_H)^k)\left(1+(\tfrac{1-q_H}{q_H})^k\right)^k \left(\textstyle\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right) \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(\omega_0^H(1-\omega_0^H)(1+(\tfrac{1-q_H}{q_H})^k)(1+(\tfrac{1-q_H}{q_H})^k)^k \left(\textstyle\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)\right) \\
z &= 2(1-\omega_0^H)^2(\frac{1}{2})^{2k^2+3k}(2)^{2k} = 2(1-\omega_0^H)^2(\frac{1}{2})^{2k^2+k} \\
&= q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}} \left(2(1-\omega_0^H)^2(\frac{1}{2q_H})^{\frac{3k(k+1)}{2}}(\frac{1}{2})^{\frac{k(k-1)}{2}}\right)
\end{aligned}
$$

The term $q_H^{\frac{3k(k+1)}{2}}(1-q_L)^{k+\frac{k(k+1)}{2}}$ drops out since it is in every term when calculating the joint posteriors. A necessary and sufficient condition for $\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)$ to converge is that $\sum_{m=1}^{\infty}(\frac{1-q_H}{q_H})^m$ is absolutely convergent, which is clearly satisfied when $q_H > \frac{1}{2}$.

Terms w, x, and z converge to 0. Term y converges to $\omega_0^H(1-\omega_0^H)\left(\prod_{m=1}^{\infty}(1+(\frac{1-q_H}{q_H})^m)\right)$, which is a finite number, because $\lim_{k\to\infty}(1+(\frac{1-q_H}{q_H})^k)^k = 1$ (re-arranging and using

L'Hopital's rule several times):

$$\lim_{k\to\infty}(1+(\frac{1-q_H}{q_H})^k)^k = \lim_{k\to\infty}\left(\exp\left(\ln(1+(\frac{1-q_H}{q_H})^k)\right)^k\right) = \exp\lim_{k\to\infty}\left(k\ln(1+(\frac{1-q_H}{q_H})^k)\right)$$
$$= \exp\lim_{k\to\infty}\frac{\ln(1+(\frac{1-q_H}{q_H})^k)}{\frac{1}{k}} = \exp\lim_{k\to\infty}\frac{\frac{(\frac{1-q_H}{q_H})^k\ln(\frac{1-q_H}{q_H})}{1+(\frac{1-q_H}{q_H})^k}}{-(\frac{1}{k})^2}$$
$$= \exp\lim_{k\to\infty}(\ln(\frac{1-q_H}{q_H}))\left(\frac{-k^2}{\frac{1+(\frac{1-q_H}{q_H})^k}{(\frac{1-q_H}{q_H})^k}}\right) = \exp\lim_{k\to\infty}(\ln(\frac{1-q_H}{q_H}))\left(\frac{-2k}{-\frac{\ln(\frac{1-q_H}{q_H})}{(\frac{1-q_H}{q_H})^k}}\right)$$
$$= \exp\lim_{k\to\infty}2\left(\frac{k}{\frac{1}{(\frac{1-q_H}{q_H})^k}}\right) = \exp\lim_{k\to\infty}2\left(\frac{1}{-\frac{\ln(\frac{1-q_H}{q_H})}{(\frac{1-q_H}{q_H})^k}}\right) = \exp\lim_{k\to\infty}2\left(\frac{(\frac{1-q_H}{q_H})^k}{-\ln(\frac{1-q_H}{q_H})}\right)$$
$$= \exp(0) = 1.$$

This implies that when $1 > q_H > q_L = \frac{1}{2}$, $\lim_{k\to\infty} P^b(H_1, H_2|\mathbf{s}_{n_1,n_2}) = 0$, $\lim_{k\to\infty} P^b(L_1, H_2|\mathbf{s}_{n_1,n_2}) = 0$, $\lim_{k\to\infty} P^b(H_1, L_2|\mathbf{s}_{n_1,n_2}) = 1$, and $\lim_{k\to\infty} P^b(L_1, L_2|\mathbf{s}_{n_1,n_2}) = 0$. An extremely similar proof applies to show that $\lim_{k\to\infty} P^b(\theta = A|\mathbf{s}_{n_1,n_2}) = 1$ where $n_1 = n_2 = k$ when $1 > q_H > q_L = \frac{1}{2}$.

Proof of Proposition 10

Any combination of reliability and slant is summarized by the formulation that a source type c reports state A with probability $p_{ac} \equiv P(s_t = a|c, A)$ and she reports state B with probability $p_{bc} = P(s_t = b|c, B)$. A source type c is (p_{ac}, p_{bc}), which is fully defined by $(q_c, \alpha_c, \gamma_c)$ given our mapping.

The pre-screener's joint posterior beliefs on source types are given by $P^b(v, A|\mathbf{s}_{2k} = \{a, a, \dots, b, b\dots\}) = \frac{a}{a+b+c+d}$, $P^b(w, A|\mathbf{s}_{2k} = \{a, a, \dots, b, b\dots\}) = \frac{b}{a+b+c+d}$, $P^b(v, B|\mathbf{s}_{2k} = \{a, a, \dots, b, b\dots\}) = \frac{c}{a+b+c+d}$, $P^b(w, B|\mathbf{s}_{2k} = \{a, a, \dots, b, b\dots\}) = \frac{d}{a+b+c+d}$, where, using the fact that $p_{av} = p_{bw} > 1/2 > p_{aw} = p_{bv}$, we have

$$a = \beta_v[(p_{av}(1-p_{av}))^k]$$
$$= \left(\prod_{i=1}^{k}(p_{av}^i + (1-p_{bv})^i)(p_{av}^k(1-p_{av})^i + (1-p_{bv})^k p_{bv}^i)\right)[(p_{av}(1-p_{av}))^k]$$
$$= (p_{av}(1-p_{av}))^{\frac{k(k+1)}{2}}(p_{av})^{k^2}(p_{av}(1-p_{av}))^k\left(\prod_{i=1}^{k}\left((1+(\frac{1-p_{bv}}{p_{av}})^i)(1+(\frac{1-p_{bv}}{p_{av}})^k(\frac{p_{bv}}{1-p_{av}})^i)\right)\right)$$

$$
\begin{aligned}
b &= \beta_w[(p_{bv}(1-p_{bv}))^k] \\
&= \left(\prod_{i=1}^{k}\left((p_{av}^i + (1-p_{bv})^i)(p_{av}^k(1-p_{bv})^i + (1-p_{bv})^k p_{av}^i)\right)\right)[(1-p_{bv})p_{bv}]^k \\
&= (p_{bv}(1-p_{bv}))^{\frac{k(k+1)}{2}}(p_{bv})^{k^2}(p_{bv}(1-p_{bv}))^k\left(\prod_{i=1}^{k}\left((1+(\frac{1-p_{av}}{p_{bv}})^i)(1+(\frac{1-p_{av}}{p_{bv}})^{k-i})\right)\right) \\
c &= \beta_v[(p_{bv}(1-p_{bv}))^k] \\
&= \left(\prod_{i=1}^{k}(p_{av}^i + (1-p_{bv})^i)(p_{av}^k(1-p_{av})^i + (1-p_{bv})^k p_{bv}^i)\right)[(p_{bv}(1-p_{bv}))^k] \\
&= (p_{av}(1-p_{av}))^{\frac{k(k+1)}{2}}(p_{av})^{k^2}(p_{bv}(1-p_{bv}))^k\left(\prod_{i=1}^{k}\left((1+(\frac{1-p_{bv}}{p_{av}})^i)(1+(\frac{1-p_{bv}}{p_{av}})^k(\frac{p_{bv}}{1-p_{av}})^i)\right)\right) \\
d &= \beta_w[(p_{av}(1-p_{av}))^k] \\
&= \left(\prod_{i=1}^{k}\left((p_{av}^i + (1-p_{bv})^i)(p_{av}^k(1-p_{bv})^i + (1-p_{bv})^k p_{av}^i)\right)\right)[(p_{av}(1-p_{av}))^k] \\
&= (p_{bv}(1-p_{bv}))^{\frac{k(k+1)}{2}}(p_{bv})^{k^2}(p_{av}(1-p_{av}))^k\left(\prod_{i=1}^{k}\left((1+(\frac{1-p_{av}}{p_{bv}})^i)(1+(\frac{1-p_{av}}{p_{bv}})^{k-i})\right)\right).
\end{aligned}
$$

Moreover, we know that $p_{av} > 1-p_{bv} > 1/2 > p_{bv}$, $p_{bv} > 1-p_{av})$, and $p_{bv}(1-p_{bv}) > p_{av}(1-p_{av})$. Using this fact to factor out the terms $[p_{bv}(1-p_{bv})]^{\frac{k(k+1)}{2}}[p_{bv}(1-p_{bv})]^k p_{av}^{k^2}$ and take the limit, we have

$$
\begin{aligned}
\lim_{k\to\infty} a &= \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) \\
\lim_{k\to\infty} b &= \left(\frac{p_{bv}}{p_{av}}\right)^{k^2}\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right) \\
\lim_{k\to\infty} c &= \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) \\
\lim_{k\to\infty} d &= \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right).
\end{aligned}
$$

To find the limit of this, we need to compare $\lim_{k\to\infty}\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}$ and $\lim_{k\to\infty}\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}$

to determine what else to factor out. We can re-write the ratio of these two terms as

$$\lim_{k\to\infty} \frac{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}} = \lim_{k\to\infty} \frac{\exp(k^2 \ln\left(\frac{p_{bv}}{p_{av}}\right))}{\exp(\frac{k(k+1)}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right))}$$

$$= \lim_{k\to\infty} \exp(k^2 \ln\left(\frac{p_{bv}}{p_{av}}\right)) - \exp(\frac{k(k+1)}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right))$$

$$= \lim_{k\to\infty} \exp\left(k\left[k\left(\ln\left(\frac{p_{bv}}{p_{av}}\right) - \frac{1}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)\right) - \frac{1}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)\right]\right).$$

Note that $\lim_{k\to\infty} \frac{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}} = 0$ if $\ln\left(\frac{p_{bv}}{p_{av}}\right) - \frac{1}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right) < 0$, and $\lim_{k\to\infty} \frac{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}} = \infty$ if $\ln\left(\frac{p_{bv}}{p_{av}}\right) - \frac{1}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right) \geq 0$. We have

$$\ln\left(\frac{p_{bv}}{p_{av}}\right) - \frac{1}{2}\ln\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right) = \ln\left(\frac{\frac{p_{bv}}{p_{av}}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{1}{2}}}\right),$$

which is negative if $\left(\frac{p_{bv}}{p_{av}}\right)^2 < \frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}$, which is equivalent to the condition that $p_{bv}^3(1-p_{bv}) < p_{av}^3(1-p_{av})$.

Suppose that $p_{bv}^3(1-p_{bv}) < p_{av}^3(1-p_{av})$. Then we can also factor out $\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}$ from each of our terms, which yields:

$$\lim_{k\to\infty} a = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k \left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) = 0$$

$$\lim_{k\to\infty} b = \left(\frac{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}}\right)\left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right) = 0$$

$$\lim_{k\to\infty} c = \left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right)$$

$$\lim_{k\to\infty} d = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\infty}\left(\frac{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}}\right)\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right) = 0.$$

Thus, if $p_{bv}^3(1-p_{bv}) < p_{av}^3(1-p_{av})$, then $\lim_{k\to\infty} P^b(B, v|\mathbf{s}_{2k} = \{a, a, a, \ldots b, b, b, \ldots\}) = 1$.

Since $p_{bv}(1-p_{bv}) > p_{av}(1-p_{av})$, then $p_{bv}^3(1-p_{bv}) < p_{av}^3(1-p_{av})$ only if type v is sufficiently A-slanted (i.e., α is sufficiently high) relative to reliability q.

A very similar exercise applies to the subcases of $p_{bv}^3(1-p_{bv}) > p_{av}^3(1-p_{av})$ and $p_{bv}^3(1-p_{bv}) = p_{av}^3(1-p_{av})$.

Suppose that $p_{bv}^3(1-p_{bv}) > p_{av}^3(1-p_{av})$. Then we can instead factor out $\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}$ from each of our terms, which yields

$$\lim_{k\to\infty} a = \left(\frac{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}}{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}\right)\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k\left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) = 0$$
$$\lim_{k\to\infty} b = \left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right)$$
$$\lim_{k\to\infty} c = \left(\frac{\left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}}{\left(\frac{p_{bv}}{p_{av}}\right)^{k^2}}\right)\left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) = 0$$
$$\lim_{k\to\infty} d = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\infty}\left(\prod_{i=1}^{k}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right) = 0.$$

Thus, if $p_{bv}^3(1-p_{bv}) > p_{av}^3(1-p_{av})$, then $\lim_{k\to\infty} P^b(A, w|\mathbf{s}_{2k} = \{a,a,a,\ldots b,b,b,\ldots\}) = 1$. Since $p_{bv}(1-p_{bv}) > p_{av}(1-p_{av})$, then $p_{bv}^3(1-p_{bv}) > p_{av}^3(1-p_{av})$ only if type v is not too A-slanted (i.e., α is sufficiently low) relative to reliability q.

If $p_{bv}^3(1-p_{bv}) = p_{av}^3(1-p_{av})$, then we can factor out $\left(\frac{p_{bv}}{p_{av}}\right)^{k^2} = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^{\frac{k(k+1)}{2}}$ from each of our terms, which yields,

$$\lim_{k\to\infty} a = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k\left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right) = 0$$
$$\lim_{k\to\infty} b = \left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right)$$
$$\lim_{k\to\infty} c = \left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{bv}}{p_{av}})^i\right)\right)$$
$$\lim_{k\to\infty} d = \left(\frac{p_{av}(1-p_{av})}{p_{bv}(1-p_{bv})}\right)^k\left(\prod_{i=1}^{\infty}\left(1+(\frac{1-p_{av}}{p_{bv}})^i\right)\right) = 0.$$

Since $p_{bv}(1-p_{bv}) > p_{av}(1-p_{av})$, then we know that $\frac{1-p_{av}}{p_{bv}} < \frac{1-p_{bv}}{p_{av}}$. This means that when $k \to \infty$, then $0 < P^b(B, v|\mathbf{s}_{2k} = \{a,a,a,\ldots b,b,b,\ldots\}) < P^b(A, w|\mathbf{s}_{2k} = \{a,a,a,\ldots b,b,b,\ldots\}) <$

1 where $P^b(A, w|\mathbf{s}_{2k} = \{a, a, a, \ldots b, b, b, \ldots\}) + P^b(B, v|\mathbf{s}_{2k} = \{a, a, a, \ldots b, b, b, \ldots\}) = 1$.

To interpret the three conditions for Proposition 10 in terms of reliability and slant, define $G(q, \alpha) \equiv p_{bv}^3(1 - p_{bv}) - p_{av}^3(1 - p_{av})$ where $p_{bv} = q(1 - \alpha)$ and $p_{av} = q + (1 - q)\alpha$. We can re-write $G(q, \alpha)$ to obtain:

$$G(q, \alpha) = (1 - \alpha) \left(q^3(1 - \alpha)^2(1 - q(1 - \alpha)) - (1 - q)(q + (1 - q)\alpha)^3\right) \tag{34}$$

We will consider how G varies with α for fixed q, where we only consider the relevant range of $q \in (\frac{1}{2}, 1)$. First, note that $G(q, 0) = G(q, 1) = 0$ for all $q \in (\frac{1}{2}, 1)$. Second, we can have

$$\frac{\partial G}{\partial \alpha} = q^3(1 - \alpha)\left(-2(1 - q(1 - \alpha)) + (1 - \alpha)q\right) - 3(1 - q)^2(q + (1 - q)\alpha)^2.$$

Evaluating $\frac{\partial G}{\partial \alpha}$ for $\alpha = 0$ and $\alpha = 1$ yields $\frac{\partial G}{\partial \alpha}|_{\alpha=0} = q^2(4q - 3)$ and $\frac{\partial G}{\partial \alpha}|_{\alpha=1} = 1 - q$. Thus, $\frac{\partial G}{\partial \alpha}|_{\alpha=0} > 0$ for $q \in [3/4, 1]$, $\frac{\partial G}{\partial \alpha}|_{\alpha=0} \leq 0$ for $q \in (\frac{1}{2}, \frac{3}{4}]$ and $\frac{\partial G}{\partial \alpha}|_{\alpha=1} > 0$ for $q \in (\frac{1}{2}, 1)$.

Suppose there exists some interior $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$. Evaluating $\frac{\partial G}{\partial \alpha}$ at an $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$ yields

$$\begin{aligned}\frac{\partial G}{\partial \alpha}|_{\alpha=\bar{\alpha}} &= (1 - q)(q + (1 - q)\bar{\alpha})^2 \left(-3(1 - q) + (q + (1 - q)\bar{\alpha}) \left(\frac{-2}{1 - \bar{\alpha}} + \frac{q}{1 - q(1 - \alpha)}\right)\right) \\ &= \frac{(1 - q)(q + (1 - q)\bar{\alpha})^2}{(1 - \bar{\alpha})(1 - q(1 - \bar{\alpha}))}\left(4q - 3 - \bar{\alpha}(4q - 1)\right).\end{aligned}$$

This implies that if there exists an $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$, then $\frac{\partial G}{\partial \alpha}|_{\alpha=\bar{\alpha}} < 0$ when $\bar{\alpha} > \frac{4q-3}{4q-1}$, where we know that $4q - 1 > 0$. Moreover, $\frac{4q-3}{4q-1} \leq 0$ when $q \in (\frac{1}{2}, \frac{3}{4}]$. This implies that any such $\bar{\alpha}(q) \in (0, 1)$ must satisfy $\frac{\partial G}{\partial \alpha}|_{\alpha=\bar{\alpha}} < 0$ when $q \in (\frac{1}{2}, \frac{3}{4}]$.

Since $G(q, 0) = G(q, 1) = 0$ for all $q \in (\frac{1}{2}, 1)$, $\frac{\partial G}{\partial \alpha}|_{\alpha=0} \leq 0$ for $q \in (\frac{1}{2}, \frac{3}{4}]$, $\frac{\partial G}{\partial \alpha}|_{\alpha=1} > 1$ for $q \in (\frac{1}{2}, 1)$, and any such $\bar{\alpha}(q) \in (0, 1)$ must satisfy $\frac{\partial G}{\partial \alpha}|_{\alpha=\bar{\alpha}} < 0$ when $q \in (\frac{1}{2}, \frac{3}{4}]$, then there exists no $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$ when $q \in (\frac{1}{2}, \frac{3}{4}]$. That is, for any given $q \in (\frac{1}{2}, \frac{3}{4}]$, $p_{bv}^3(1 - p_{bv}) < p_{av}^3(1 - p_{av})$ for all $\alpha \in (0, 1)$.

Consider $q \in (\frac{3}{4}, 1)$. Since $G(q, 0) = G(q, 1) = 0$ for all $q \in (\frac{1}{2}, 1)$, $\frac{\partial G}{\partial \alpha}|_{\alpha=0} > 0$ for $q \in [3/4, 1)$, and $\frac{\partial G}{\partial \alpha}|_{\alpha=1} > 0$ for $q \in (\frac{1}{2}, 1)$, then we know that there exists an $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$ for a given $q \in (\frac{3}{4}, 1])$. Moreover, we know from Equation 34 that $G(q, \alpha)$ is a 4th order polynomial in α. Thus for a fixed q, $G(q, \alpha) = 0$ for at most four values of $\alpha \in \mathbb{R}$. Given the fact that $G(q, 0) = G(q, 1) = 0$ for all $q \in (\frac{1}{2}, 1)$, $\frac{\partial G}{\partial \alpha}|_{\alpha=0} > 0$ for $q \in [3/4, 1)$, and $\frac{\partial G}{\partial \alpha}|_{\alpha=1} > 0$ for $q \in (\frac{1}{2}, 1)$, then $G(q, \alpha) = 0$ for at least five values of $\alpha \in \mathbb{R}$ if $\bar{\alpha}(q)$ is not unique. Thus it is not possible that $\bar{\alpha}(q)$ is not unique. Therefore, for a given $q \in (\frac{3}{4}, 1)$, there exists a unique $\bar{\alpha}(q) \in (0, 1)$ such that $G(q, \bar{\alpha}) = 0$. This implies that for any given $q \in (\frac{3}{4}, 1)$, there exists a unique $\bar{\alpha}(q) \in (0, 1)$ such that $p_{bv}^3(1 - p_{bv}) > p_{av}^3(1 - p_{av})$ for $\alpha < \bar{\alpha}(q)$, $p_{bv}^3(1 - p_{bv}) = p_{av}^3(1 - p_{av})$ for $\alpha = \bar{\alpha}(q)$, and $p_{bv}^3(1 - p_{bv}) < p_{av}^3(1 - p_{av})$ for $\alpha > \bar{\alpha}(q)$.

To summarize:

1. If $q \leq 3/4$, $\lim_{k\to\infty} P^b(B, v|\mathbf{s}_{2k}) = 1$;
2. If $q > 3/4$, $\exists$ a unique $\bar{\alpha}(q) \in (0, 1)$ such that:
   (a) $\forall\ \alpha > \bar{\alpha}(q), \lim_{k\to\infty} P^b(B, v|\mathbf{s}_{2k}) = 1$,
   (b) $\forall\ \alpha < \bar{\alpha}(q), \lim_{k\to\infty} P^b(A, w|\mathbf{s}_{2k}) = 1$,
   (c) If $\alpha = \bar{\alpha}(q)$, then $\lim_{k\to\infty}[P^b(A, w|\mathbf{s}_{2k}) + P^b(B, v|\mathbf{s}_{2k})] = 1$ and $P^b(B, v|\mathbf{s}_{2k}) > 1/2$.

Proof of Proposition 11

To show the results when agents receive signals from multiple sources, note that Equation (28) can also be re-written as

$$P^b(c_1, c_2, \theta|\mathbf{s}_{n_1,n_2}) = \frac{\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{c_1}\omega_0^{c_2}\omega_0^{\theta}\beta_{c_1}(\mathbf{s}_{n_1})\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2})}{\sum_{c_2}\sum_{c_1}\sum_{\theta}\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{c_1}\omega_0^{c_2}\omega_0^{\theta}\beta_{c_1}(\mathbf{s}_{n_1})\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2})}, \quad (35)$$

where $\beta_{c_1}(\emptyset) = 1$ and:

$$\beta_{c_1}(\mathbf{s}_{n_1}) = \prod_{m=1}^{n_1}\left(\sum_{\theta}\left(\prod_{t=1}^{m} P(s_{t1}|c_1,\theta)\right)\omega_0^{\theta}\right)$$

$$\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2}) = \prod_{m=n_1+1}^{n_1+n_2}\left(\sum_{\theta}\left(\prod_{t=n_1+1}^{m} P(s_{t2}|c_2,\theta)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1,\theta)\right)\omega_0^{\theta}\right).$$

Given $\mathbf{s}_{n_1} = (a, \ldots, a)$ and $\mathbf{s}_{n_2} = (b, \ldots, b)$ where $n_1 = n_2 = k$ and $k > 1$, we can write the pre-screener's beliefs using:

$$\beta_{c_1}(\mathbf{s}_{n_1}) = \prod_{m=1}^{k}\left(p_{ac_1}^m + (1-p_{bc_1})^m\right) = \prod_{m=1}^{k} p_{ac_1}^m\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^m\right)$$

$$\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2}) = \prod_{m=1}^{k}\left(p_{ac_1}^k(1-p_{ac_2})^m + (1-p_{bc_1})^k p_{bc_2}^m\right) = \prod_{m=1}^{k} p_{ac_1}^k(1-p_{ac_2})^m\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^k\left(\frac{p_{bc_2}}{1-p_{ac_1}}\right)^m\right)$$

Therefore we can write

$$\begin{aligned}\beta_{c_1}(\mathbf{s}_{n_1})\beta_{c_2c_1}(\mathbf{s}_{n_1,n_2}) &= \left(\prod_{m=1}^{k} p_{ac_1}^m\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^m\right)\right)\left(\prod_{m=1}^{k} p_{ac_1}^k(1-p_{ac_2})^m\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^k\left(\frac{p_{bc_2}}{1-p_{ac_1}}\right)^m\right)\right) \\ &= (p_{ac_1}(1-p_{ac_2}))^{\frac{k(k+1)}{2}} p_{ac_1}^{k^2}\prod_{m=1}^{k}\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^m\right)\left(1 + \left(\frac{1-p_{bc_1}}{p_{ac_1}}\right)^k\left(\frac{p_{bc_2}}{1-p_{ac_1}}\right)^m\right).\end{aligned} \quad (36)$$

We also have that

$$\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2, A)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1, A)\right) = (p_{ac_1}(1-p_{ac_2}))^k \tag{37}$$

$$\left(\prod_{t=n_1+1}^{n_1+n_2} P(s_{t2}|c_2, B)\right)\left(\prod_{t=1}^{n_1} P(s_{t1}|c_1, B)\right) = ((1-p_{bc_1})p_{bc_2}))^k \tag{38}$$

Assume that each state and each type is equally likely ex ante: $\omega_0^A = 1/2$ and $\omega_0^c = 1/4\ \forall c \in \{v, w, x, z\}$. Suppose that there are four possible source types ($c \in \{v, w, x, z\}$), where reliability and slant are uncorrelated and slant is symmetric ($\alpha = \gamma \geq 0$ and $q_H \geq q_L \geq 1/2$): $v = (q_H, \alpha, 0)$, $w = (q_L, \alpha, 0)$, $x = (q_H, 0, \gamma)$, $w = (q_L, 0, \gamma)$.

Substituting Equations (36), (37), and (38) into Equation (35), we can then take the limit as $k \to \infty$. The relevant cases are:

1. Uncertainty about reliability only: $q_H > q_L$, $\alpha = \gamma = 0$

   This implies that $v = x$ and $w = z$, so we can write everything just in terms of x and z. This implies $p_{ax} = p_{bx}, p_{az} = p_{bz}$. We can also show that $p_a z > 1 - p_{ax}$, which implies that $p_{ax}(1-p_{ax}) > p_{az}(1-p_{az})$. This yields $\lim_{k\to\infty} P(x_1, z_2, A) = \lim_{k\to\infty} P(z_1, x_2, B) = \frac{1}{2}$ and $\lim_{k\to\infty} P^b(x_1, z_2, A) = 1$.

2. Uncertainty about reliability and slant: $q_H > q_L$ and $\alpha = \gamma > 0$

   This implies the following properties: $p_{av} = p_{bx}, p_{bv} = p_{ax}, p_{aw} = p_{bz}, p_{bw} = p_{az}$. Given this symmetry we can write everything just in terms of x and z. We can also show that $p_{bx} > p_{bz} > p_{az}, p_{bx} > p_{ax} > p_{az}, p_{ax} > 1 - p_{bx}, p_{az} > 1 - p_{bz}$. Therefore, $p_{bx}(1-p_{az}) = \max\ \{p_{bc_1}(1-p_{ac_2}), p_{bc_2}(1-p_{ac_1}), p_{ac_1}(1-p_{bc_2}), p_{ac_2}(1-p_{bc_1})\}$. This yields $\lim_{k\to\infty} P(v_1, z_2, A) = \lim_{k\to\infty} P(w_1, x_2, B) = \frac{1}{2}$ and $\lim_{k\to\infty} P^b(v_1, z_2, A) = 1$.

3. Uncertainty about slant only: $q_H = q_L$, $\alpha = \gamma > 0$

   This implies that $v = w$ and $x = z$. Also, $p_{av} = p_{bx} > 1/2 > p_{ax} = p_{bv}$ so we can write everything just in terms of x. We can also show that $p_{bx} > 1 - p_{ax}$. This yields $\lim_{k\to\infty} P(v_1, x_2, A) = \lim_{k\to\infty} P(v_1, x_2, B) = \frac{1}{2}$ and $\lim_{k\to\infty} P(v_1, x_2, A) = \lim_{k\to\infty} P(v_1, x_2, B) = \frac{1}{2}$.