Disliking to disagree*

Florian Hoffmann† Kiryl Khalmetski‡ Mark T. Le Quement§

September 29, 2022

Abstract

We propose a theory of perceived disagreement aversion, capturing a preference for not being viewed by one's interlocutor as holding a fundamentally different opinion. We study the impact of this preference on information sharing in a model of strategic disclosure with heterogeneous prior beliefs. Equilibrium disclosure is partial and biased towards the prior opinion of the most confident party. The informativeness of equilibrium disclosure is higher the larger the difference in the interlocutors' opinions and the more similar their confidence levels. Senders' preferences over communication partners induce assortative matching in prior opinions, which hurts information sharing.

Keywords: strategic disclosure, psychological games, disagreement aversion

JEL classification: D81, D83, D91

Decentralized information exchange within social networks, for example, in the form of informal political talk, is an important source of information for many. While it avoids some of the distortions associated with centralized information provision (e.g., via mass media), it is not bias free. Individuals are not equally likely to discuss all topics or disclose all opinions and facts, and the average profile of conversation partners is far from random. A robust pattern is that people avoid conflictual topics and prefer to interact with people who share their worldview (see, for example, Mutz, 2006; Sunstein, 2007). According to a recent survey by Pew Research Center (2018), more than 50% of Americans find it stressful and frustrating to discuss politics with people whom they disagree with.[1]

We propose a theory of disagreement aversion which provides a possible explanation of the above patterns of selective information sharing and homophily. Our starting point is that disagreement often mechanically spurs negative emotions (antipathy, resentment) towards the other. This immediate emotional response might in turn trigger a set of higher order negative emotions (stemming from disliking to be disliked, disliking to dislike, disliking to being sensed to dislike, etc.) which contribute to creating an unpleasant social

*We thank participants at various seminars and conferences for helpful comments as well as Peter Norman Sørensen (IAST Toulouse discussant) and Joel Sobel for detailed suggestions on an earlier draft. Khalmetski gratefully acknowledges financial support of the German Research Foundation (DFG) through the Research Unit "Design and Behavior" (FOR 1371).

†Faculty of Economics and Business, KU Leuven. E-mail: florian.hoffmann@kuleuven.be

‡Department of Economics, University of Bath. E-mail: kk917@bath.ac.uk

§School of Economics, University of East Anglia. E-mail: m.le-quement@uea.ac.uk.

[1] Similarly, Dorison et al. (2019) find that *"people who hold strong opinions on an issue rated policy discussion with holders of opposing views as more aversive than any other activity listed, including household chores, yard work, and a visit to the dentist."*

experience. In daily conversations, keeping interaction pleasant typically ranks high among a speaker's priorities, whilst trying to persuade the other is often secondary.[2] As a consequence, the speaker might often prefer to bite his tongue and omit facts that increase perceived disagreement while instead eagerly sharing information that reduces it. Generally, belief homophily or a preference for "belief consonance" with communication partners has been widely documented in social psychology, being often linked to fundamental individual motivations like protecting own identity, maintaining harmonic social relations or aversion to cognitive dissonance (Golman et al., 2016).[3]

We suggest a simple formalization of perceived disagreement and study a disagreement averse sender's incentive to share privately held hard information with a receiver in situations with heterogeneous prior beliefs. Our main objective is to determine the effect of differences in prior beliefs on information sharing in equilibrium In particular, are there systematic biases regarding which information is shared and which profiles of priors lead to the most information sharing?

We consider a standard normal-normal environment in which agents' prior beliefs over the state and the conditional distributions of signals are normal. We interpret an agent's prior mean as her prior opinion and the prior variance as her confidence in her prior opinion. Agents' priors are heterogeneous and commonly known. Our baseline setting is a model of voluntary disclosure with uncertainty about the sender's information endowment (e.g., Dye, 1985; Jung and Kwon, 1988).

We define disagreement, given a signal, as the absolute difference between the sender's and the receiver's posterior means given the signal. The receiver's *perceived* disagreement given disclosure is simply his expected value of posterior disagreement conditional on the sender's disclosure. In case no information is disclosed, perceived disagreement is thus the expected value of the magnitude by which the sender's posterior opinion deviates from what would be the receiver's posterior opinion given the sender's information. The sender's utility decreases in the receiver's perceived disagreement while the receiver simply aims to learn the state as accurately as possible.

An essential property of the setting is that different prior confidence levels – by leading players to update with different intensities – imply a mixed effect of information on disagreement. There is a closed interval of signals which strictly reduce disagreement and a unique signal yielding zero ex post disagreement. Beyond this interval, posterior disagreement is larger than prior disagreement.[4]

[2]See for example Pennington and Winfrey (2021) on motivations behind political posting on Facebook, as well as Lyons et al. (2016) on the relation between the agreeableness personality trait and disagreement aversion.

[3]In particular, besides interpersonal conflict, a distinct potential source of aversion to conversational disagreement is intrapersonal conflict. An individual may be intrinsically averse to contradicting information, which confronts her with a choice between the two unattractive options of either revising beliefs (thereby acknowledging that she was wrong) or experiencing cognitive dissonance (Festinger, 1957).

[4]This is most easily seen by comparing the belief adjustment of two agents with the same prior mean, but whose prior variances are respectively zero and some positive number. The agent with the higher prior variance will update her belief after any informative signal, thus deviating from the belief of the other agent which remains invariant to any signal.

Our first set of findings concerns the qualitative nature of equilibrium disclosure. If agents have different prior variances, the equilibrium disclosure is partial and features what can be termed an opinion-corridor. There is a finite threshold such that a signal is disclosed only if it yields a disagreement smaller than this threshold level, the latter being strictly larger than prior disagreement. This translates into a unique interval of disclosed signals, symmetric around the zero disagreement signal and containing all signals that reduce prior disagreement. The disclosure interval is biased towards the more confident player's prior opinion, its center being located closer to the prior mean of the latter. The reason is that the more confident party moves less, when contradicted, than the less confident party. Signals that contradict the former's opinion thus generate larger disagreement and are thus less attractive to disclose. If players' prior means are very close, the center of the disclosure interval is close to these means, generating an echo-chamber effect whereby homogeneity in prior opinions leads to endogenous confirmation bias.

Our second set of findings concerns the parties' preference over communication partners. This captures the situation of a decision maker (DM) who picks an advisor among a rich pool of agents. These differ in terms of their prior beliefs but share the same expertise level and they are all disagreement averse. The DM's objective is to obtain as much information as possible. Which expert should he pick? We identify a simple sufficient condition such that the sender's informativeness increases in the difference in prior means and decreases in the difference in prior variances.

We also study a sender's preference across various receivers if she evaluates these before observing her signal. This captures, e.g., the choice of long term communication partners on social media (e.g., Facebook). An agent creates a link with someone with an eye to sharing information with this person in the future, being motivated by an image concern (to be perceived as agreeing and therefore liked) rather than a persuasion motive. We find that the sender prefers receivers with closer prior means. There is thus a conflict of interest between senders' preference for minimizing the difference in prior means and receivers' preference for maximizing it.

Next, we apply our model to sender-receiver matching in a population with multiple senders and receivers. We contrast endogenous matching when senders pick their most preferred receivers to the case of random matching, focusing on the consequences for the informativeness of equilibrium disclosure. We identify when the sender's belief homophily in selecting communication partners is detrimental for the informativeness of subsequent communication such that receivers prefer random over endogenous matching.

We further discuss several extensions. First, our notion of perceived disagreement aversion easily extends beyond disagreement in means to more general measures of disagreement which also take into account difference in variances such as relative entropy without affecting the fundamental features of equilibrium disclosure. Second, perceived disagreement aversion also provides interesting implications for information acquisition. In that respect we study a simple model of conversation to generate public information. In such a setting perceived disagreement averse players often assign negative ex ante value to information if their prior opinions are too close. That is, certain prior configurations are an obstacle not only for information sharing but also for information acquisition.

Literature review Our paper relates to three literatures. An essential starting point of our paper lies in the social psychology literature on conformism and homophily in social networks (see Newcomb, 1961; Asch, 1955; Lazarsfeld and Merton, 1954; Goffman, 1959). In the seminal experiments conducted by Asch (1955), subjects wrongly evaluated the length of a line in public after being exposed to others' (artificially induced) wrong assessment. Deutsch and Gerard (1955) show a weaker effect if judgments are reported privately, so that others' *perceived* disagreement is unaffected. Bursztyn et al. (2020) found that subjects were more likely to publicly state negative views on immigration after Donald Trump's victory than before it. Prentice and Miller (1993) established that a large fraction of students refrained from expressing dissent with campus alcohol practices for fear of stigma, underestimating how many shared their opinion. Field studies conducted by Lazarsfeld and Merton's in the 1950s revealed a large degree of homophily in social networks in the USA and Mutz (2006) documents similar patterns in contemporary USA. A burgeoning contemporary literature in neuroscience and psychology studies disagreement aversion, as illustrated for example by Domínguez et al. (2016). Golman et al. (2016) provide a rich review of the behavioral and economic literature on the preference for belief consonance.

Our paper also relates to a literature studying the relation between public information and disagreement in beliefs. Disagreement in posterior beliefs in these settings may result from different prior beliefs (Dixit and Weibull, 2007; Acemoglu et al., 2007) or private signals (Andreoni and Mylovanov, 2012) as well as ambiguity (Baliga et al., 2013).[5] Disagreement in beliefs may under some conditions persist in the long run, i.e., asymptotically (Acemoglu et al., 2016; Andreoni and Mylovanov, 2012).[6] A related approach is the theory of correlation neglect (see Levy and Razin, 2015), which posits that agents underestimate the correlation between others' information and their own, leading them to overweight the former.

Our paper finally relates to the literature on strategic information transmission. An extensive body of research dating back to Grossman (1981) and Milgrom (1981) studies strategic disclosure of verifiable signals,[7] assuming misaligned preferences over the receiver's action conditional on the state. Newer papers assume different prior beliefs, such as Banerjee and Somanathan (2001) and Kartik et al. (2021). While these papers share important elements of our analysis, perceived disagreement does not affect equilibrium disclosure incentives. Che and Kartik (2009) (CK) examine the effect of prior belief misalignment (in means only) on the sender's incentives to both privately acquire and share hard information. Prior misalignment unambiguously hurts disclosure but increases information acquisition, so that the receiver may benefit from more misalignment. There are a number of key difference to our analysis. The objectives of the sender are different: Induce the right action vs. minimize perceived disagreement. Besides, our analysis mostly assumes different prior

[5] Several papers in network economics consider the effect of individual conformity to the beliefs or opinions of others on belief polarization (Dandekar et al., 2013; Buechel et al., 2015; Golub and Jackson, 2012).

[6] Sethi and Yildiz (2016) focus on the fact that observing others' opinion over time, an observer learns both about their subjective prior and about their private information concerning some objective state, thereby triggering non-trivial dynamics in belief updating.

[7] See Sobel (2013) for a general review of the literature on strategic information transmission.

variances while CK assume identical variances throughout their analysis. This leads to important differences in equilibrium predictions. In CK, if priors differ, the sender discloses all signals except some closed interval, while our setup delivers the reverse prediction. Further, in our model, different prior means lead to higher informativeness of communication (under a simple sufficient condition). In contrast, in CK a difference in prior opinions under exogenous information always hurts disclosure. Thus, in CK the benefit of prior misalignment in means lies exclusively in incentivizing information (while it hurts disclosure), whereas in our setup prior misalignment in means incentivizes disclosure directly.[8]

A strand of the literature on strategic information transmission features an endogenous preference for belief conformity arising from reputational concerns. In Gentzkow and Shapiro (2006) and Ottaviani and Sørensen (2006a,b), the sender wishes to signal high expertise to the receiver (who observes the state with positive probability). This incentivizes her to bias her cheap talk message towards the receiver's prior belief (which is identical to the sender's prior belief in Ottaviani and Sørensen, 2006a,b). Visser and Swank (2007) study deliberative committees whose members want to signal high expertise. This gives them an incentive to pretend to have similar signals (i.e., to agree) and to decide against the prior. Within a similar setup Levy (2007) focuses on the impact of transparency rules on decision making. As in Gentzkow and Shapiro (2006), in our setup if the sender is less confident than the receiver, she tends to omit signals contradicting the receiver's prior opinion, but here her motivation is simply to mitigate perceived disagreement (her information quality being known). This same objective will drive her to omit signals that confirm the receiver's prior if the sender is more confident than the receiver.[9]

There is a significant and growing literature on strategic information transmission and acquisition with belief-based utility (see Caplin and Leahy, 2001, 2004; Kőszegi, 2006; Ely et al., 2015).[10] Here, information affects players in a non-instrumental way. For example, a receiver might experience anxiety and therefore have preferences over the timing of uncertainty resolution.[11]

[8] The heterogeneous prior assumption has become common in economic theory. Apart from the above, see also for example Yildiz (2004), Acemoglu et al. (2016) and Gentzkow and Shapiro (2006) for interesting applications and Morris (1995) for a conceptual discussion.

[9] In Morris (2001) (see also Sobel, 1985; Benabou and Laroque, 1992; Ely and Välimäki, 2003), the sender wants to be perceived as unbiased, which leads to distorted communication. The following contributions also bear an indirect relation to the above. In the principal-agent setting studied by Prendergast (1993), the use of subjective performance opinion creates incentives for the agent to inefficiently conform to the opinion of the supervisor. In consequence, the principal may be better off avoiding the use of incentive contracts in order to induce better reporting. Bursztyn et al. (2020) consider a setting where a sender has to communicate her type to a receiver and has an incentive to appear of the same type as the receiver. Bénabou (2012) shows that agents with anticipatory utility may converge to each other's wrong beliefs due to the dependence of one's payoffs on the actions of the others.

[10] See Battigalli and Dufwenberg (2022) for a general overview of models incorporating belief-dependent preferences.

[11] Overall, our approach offers a novel methodological perspective on the strategic communication as disagreement aversion generates *both* types of incentives: to signal one's own beliefs about the state of the world *and* to shift the opponent's beliefs about the state. Most of the literature so far has primarily focused on one of these two incentives.

1 Setup

We study a simple model of voluntary disclosure with a *perceived disagreement* averse sender. There are two players, a sender S (she) and a receiver R (he). The two players initially disagree, in the sense of having different priors over an underlying state of nature ω drawn from state space Ω. For concreteness, we assume $\Omega = \mathbb{R}$ and that each player $i \in \{S, R\}$ has a commonly known normal prior $N(\mu_i, \gamma_i^2)$, $\gamma_i > 0$. We call μ_i i's prior *opinion*, and $1/\gamma_i^2$ i's *confidence* in her prior opinion.

S can disclose information about ω, but as in Dye (1985) or Jung and Kwon (1988) there is uncertainty about her information endowment. In particular, with probability φ, S privately observes an informative signal $\sigma = \omega + \varepsilon$, where ε is commonly known to be distributed according to $N(0, \gamma_\varepsilon^2)$. Thus, $1/\gamma_\varepsilon^2$ measures the precision of S's signal. If S observes no information about ω, we denote this by $\sigma = \varnothing$ (empty signal).

R does not observe signal σ nor whether S is informed or not. If $\sigma \neq \varnothing$, S decides whether to disclose the signal to R or not. We denote by d the disclosed information, where $d \in \{\sigma, \varnothing\}$ if S is informed and $d = \varnothing$ otherwise. R only observes the disclosed information d, based on which he updates his prior belief using Bayes rule. Subsequently, R takes an action $a \in \mathbb{R}$ so as to maximize the utility function:[12]

$$u_R(\omega, a) = -(\omega - a)^2 . \tag{1}$$

While R thus prefers a more informative disclosure strategy that allows for a lower expected loss, S's payoff is not affected by R's action choice. Instead, S values lower disagreement in posterior opinions as perceived by R. Concretely, for any disclosed information d, we define R's *ex post perceived disagreement* (PD) as

$$\widetilde{\Delta}(d) = E_R[\Delta(\sigma)\,|d] \tag{2}$$

where

$$\Delta(\sigma) = |E_S[\omega|\sigma] - E_R[\omega|\sigma]| \tag{3}$$

is the *actual ex post disagreement* between S and R given $\sigma \in \{\mathbb{R}, \varnothing\}$. Put differently, $\Delta(\sigma)$ reflects the magnitude by which S's posterior belief about ω given signal σ deviates from R's posterior belief given the same signal (i.e., the most accurate belief conditional on σ from R's perspective).[13] The key novelty in our analysis is that S dislikes perceived disagreement on the part of R (i.e., dislikes being perceived as having inaccurate beliefs), which we model by stipulating that S chooses disclosure d to minimize $\widetilde{\Delta}(d)$ defined in (2).

Since R's information is controlled by S via her disclosure choices, R might not have suf-

[12] As shall be made clear next, the utility function of R is irrelevant for the characterization of S's optimal disclosure strategy. It will become relevant when assessing the value of equilibrium disclosure for R and associated comparative statics, such as R's preferences over senders with heterogeneous priors. The standard quadratic loss function is for tractability only as most of our results extend to more general specifications of u_R.

[13] While our baseline model considers disagreement in means, the key features of our equilibrium characterisation extend to more general measures of differences in posteriors such as the Kullback-Leibler divergence (Relative Entropy) or Bhattacharya distance (see Section 5).

ficient information to compute actual disagreement. In particular, if S does not disclose any information (i.e., $d = \varnothing$), then R's perceived disagreement $\widetilde{\Delta}(\varnothing)$ is his expectation of actual disagreement $\Delta(\sigma)$ across all signals σ that induce non-disclosure $d = \varnothing$ in equilibrium, i.e., $\widetilde{\Delta}(\varnothing) = E_R[\Delta(\sigma)\,|d = \varnothing]$. Instead, if $d = \sigma \neq \varnothing$, R's perceived disagreement simply equals the actual disagreement, i.e., $\widetilde{\Delta}(d) = \Delta(\sigma)$. Thus, perceived disagreement $\widetilde{\Delta}(d)$ as defined in (2) is R's posterior estimate of the magnitude by which S's actual posterior belief about the state, i.e., $E_S[\omega|\sigma]$, deviates from what would be the most accurate belief from R's perspective given S's information, i.e., $E_R[\omega|\sigma]$. Put differently, $\widetilde{\Delta}(d)$ measures by how much S's posterior belief is on average incorrect as estimated by R. Another way to interpret our measure of perceived disagreement $\widetilde{\Delta}(d)$ is that it reflects R's estimate, given the disclosed information, of what would be the actual disagreement $\Delta(\sigma)$ if R observed the same information as S.[14]

To build intuition about the concept of perceived disagreement, note how it differs from what could be termed plain perceived disagreement after disclosure d, i.e., $E_R[|E_S[\omega|\sigma] - E_R[\omega|d]|]$, which is simply R's expectation of the absolute distance between S's posterior mean conditional on σ and R's posterior mean conditional on d. Suppose for example that S would be known to receive with probability one a perfectly informative signal. Then, perceived disagreement after non-disclosure would always be zero given our definition (2), whatever the assumed equilibrium disclosure strategy. Indeed, R would simply recognize that S's belief about the state is correct. In contrast, the plain disagreement measure conditional on non-disclosure would yield a positive value since in this case R would know that S's actual posterior belief would (almost) always deviate from R's posterior belief conditional on non-disclosure.[15]

2 Equilibrium analysis

In this section we characterize equilibria of the disclosure game. Our equilibrium concept throughout is sequential equilibrium.[16] An equilibrium of the disclosure game is characterized by a disclosure strategy of S (a probability of disclosing at each of S's information sets) and a decision rule of R (a distribution over $a \in \mathbb{R}$ at each of R's information sets) that are

[14] As Golman et al. (2016, p. 170) put it, "*If other people believe something different because they do not have access to the same information, it is easier to assume that they would believe the same as oneself if they had access to one's own information. If they have come to different beliefs from the same information, this poses a much greater challenge to one's own interpretation of reality.*"

[15] Another alternative measure of perceived disagreement would be $|E_R\,[E_S[\omega|\sigma]\,|d] - E_R[\omega|d]|$, i.e., the absolute distance between R's posterior mean and R's expectation of S's posterior mean. However, this measure would also yield unnatural estimates. Indeed, suppose that S and R have the same prior mean μ and different prior variances. Consider a putative equilibrium in which S would never disclose. Here, perceived disagreement after non-disclosure would be strictly positive given our definition, since R would recognise that if S receives a signal, then $E_S[\omega|\sigma]$ deviates from R's belief conditional on σ, $E_R[\omega|\sigma]$ (i.e., the correct belief conditional on σ from the perspective of R), due to different prior variances and hence different speeds of belief updating. Yet, the alternative perceived disagreement measure defined above yields a value of 0 after non-disclosure, as we would have $E_R\,[E_S[\omega|\sigma]\,|d = \varnothing] = E_R[\omega|d = \varnothing] = \mu$.

[16] The concept of sequential equilibrium was extended for psychological games, as in our model, by Battigalli and Dufwenberg (2009).

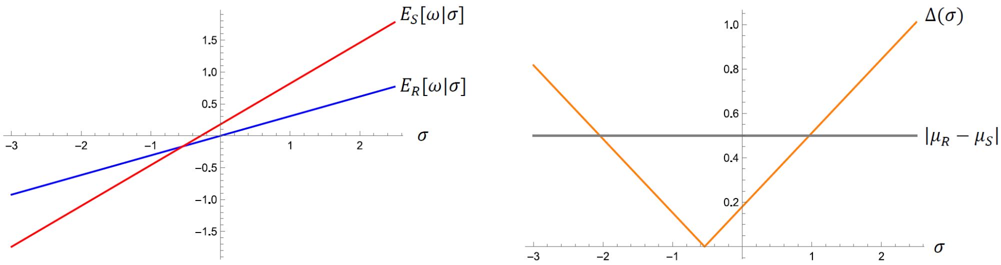

Figure 1: (*posterior disagreement*) The left panel plots the sender's (red line) and receiver's (blue line) posterior opinions as a function of the signal σ. The right panel depicts the implied ex post disagreement.

sequentially rational given players' beliefs. Beliefs are determined via Bayes rule whenever possible.[17] For simplicity of exposition and without loss of generality for the qualitative results, we assume that when indifferent, S prefers disclosure over non-disclosure, i.e., $d = \sigma$ if $\widetilde{\Delta}(\sigma) = \widetilde{\Delta}(\varnothing)$.[18]

2.1 Updating and ex post disagreement

As is well known, upon observing an informative signal $\sigma \in \mathbb{R}$, player i's posterior over the state of the world ω is also normal (conjugate prior property) with mean

$$E_i\left[\omega\,|\sigma\right] = \left(\frac{1}{1+\frac{\gamma_i^2}{\gamma_\varepsilon^2}}\right)\mu_i + \left(\frac{1}{1+\frac{\gamma_\varepsilon^2}{\gamma_i^2}}\right)\sigma \tag{4}$$

and variance

$$Var_i\left[\omega\,|\sigma\right] = \frac{1}{\frac{1}{\gamma_i^2}+\frac{1}{\gamma_\varepsilon^2}}. \tag{5}$$

Intuitively, the posterior mean in (4) is a convex combination of i's prior mean μ_i – i.e., his prior opinion – and the signal σ, where the weight on μ_i is decreasing in the signal to noise ratio $\frac{\gamma_i^2}{\gamma_\varepsilon^2}$. The posterior variance in (5) is decreasing in both i's confidence level $1/\gamma_i^2$ and signal precision $1/\gamma_\varepsilon^2$.

Following any informative signal $\sigma \in \mathbb{R}$, actual ex post disagreement as defined in (3) is given by the difference between the S's and R's posterior opinion, which from (4) depends on their respective prior opinion and confidence. We first focus on the generic case of non-identical prior confidences. Figure 1 shows a numerical example of this case, depicting the posterior opinions $E_S\left[\omega\,|\sigma\right]$ and $E_R\left[\omega\,|\sigma\right]$ (left panel) and their difference, ex post disagreement $\Delta(\sigma)$ (right panel), as functions of $\sigma \in \mathbb{R}$.

As is apparent from the left panel, $E_S\left[\omega\,|\sigma\right]$ (red line) and $E_R\left[\omega\,|\sigma\right]$ (blue line) are both linear functions of σ. Their intercepts differ, since S has a higher prior mean than R. Crucially, also their slopes differ as S and R update at different speeds. This is due to the

[17] Off-equilibrium beliefs in our setting are trivial. In the sequential equilibrium, R's belief about σ is always equal to σ if $\sigma \neq \varnothing$ is disclosed, independently of whether such disclosure is on the equilibrium path or not. Non-disclosure $\varnothing$ is always on the equilibrium path since S is uninformed with a positive probability.

[18] The set of such signals has zero measure in equilibrium unless S and R have identical priors.

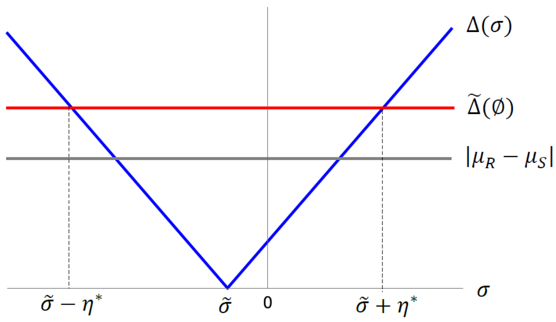

Figure 2: (*equilibrium disclosure*) The graph illustrates the equilibrium disclosure interval, with endpoints at the intersection of perceived disagreement upon disclosure (blue line) and upon non-disclosure (red line).

fact that R is more confident in his prior opinion than S and, hence, puts relatively less weight on the signal (see (4)). Accordingly the red line is steeper in the signal σ than the blue line. As a consequence, the lines cross exactly once and the distance between posterior opinions – actual ex post disagreement $\Delta(\sigma)$ – has the V-shaped form shown in the right panel with minimum at $\widetilde{\sigma}$. That is, given S and R's priors, there is a unique signal

$$\widetilde{\sigma} = \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2) - \mu_R(\gamma_S^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} \tag{6}$$

for which $\Delta(\widetilde{\sigma}) = 0$. We refer to this signal as the *zero disagreement signal.* Now, since R and S update their prior at different (fixed) speeds for any given signal $\sigma \in \mathbb{R}$, actual ex post disagreement $\Delta(\sigma)$ increases linearly and tends to infinity as σ moves away from $\widetilde{\sigma}$. As a consequence, $\Delta(\sigma)$ is also symmetric around $\widetilde{\sigma}$. Further, there exist two signals $\underline{\sigma} < \overline{\sigma}$ such that actual ex post disagreement exactly equals prior disagreement, i.e., $\Delta(\underline{\sigma}) = \Delta(\overline{\sigma}) = |\mu_S - \mu_R|$. We refer to these signals as *status quo signals.* They satisfy $\underline{\sigma} < \min\{\mu_S, \mu_R\} < \max\{\mu_S, \mu_R\} < \overline{\sigma}$.

2.2 Equilibrium disclosure

We next characterize the equilibrium disclosure strategy. Full unravelling is prevented since a sender holding unfavorable information can claim to be uninformed. Hence, equilibrium features partial disclosure. Recalling that $\widetilde{\Delta}(\varnothing)$ is perceived disagreement conditional on non-disclosure, S will disclose signal $\sigma \neq \varnothing$ if and only if $\Delta(\sigma) \leq \widetilde{\Delta}(\varnothing)$. Given that $\Delta(\sigma)$ is V-shaped and symmetric around the zero disagreement signal $\widetilde{\sigma}$, the set of signals that are optimally disclosed is an interval symmetric around $\widetilde{\sigma}$, $I = [\widetilde{\sigma} - \eta, \widetilde{\sigma} + \eta]$, as illustrated in Figure 2.

At the boundaries of the disclosure interval $(\widetilde{\sigma} \pm \eta)$, S must be indifferent between disclosure and non-disclosure. The equilibrium value η^* hence satisfies

$$\underbrace{\Delta(\widetilde{\sigma} + \eta^*) = \Delta(\widetilde{\sigma} - \eta^*)}_{\text{PD if disclosing a boundary signal}} = \underbrace{\widetilde{\Delta}(\varnothing)}_{\text{PD given non-disclosure}}, \tag{7}$$

where $\widetilde{\Delta}(\varnothing)$ is a weighted average of prior disagreement and the disagreement conditional on concealed signals:

$$\begin{aligned}
\widetilde{\Delta}(\varnothing) &= E_R[\Delta(\sigma)\,|d=\varnothing,\eta] \\
&= \left[\frac{\varphi\left[F_R(\widetilde{\sigma}-\eta)\right]}{\varphi\left[F_R(\widetilde{\sigma}-\eta)+(1-F_R(\widetilde{\sigma}+\eta))\right]+1-\varphi}\right] E_R[\Delta(\sigma)\,|\sigma<\widetilde{\sigma}-\eta] \\
&\quad+\left[\frac{\varphi\left[1-F_R(\widetilde{\sigma}+\eta)\right]}{\varphi\left[F_R(\widetilde{\sigma}-\eta)+(1-F_R(\widetilde{\sigma}+\eta))\right]+1-\varphi}\right] E_R[\Delta(\sigma)\,|\sigma>\widetilde{\sigma}+\eta] \\
&\quad+\left[\frac{(1-\varphi)}{\varphi\left[F_R(\widetilde{\sigma}-\eta)+(1-F_R(\widetilde{\sigma}+\eta))\right]+1-\varphi}\right] |\mu_S-\mu_R|\,. \qquad (8)
\end{aligned}$$

Here F_R is the cumulative distribution function of $\sigma \in \mathbb{R}$ from R's ex ante perspective. The following proposition provides a characterization of equilibrium.

Proposition 1 *Let $\gamma_S \neq \gamma_R$. There exists a unique equilibrium and it features partial disclosure. In particular,*

(a) There exists a unique $\eta^* > 0$ pinned down by (7) such that S discloses signal σ if and only if it falls in the disclosure interval $I = [\widetilde{\sigma}-\eta^*, \widetilde{\sigma}+\eta^*]$.

(b) The disclosure interval I contains the status quo signals, i.e., $[\underline{\sigma}, \overline{\sigma}] \subset I$.

(c) If $\mu_S \neq \mu_R$, then the center of the disclosure interval I is located closer to the prior opinion of the more confident player.

Part (a) of the Proposition states that equilibrium disclosure satisfies an opinion corridor property: Only signals that fall within a bounded interval are disclosed. So a limited range of "facts" are disclosed, while any more extreme facts remain "off limits".

By point (b), the disclosure interval contains all signals that improve on prior disagreement as well as a limited range of signals that increase disagreement relative to the prior. This sets a lower bound on disclosure. Technically, the critical disagreement level $\Delta(\widetilde{\sigma}+\eta^*)$ defining the boundaries of the disclosure interval is strictly larger than prior disagreement $|\mu_S-\mu_R|$. The reason for this is that in equilibrium, the perceived disagreement conditional on non-disclosure is a weighted average of prior disagreement $|\mu_S-\mu_R|$ – reflecting the possibility of an uninformed sender – and the disagreement conditional on concealed signals (see (8)). Since any signal S conceals must satisfy $\Delta(\sigma) > \widetilde{\Delta}(\varnothing)$ by optimality, for such signals it should hold $\Delta(\sigma) > \widetilde{\Delta}(\varnothing) > |\mu_S-\mu_R|$.

Point (c) implies that the set of signals that are disclosed is more aligned with the prior opinion of the more confident player. Intuitively, this follows from the basic property that the more confident player adjusts his opinion less to new facts (signals) than the less confident player. It follows that signals that contradict the more confident party's prior opinion generate more disagreement than signals that contradict the less confident party's prior opinion. Hence, the zero disagreement signal $\widetilde{\sigma}$ as defined in (6), which is the center of the disclosure interval, must be closer to the prior of the more confident player.

Note that if prior opinions μ_S and μ_R are very close, then the zero disagreement signal $\widetilde{\sigma}$ is also close to these prior opinions (in particular, if $\mu_S = \mu_R = \mu$, then $\widetilde{\sigma} = \mu$). Thus, the equilibrium disclosure interval $[\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ will be centred around a point close to the prior opinions of both players. Put differently, if S and R are close in their prior opinion, then the disclosed information will tend to confirm this opinion on average. This mirrors the real-world phenomenon of "echo chambers", i.e., social groups where any information significantly deviating from the normative belief in this group is suppressed from social discourse.

Consider now the knife-edge case in which both players are equally confident. Identical prior variances imply that S and R update their opinions at the same speed. Formally, posterior means $E_S[\omega|\sigma]$ and $E_R[\omega|\sigma]$ have identical slopes with respect to $\sigma \in \mathbb{R}$. Hence, ex post disagreement $\Delta(\sigma)$ is identical for all $\sigma \in \mathbb{R}$, with $\Delta(\sigma) = \frac{\gamma_\varepsilon^2}{\gamma_\varepsilon^2 + \gamma_S^2} |\mu_S - \mu_R|$, which is strictly smaller than the prior disagreement $|\mu_S - \mu_R|$ in case prior means differ. The following can be shown.

Proposition 2 *Let $\gamma_S = \gamma_R$. Then the unique equilibrium features full disclosure.*

3 Preferences over communication partners

We have characterized the unique equilibrium of the disclosure game for any combination of priors. This allows us to discuss preferences over the prior characteristics of the communication partner from both S's and R's perspective. A receiver's preferences over different senders is determined by the informativeness of the respective equilibrium disclosure strategy. In contrast, the sender prefers receivers for which ex post perceived disagreement in equilibrium is lowest in expectation.

3.1 Receiver preferences over senders

A receiver prefers a disclosure strategy that yields a higher expected utility in the subsequent action choice problem. Below, we consider a given receiver – as characterized by a fixed prior (μ_R, γ_R) – and analyze how changes in the prior (μ_S, γ_S) of the faced sender affect the receiver's expected utility via its effect on the equilibrium disclosure strategy.[19] This captures the case of a decision maker who picks an advisor from a large pool, where candidates differ in their prior beliefs, have the same ex ante level of expertise (i.e., the same γ_ε^2 and φ) and are all perceived disagreement averse. The decision maker's objective is to maximize learning from the hired advisor.

3.1.1 Senders with different prior opinions

We first consider receiver preferences over senders with different prior opinions, μ_S, holding constant prior confidence $1/\gamma_S^2$. By the fixed order of μ_S and μ_R we mean that the ordinal

[19] While we consider separately the effect of changes in S's prior opinion and confidence, results can be combined to derive R's preference between senders who differ both in prior opinions and confidence levels.

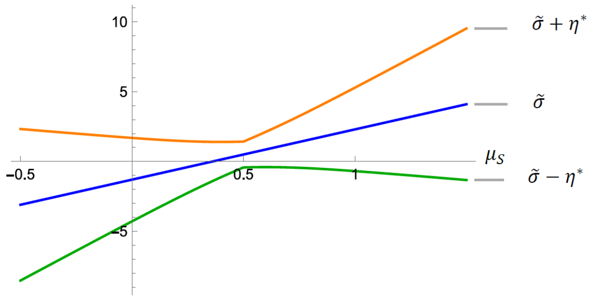

Figure 3: (*prior disagreement and equilibrium disclosure*) The graph plots the upper (yellow line) and lower (green line) bounds of the disclosure interval as well as the zero disagreement signal (blue line) as a function of the sender's prior opinion μ_S.

relation $\mu_S \geq \mu_R$ or $\mu_S \leq \mu_R$ remains the same as μ_S changes.

Proposition 3 *Let $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Fix the receiver's prior (μ_R, γ_R) as well as the sender's prior confidence $1/\gamma_S^2$, with $\gamma_R \neq \gamma_S$.*

a) Keeping the order of μ_R and μ_S fixed, the equilibrium disclosure strategy becomes more Blackwell informative when prior disagreement $|\mu_R - \mu_S|$ strictly increases.

b) Independently from the order of μ_R and μ_S, the receiver's expected utility is strictly increasing in prior disagreement $|\mu_R - \mu_S|$. I.e., the receiver strictly prefers a sender whose prior opinion is more distant from his own.

Our results are obtained under the sufficient (but not necessary) condition $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$.[20] The condition ensures that the asymmetric information problem is sufficiently severe. First, the sender's private signal is sufficiently informative, as captured by a sufficiently high signal to noise ratio $\frac{\gamma_S^2}{\gamma_\varepsilon^2}$. Second, the sender can strategically use this private information since unravelling is restricted by a sufficiently high probability of being uninformed, as captured by a sufficiently low $\frac{\varphi}{1-\varphi}$.

Under this condition, the effect of higher prior disagreement on equilibrium disclosure is illustrated in Figure 3. As can be seen, the disclosure interval widens as prior disagreement increases from zero, i.e., as μ_S moves away from $\mu_R = 0.5$ in either direction, in line with point a). Formally, making the dependence on S's prior opinion explicit, the equilibrium disclosure interval $I(\mu_S) = [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ satisfies the following ranking: For any $\mu_S'' > \mu_S' \geq \mu_R$, the disclosure interval given μ_S'' is a superset of the one given μ_S', i.e., $I(\mu_S') \subset I(\mu_S'')$, and similarly for any $\mu_S'' < \mu_S' \leq \mu_R$. That is, as μ_S increases (decreases) starting from $\mu_S = \mu_R$, equilibrium disclosure becomes more informative in the sense of Blackwell (1953), since any signal disclosed given μ_S'' is also disclosed given μ_S' but not vice-versa. Thus, when comparing two senders who are both either more optimistic than R ($\mu_S'' > \mu_S' \geq \mu_R$) or more pessimistic ($\mu_S'' < \mu_S' \leq \mu_R$), it directly follows from the Blackwell theorem that R prefers communicating with the sender whose prior disagreement $|\mu_R - \mu_S|$ is larger. This is true

[20] If this condition does not hold, then the equilibrium disclosure intervals corresponding to different levels of prior disagreement cannot generally be ranked in terms of Blackwell informativeness for a fixed order of prior means.

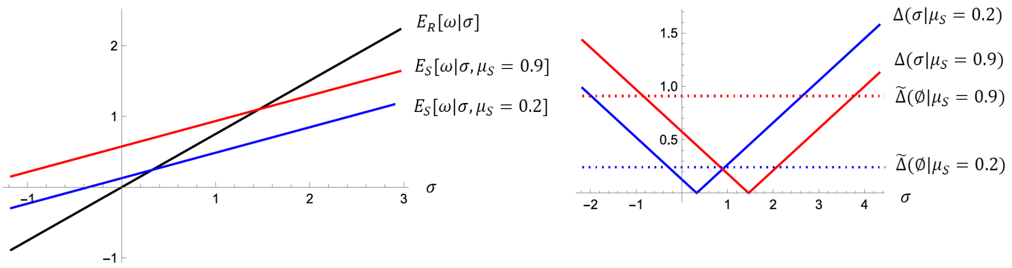

Figure 4: (*increasing prior disagreement*) The graph plots the effects of an increase in prior disagreement. The left panel illustrates the shift in the sender's posterior opinion due to an increase in μ_S from 0.2 (blue line) to 0.9 (red line) holding constant $\gamma_S = 0.3$, together with the receiver's posterior opinion (black line) for $\mu_R = 0$ and $\gamma_R = 0.7$. The right panel plots the effect of the increase in prior disagreement on posterior disagreement upon disclosure (shift from solid blue to solid red line) as well as on perceived disagreement upon non-disclosure (shift from dashed blue to dashed red line). The other parameter values are $\gamma_\varepsilon = 0.4$ and $\varphi = 0.3$.

independently of the concrete specification of R's decision problem.[21]

Instead, ranking senders with different prior opinions according to the Blackwell-informativeness of the respective equilibrium disclosure strategy is not generally possible when we compare two senders who are respectively more optimistic ($\mu_S > \mu_R$) and more pessimistic ($\mu_S < \mu_R$) than R, since in this case disclosure intervals need not satisfy an inclusion relation (cf., for example, disclosure intervals when $\mu_S = 0$ and $\mu_S = 1$ in Figure 3). In this case, characterizing R's preference over senders requires specifying R's utility function. The symmetry properties of the utility function (1) allow us to establish that R prefers higher prior disagreement also in this case (see proof of Proposition 3). A key property of equilibrium disclosure that we use here is that if $\mu'_S \neq \mu''_S$ are both equidistant from μ_R, then the respective equilibrium disclosure intervals are symmetric around μ_R.[22]

We next provide some intuition for why moving S's prior opinion away from μ_R can lead to more Blackwell informative equilibrium disclosure as illustrated in Figure 3. Intuitively, a main driving force is that an increase in prior disagreement robustly increases *perceived disagreement upon non-disclosure*, $\widetilde{\Delta}(\varnothing)$. This in turn makes non-disclosure less attractive. Indeed, recall that conditional on non-disclosure, R assigns a strictly positive probability (bounded below by $1 - \varphi$) to the event that S is uninformed, in which case actual disagreement equals prior disagreement (see the last line in (8)). Via this channel, higher prior disagreement pushes perceived disagreement conditional on non-disclosure upwards.[23]

Higher prior disagreement however not only increases (perceived) disagreement upon non-disclosure, $\widetilde{\Delta}(\varnothing)$, but it also affects actual disagreement upon disclosure, $\Delta(\sigma)$, for

[21] Note that Blackwell theorem would imply a weak increase of R's utility as the disclosure interval widens on both sides. The strict relation is additionally shown in the proof of Lemma A.8 in the online Appendix.

[22] Generally, the following condition guarantees that two disclosure intervals that are symmetric around μ_R yield the same expected utility for R: For any ω, ω', a and a' s.t. $\mu_R - \omega = \omega' - \mu_R$ and $\mu_R - a = a' - \mu_R$, it holds true that $u(\omega, a) = u(\omega', a')$.

[23] We note that this disclosure-enhancing effect of higher ex ante disagreement becomes more important the larger the ex ante likelihood that S is uninformed, $1 - \varphi$. Accordingly, the sufficient condition for the result in Proposition 3 to hold, $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, is more likely to be satisfied the larger $1 - \varphi$.

$\sigma \neq \varnothing$. Whether $\Delta(\sigma)$ is increasing or decreasing in $|\mu_R - \mu_S|$ depends on the signal σ considered. Perhaps surprisingly, for some signals higher prior disagreement leads to lower posterior disagreement. In the example shown on Figure 4, the increase in $\mu_S > \mu_R$ yields a reduced posterior disagreement for high signal values and an increased posterior disagreement for low signal values. Indeed, the divergence of posterior beliefs after getting a sufficiently high signal is mitigated if S's prior mean (and hence her posterior mean) is shifted in the direction of the updated posterior mean of R (who updates upwards at a higher speed), i.e., increases. Yet, simultaneously such shift in μ_S increases posterior disagreement for sufficiently low signals, where R's posterior is always lower.

Summarizing, an increase in prior disagreement has two effects on the sender's incentives. First, there is a robust effect that makes disclosure of any signal relatively more attractive by increasing perceived disagreement upon non-disclosure, $\widetilde{\Delta}(\varnothing)$ (see the upward shift of the dashed horizontal line in the right panel of Figure 4). Second, there is a more composite effect operating via $\Delta(\sigma)$, disagreement upon disclosure. $\Delta(\sigma)$ decreases at one end of the disclosure interval, where it incentivizes more disclosure, but $\Delta(\sigma)$ simultaneously increases at the other end of the disclosure interval, where it incentivizes less disclosure (see the shift of $\Delta(\sigma)$ at the upper and lower ends of the initial disclosure interval in the right panel of Figure 4). As we show in the proof of Lemma A.12 in online Appendix, the robust disclosure-enhancing effect dominates the latter negative effect whenever $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. In consequence, if this condition holds, the disclosure interval expands on both sides following an increase in $|\mu_R - \mu_S|$ as illustrated in the right panel of Figure 4.

3.1.2 Senders with different prior confidence levels

We next consider receiver preferences over senders with different prior confidence levels $1/\gamma_S^2$, holding constant μ_S. By the fixed order of γ_S and γ_R we mean that the relation $\gamma_S \geq \gamma_R$ or $\gamma_S \leq \gamma_R$ remains the same as γ_S changes.

Proposition 4 *Let $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Fix the receiver's prior (μ_R, γ_R) as well as the sender's prior opinion μ_S, with $\mu_S \neq \mu_R$.*

a) Keeping the order of γ_R and γ_S fixed, as the absolute difference $|\gamma_R - \gamma_S|$ strictly increases, the equilibrium disclosure strategy becomes less Blackwell informative and the receiver's expected utility strictly decreases. I.e., the receiver strictly prefers a sender whose confidence level is less distant to his own.

b) For any $\delta \in (0, \gamma_R)$, the receiver's expected utility is strictly higher if $\gamma_S^2 = \gamma_R^2 + \delta$ than if $\gamma_S^2 = \gamma_R^2 - \delta$. I.e., between two senders whose prior variances are equidistant from his own, the receiver strictly prefers the sender whose prior variance is the highest (and whose confidence is thus the lowest).

Point a) states that, conditional on $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and preserving the ordinal ranking of γ_R and γ_S, an increase in the difference in prior variances (and hence, in confidences) has exactly the opposite effect than an increase in the difference in prior means (i.e., opinions) in that it

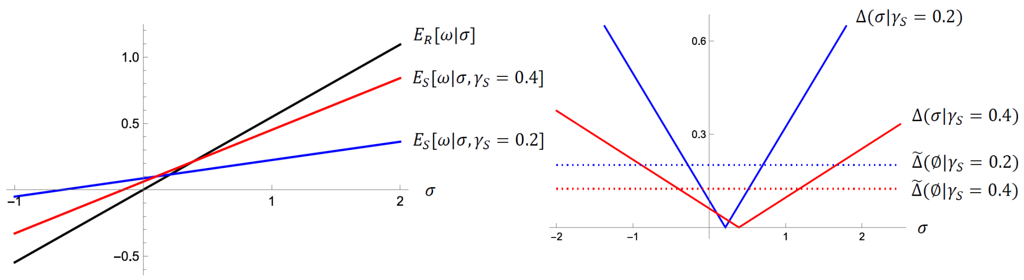

Figure 5: (*decreasing difference in prior confidence*) The graph plots the effects of a decrease in the difference of prior confidence. The left panel illustrates the shift in the sender's posterior opinion due to an increase in γ_S from 0.2 (blue line) to 0.4 (red line) holding constant $\mu_S = 0.1$, together with the receiver's posterior opinion (black line) for $\mu_R = 0$ and $\gamma_R = 0.55$. The right panel plots the effect of the decrease in the difference of prior confidence on posterior disagreement upon disclosure (shift from solid blue to solid red line) as well as on perceived disagreement upon non-disclosure (shift from dashed blue to dashed red line). The other parameter values are $\gamma_\varepsilon = 0.5$ and $\varphi = 0.5$.

reduces the informativeness of equilibrium disclosure.[24] As γ_S moves closer to γ_R the speeds of updating of the players converge to each other so that posterior disagreement conditional on disclosure remains comparatively small even for very high or low signals. Graphically, this is illustrated in the left panel of Figure 5. An increase in γ_S towards γ_R (from below) causes the slope of $E_S[\omega|\sigma]$ to increase (transition from blue to red line) and become closer to the slope of $E_R[\omega|\sigma]$ (black line). This maps into a flatter actual disagreement function $\Delta(\sigma)$ as shown in the right panel of Figure 5 (transition from solid blue to solid red line). Ceteris paribus, this makes disclosure of most signals more attractive for S and hence pushes towards a larger disclosure interval.

At the same time, as $|\gamma_R - \gamma_S|$ decreases, the above disclosure-enhancing effect is counteracted by a reduction in perceived disagreement conditional on non-disclosure $\widetilde{\Delta}(\varnothing)$ (see downward shift from the dashed blue to the dashed red line in Figure 5). This second effect makes non-disclosure relatively more attractive for S and follows from the same underlying intuition as the disclosure-enhancing effect. As implied by Proposition 4, the disclosure-enhancing effect dominates whenever $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. This condition again reflects situations where S has sufficiently precise information conditional on obtaining the signal (γ_ε^2 is sufficiently small), while at the same time she has a distinct strategic possibility not to disclose the signal as she can credibly mimic the uninformed type (once S's ex ante likelihood of being informed φ is sufficiently small).

The overall effect of a decreased difference in variances (preserving the ordinal ranking of γ_R and γ_S) is illustrated in the right panel of Figure 5. The disclosure interval widens in the sense of becoming a superset as γ_S moves closer to γ_R, so that disclosure becomes more Blackwell informative. This ultimately implies that R becomes better off.

[24] As in the case of Proposition 3, if the sufficient condition $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ does not hold, then the equilibrium disclosure intervals corresponding to different levels of S's prior confidence cannot be generally ranked in terms of Blackwell informativeness.

Point b) of Proposition 4 instead considers the receiver's preference over senders whose variances are equidistant from the receiver's variance. The overall intuition builds on the fact that the belief convergence between S and R after jointly observing an informative signal is enhanced if either of the players becomes less confident. This is particularly prominent if the absolute distance between players' prior variances remains fixed. This effect upon disclosure is only partly echoed by a fall in perceived disagreement upon non-disclosure, as the latter expression also depends on $|\mu_S - \mu_R|$. Accordingly, a symmetric shift in S's variance upwards around R's variance triggers a disclosure enhancing effect which dominates the accompanying disclosure discouraging effect.

Proposition 4 assumes unequal prior opinions, i.e., $\mu_S \neq \mu_R$. The case of $\mu_S = \mu_R$ is addressed next.

Proposition 5 *Fix the receiver's prior* (μ_R, γ_R) *as well as the sender's prior opinion* μ_S*, with* $\mu_S = \mu_R$*.*

a) The receiver is indifferent between any γ'_S, γ''_S *if both are unequal to* γ_R*.*

b) The receiver strictly prefers $\gamma_S = \gamma_R$ *over any* $\gamma'_S \neq \gamma_R$*.*

We see that if $\mu_S = \mu_R$, S's prior confidence generically does not affect R's expected utility unless there is a shift from some $\gamma'_S \neq \gamma_R$ to $\gamma_S = \gamma_R$. In particular, $\gamma_S = \gamma_R$ implies full disclosure in equilibrium, yielding a strictly higher expected utility for R than partial disclosure under $\gamma'_S \neq \gamma_R$.[25]

3.2 Sender preferences over receivers

From an ex ante perspective (i.e., before observing σ), senders care about the expected ex post perceived disagreement, i.e., $E_S[\widetilde{\Delta}(d)] = E_S[E_R[\Delta(\sigma)\,|d]]$, as pinned down by the disclosure equilibrium implied by the profile of priors under consideration. Senders prefer receivers who yield a lower value of this quantity.

3.2.1 Receivers with different prior opinions

Proposition 6 *Fix the sender's prior* (μ_S, γ_S) *as well as the receiver's prior confidence* $1/\gamma_R^2$*. Then,* $E_S[\widetilde{\Delta}(d)]$ *is strictly decreasing in prior disagreement* $|\mu_R - \mu_S|$*. I.e., ex ante the sender strictly prefers communicating with a receiver with whom she has smaller prior disagreement, independently from the order of* μ_R *and* μ_S*.*

[25] The intuition for point a) is as follows. Under equal prior opinions $\mu_S = \mu_R = \mu$, the 0-disagreement signal $\widetilde{\sigma}$ becomes equal to μ for any γ_S. Hence, a change in γ_S does not shift the intercept but only affects the slope of the posterior disagreement function $\Delta(\sigma)$. It follows that for any given signals σ_1 and σ_2, the ratio $\frac{\Delta(\sigma_1)}{\Delta(\sigma_2)}$ depends only on the signal ratio $\frac{\sigma_1}{\sigma_2}$ while being invariant to γ_S. In turn, this implies that for given η, the ratio between perceived disagreement conditional on concealed signals, $\widetilde{\Delta}(\varnothing|\eta)$, and disagreement conditional on boundary disclosed signals, $\Delta(\widetilde{\sigma} - \eta) = \Delta(\widetilde{\sigma} + \eta)$, does not change with γ_S. Thus, the equilibrium value of η solving $\widetilde{\Delta}(\varnothing|\eta) = \Delta(\widetilde{\sigma} - \eta) = \Delta(\widetilde{\sigma} + \eta)$ stays the same for all $\gamma_S \neq \gamma_R$. Then, the equilibrium disclosure interval $[\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*] = [\mu - \eta^*, \mu + \eta^*]$ (and hence, R's expected utility) does not change with γ_S.

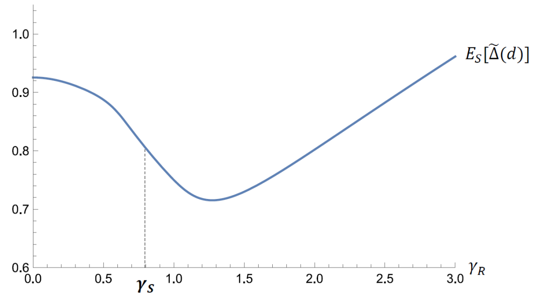

Figure 6: (*S's ex ante expected disagreement*) The graph plots S's ex ante expected perceived disagreement as a function of γ_R.

Comparing Proposition 6 with Proposition 3 we see that receiver and sender preferences over prior disagreement differ. While receivers prefer communicating with senders with whom prior disagreement in opinions is larger (under our standard sufficient condition), senders always prefer a lower value of $|\mu_R - \mu_S|$. This is intuitive since from the sender's ex ante perspective, expected posterior disagreement must decrease with prior disagreement. It is less trivial that this result holds independently of whether the reduction in prior disagreement makes equilibrium disclosure more or less informative for the receiver.

3.2.2 Receivers with different prior confidence levels

Proposition 7 *Fix the sender's prior (μ_S, γ_S) as well as the receiver's prior opinion μ_R. Then, $E_S[\widetilde{\Delta}(d)]$ is continuous in γ_R on $(0, \infty)$. Besides:*

a) If $\mu_S = \mu_R$, $E_S[\widetilde{\Delta}(d)]$ is strictly decreasing in γ_R on $(0, \gamma_S]$ and strictly increasing in γ_R on $[\gamma_S, \infty)$. I.e., ex ante the sender mostly prefers the receiver with the same prior confidence as her own.

b) If $\mu_S \neq \mu_R$, there exists $\widehat{\gamma}_R > \gamma_S$ such that $E_S[\widetilde{\Delta}(d)]$ is strictly decreasing in γ_R on $(0, \widehat{\gamma}_R]$, and strictly increasing in γ_R for γ_R sufficiently large. I.e., ex ante the sender mostly prefers the receiver(s) with a prior confidence $1/\gamma_R^2$ strictly smaller than her own.

Figure 6 shows a numerical example of the result in Proposition 7 b). $E_S[\widetilde{\Delta}(d)]$ changes non-monotonically in response to γ_R and reaches its minimum at $\gamma_R \approx 1.27 > \gamma_S = 0.8$.

Proposition 7 reveals that S's and R's preferences over the difference in prior confidence levels can be aligned in that both prefer a smaller discrepancy $|\gamma_R - \gamma_S|$ if $\gamma_R < \gamma_S$ and $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ (see Proposition 4). Yet, if $\gamma_R > \gamma_S$ (i.e., R is less confident) and $\mu_S \neq \mu_R$, then S's and R's preferences over differences in confidence levels can be misaligned. Here, keeping μ_R fixed, S's optimal conversation partner is strictly less confident than herself. In contrast, if R is less confident than S and $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, by Proposition 4 R would still always prefer a more aligned S (in terms of prior confidence).

The intuition for S's preferences over R's prior confidence is as follows. A key effect causing S to prefer similarly confident R is that this reduces the part of expected disagreement driven by different speeds of updating (call this the similarity effect). At the same time, from the ex ante perspective, S expects to receive information that will lead R to update in the direction of S's prior, thus reducing disagreement. This updating towards μ_S

will be stronger the less confident R is, i.e., the higher γ_R (call this the flexibility effect). So if $\gamma_R < \gamma_S$, an increase in γ_R implies that the two above effects (similarity and flexibility) push towards lower expected perceived disagreement. If instead $\gamma_R > \gamma_S$, an increase in γ_R implies both a more dissimilar and a less confident R so that the two above effects push in different directions in terms of expected perceived disagreement. As Proposition 7 shows, if $\mu_S \neq \mu_R$ the flexibility effect still dominates the similarity effect if $\gamma_R > \gamma_S$ for γ_R sufficiently close to γ_S.

4 Endogenous Matching

In this section, we study endogenous matching in a population featuring multiple senders and receivers, and examine implications for information transmission. We leverage our comparative statics results from the previous sections. While a general matching analysis is beyond the scope of this paper, we provide a simple example to illustrate key implications.

Previous analysis has established that depending on whether the quantity affected by matches is the difference in prior opinions or the difference in prior confidence levels, senders' and receivers' preferences are either aligned or misaligned. In particular, if matches affect the difference in prior opinions, under a simple sufficient condition senders prefer to minimize this difference while receivers prefer to maximize it.

Our matching setting is as follows. Two senders S_1, S_2 face two receivers R_1, R_2. Senders are averse to perceived disagreement and receivers' utility function is $-(a-\omega)^2$, as before. Priors are commonly known. A matching allocation specifies two bilateral matches, each sender being paired exclusively with one receiver and vice versa. Consider the following simple *endogenous matching game*. In stage 1, both senders choose a receiver to be matched with (either simultaneously or according to an exogenously given deterministic sender choice order).[26] In stage 2, senders receive a signal with probability φ and our baseline disclosure game is played within each match.[27] Contrast this with the benchmark of *a random matching game* that differs only with respect to stage 1, in which the matching is instead random with each allocation being equally likely.

For simplicity, assume identical confidence levels but heterogeneous opinions on each side, so $\gamma_{S_1} = \gamma_{S_2} = \gamma_S$ and $\gamma_{R_1} = \gamma_{R_2} = \gamma_R$. Furthermore, let $\gamma_S \neq \gamma_R$ to exclude the trivial case of full disclosure (independently of prior opinions) in stage 2. For any S_i, $i = 1, 2$, assume that receiver R_i is the one whose prior opinion is strictly closest to that of S_i, and the same is true for any R_i, $i = 1, 2$, with respect to S_i. We call a prior profile that satisfies these conditions a balanced heterogeneous means (BHM) prior profile. An example of this would be $\mu_{R_1} < \mu_{S_1} << \mu_{S_2} < \mu_{R_2}$, where $<<$ means much lower than. This would correspond to a polarized population featuring two distinct groups with significantly different opinions.[28]

[26] If both senders choose the same receiver under simultaneous decisions, the matching is determined randomly.

[27] We think of partner choice as the creation of long-term communication channels, like "friendships" on Facebook, which justifies the assumed timing of information arrival.

[28] Political polarisation between Democrats and Republicans in the current US would be a potential example. Counterparts in Europe would be the polarisation between Brexiteers and Remainers in the UK.

We obtain the following result:

Proposition 8 *Let* $\frac{\gamma_{S_i}^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ *for* $i = 1, 2$*. Then, given a BHM prior profile, each receiver obtains a strictly lower expected utility under endogenous matching than under random matching.*

Under endogenous matching, senders choose the receiver who is closest in prior opinion to themselves (as this is a strictly dominant strategy by Proposition 6). Across receivers, this allocation is Pareto dominated by the other allocation by Proposition 3.[29] Thus, from a policy perspective, an implication is that when dealing with communities with heterogeneous prior opinions and similar levels of prior confidence, policy makers should prefer random matching mechanisms over sender-driven selective matching if they are interested in the informativeness of communication between agents. Another adverse effect of endogenous matching for social learning is that receivers will be receiving only signals which are close to their prior opinion on average, as discussed at the end of Section 2.2 (the "echo chamber" effect).

For prior profiles that do not satisfy the above balancedness condition, our results on the effect of S-driven matching are weaker. Transitioning from endogenous to random matching does not generally cause any Pareto improvement or worsening across receivers. Consider, for example, the case where both senders are closest (in prior opinion) to the same receiver (say R_2) while both receivers are closest (in prior opinion) to the same sender (say S_1). An example of this would be that $\mu_{R_1} < \mu_{R_2} < \mu_{S_1} < \mu_{S_2}$. If S_1 moves first, then she chooses R_2, who is thus matched with his closest sender (in opinions) while R_1 is matched with his most distant sender. Hence, R_2 would prefer random matching as it gives him a chance to be matched with his more distant sender. R_1 instead prefers endogenous matching as it guarantees a match with his most distant sender. If S_2 picks a partner first, then receivers' preferences over matching mechanisms are simply reversed.[30]

5 Robustness

In our preceding analysis, we measure disagreement in beliefs as the difference in agents' prior means. In certain situations, more general notions of disagreement that also take into account differences in variances may be relevant. As an illustration, consider two scenarios. In the first scenario, agents have identical posterior means but their confidence levels differ starkly. In the second scenario, agents have slightly different posterior means but the same confidence. Arguably, in many contexts, the difference in beliefs embodied in the second scenario is less consequential than the one captured by the first scenario. A possible example

[29] In an N-senders and N-receivers population that satisfies a generalised version of the BHM condition, the same statement holds true. Across receivers, the matching allocation chosen by senders is Pareto-dominated by any other allocation.

[30] Under simultaneous decision making, the endogenous and random matching will be equivalent in this case.

would be an investment problem, where team members assess the expected return and the riskiness of different assets and need to agree on a portfolio.

In the following, we study how our qualitative insights extend to settings in which the measure of disagreement takes into account differences in both posterior opinion as well as posterior confidence. We introduce a standard measure of distance between distributions, the Kullback-Leibler divergence (KL divergence), also known as Relative Entropy. We show that when applied to our disclosure game with a perceived disagreement averse sender, the measure yields the same qualitative predictions as the disagreement measure (2) used in our main analysis.

The KL divergence is a directional measure. For two distributions over Ω with pdfs p and q, the KL divergence from q to p is given by:

$$D_{KL}(p,q) = \int_{\Omega} \left[\log(p(\omega)) - \log(q(\omega))\right] p(\omega) d\omega. \tag{9}$$

Given two Gaussian distributions $N(\mu_i, \gamma_i^2)$ and $N(\mu_j, \gamma_j^2)$, the measure yields a simple algebraic expression which is omitted here. As we measure R's perception of disagreement, it is natural to take the receiver's distribution as reference point. Denoting by $\widetilde{g}_i(\cdot\,|\sigma)$ the posterior of player $i \in \{S, R\}$ given signal σ, we thus use the KL divergence from $\widetilde{g}_S(\cdot\,|\sigma)$ to $\widetilde{g}_R(\cdot\,|\sigma)$ as the measure of disagreement, which is redefined as follows:

$$\Delta(\sigma) = D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma)). \tag{10}$$

The ex post perceived disagreement is still given by (2) and as in the main section, S chooses disclosure so as to minimize this quantity.

Proposition 9 *Let $\gamma_S \neq \gamma_R$. If $\Delta(\sigma)$ is given as in (10), then any equilibrium features a finite non-degenerate disclosure interval $I = [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ such that S discloses σ if and only if $\sigma \in I$, where $\widetilde{\sigma}$ is the unique signal for which the posterior means of S and R are equal.*

Equilibrium disclosure thus takes a very similar form as when defining $\Delta(\sigma)$ as the difference in posterior means. It features a disclosure interval centred around $\widetilde{\sigma}$, the signal ensuring zero disagreement in posterior means. One can show that the same statement holds true if we consider alternative measures of perceived disagreement such as the KL divergence from $\widetilde{g}_R$ to $\widetilde{g}_S$, the symmetric KL divergence, or the so-called Bhattacharyya distance.[31] The key fact is that under all these measures, $\Delta(\sigma)$ is U-shaped and symmetric around $\widetilde{\sigma}$. Hence, the construction of equilibrium follows the same logic as under our original definition of $\Delta(\sigma)$.

[31] See Proposition A.1 in the online Appendix.

6 The value of public information

We conclude with a simple alternative model which captures the team effort dimension of conversations and also provides further intuition behind our finding that differences in opinion positively affect information sharing. Beyond the strategic disclosure approach, one can think of conversations as a process which endogenously generates an informative public signal. As we show, disagreement aversion has interesting implications for the incentive to engage in such a process.

Consider a pair of disagreement averse players (1 and 2) who have to decide whether or not to engage in costless conversation and thereby acquire a public signal. Priors and the information structure are as in the main setup. A signal is generated if and only if both players decide to participate in conversation ("*it takes two to tango*"). If no signal is generated, players' utility equals $-|\mu_1 - \mu_2|$. If a signal $\sigma \in \mathbb{R}$ is generated, players' ex post utility is $-\Delta(\sigma)$, so i's ex ante expected utility from observing the signal is $-E_i[\Delta(\sigma)]$. For i, the ex ante value of observing the signal is thus:

$$V_i(\mu_1, \mu_2, \gamma_1, \gamma_2) = |\mu_1 - \mu_2| - E_i[\Delta(\sigma)]. \tag{11}$$

Clearly, there exists an equilibrium in which the public signal is generated if and only if $\min\{V_1, V_2\} \geq 0$.

Proposition 10 *Fix μ_1, γ_1, γ_2, γ_ε and assume $\gamma_1 \neq \gamma_2$.*

a) There exists $\lambda > 0$ such that for $|\mu_1 - \mu_2| < \lambda$, $V_1(\mu_1, \mu_2, \gamma_1, \gamma_2)$ and $V_2(\mu_1, \mu_2, \gamma_1, \gamma_2)$ are both strictly negative.

b) V_1 and V_2 are strictly increasing and become positive as the distance between μ_1 and μ_2 increases.

Point a) indicates that players both assign negative ex ante value to observing the signal if prior opinions are close enough. For an intuition, note that if $\mu_1 = \mu_2$, any signal realization (other than the zero-measure signal $\widetilde{\sigma} = \mu_1 = \mu_2$) leads to a strictly positive posterior disagreement (as $\gamma_1 \neq \gamma_2$), which is strictly larger than the zero prior disagreement. By continuity of the V functions, the same is still true if μ_2 slightly deviates from μ_1. When prior opinions are too close, players' negative attitude towards information is well summarized by the expression "*don't rock the boat*".

Point b) states that larger prior disagreement leads to larger gains from jointly acquiring a public signal, ultimately guaranteeing that the signal has positive ex ante value for both players starting from a certain prior disagreement level, as illustrated in Figure 7. Thus, heterogeneity in prior opinions is helpful to stimulate communication also in this setting.

The intuition for point b) is as follows. Consider the impact of a change in μ_2 on V_1, keeping μ_1 fixed. By definition (11), the total effect of a change in μ_2 on V_1 consists of two partial effects: a change in the prior disagreement $|\mu_1 - \mu_2|$ and a change in $E_1 |\Delta(\sigma)| = \int_{-\infty}^{\infty} \Delta(\sigma) f_1(\sigma) d\sigma$. Clearly, the absolute magnitude of the first effect (in terms of the corresponding derivative) is 1. At the same time, the absolute magnitude of the second effect is less than 1 since the effect of prior disagreement on posterior disagreement $\Delta(\sigma)$

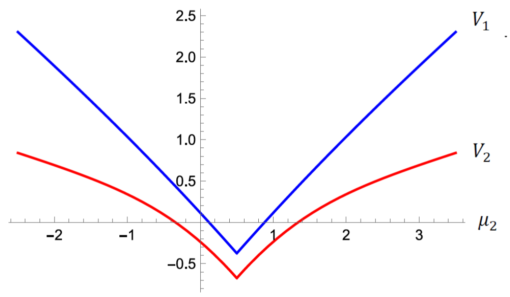

Figure 7: (*prior disagreement and value of public signal*) The graph plots the ex ante values of acquiring a public signal for players 1 and 2 as a function of the prior opinion μ_2.

(for any given σ) is alleviated by learning. For instance, if the signal is very precise, $\Delta(\sigma)$ is very close to 0 (except for the extreme values of σ) and hence does not react much to a change in μ_2. Algebraically, this is reflected in the fact that $\left|\frac{d\Delta(\sigma)}{d\mu_2}\right| = \left|\frac{\gamma_\varepsilon^2}{\gamma_2^2+\gamma_\varepsilon^2}\right| < 1$. It follows that the direct positive effect of μ_2 on V_1 through changing the prior disagreement $|\mu_1 - \mu_2|$ always dominates the negative effect through changing $E_1 |\Delta(\sigma)|$. Thus, V_1 increases as μ_2 moves away from μ_1. The same intuition explains the effect of μ_2 on V_2, keeping μ_1 fixed.

7 Conclusion

This paper introduces a new type of belief-dependent preferences capturing an aversion to perceived disagreement, which is well documented in social psychology. Our analysis has identified a range of implications for strategic communication and patterns of social matching. First, for a large set of parameter values, information disclosure is relatively low when communication partners share similar prior opinions, while similar prior confidence levels instead have a positive effect on disclosure. Second, very confident individuals are less likely to be exposed to contradictory information than less confident individuals. Finally, ceteris paribus, informed individuals prefer to talk to individuals whose opinion is very close to their own, which is counterproductive from an information sharing perspective.

Further empirical and theoretical research is called upon to test and refine the theory outlined in this paper. In a first step, experimental work should aim at deepening our understanding of the preference itself. How is the aversion affected by social context (e.g., whether S and R are friends or colleagues)? How do individuals quantify perceived disagreement and how exactly does it affect their utility? In a second step, experiments could test qualitative implications of the theory in terms of selective disclosure and matching patterns. From a theoretical perspective, perceived disagreement aversion could be studied within more general setups, for example featuring many senders and receivers, dynamic interaction, or richer matching processes.

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2016). Fragility of asymptotic agreement under bayesian learning. *Theoretical Economics 11*(1), 187–225.

Acemoglu, D., V. Chernozhukov, M. Yildiz, et al. (2007). Learning and disagreement in an uncertain world. Technical report, Collegio Carlo Alberto.

Amershi, A. H. (1988). Blackwell informativeness and sufficient statistics with applications to financial markets and multiperson agencies. In *Economic analysis of information and contracts*, pp. 25–93. Springer.

Andreoni, J. and T. Mylovanov (2012). Diverging opinions. *American Economic Journal: Microeconomics 4*(1), 209–232.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American 193*(5), 31–35.

Baliga, S., E. Hanany, and P. Klibanoff (2013). Polarization and ambiguity. *The American Economic Review 103*(7), 3071–3083.

Banerjee, A. and R. Somanathan (2001). A simple model of voice. *The Quarterly Journal of Economics 116*(1), 189–227.

Battigalli, P. and M. Dufwenberg (2009). Dynamic psychological games. *Journal of Economic Theory 144*(1), 1–35.

Battigalli, P. and M. Dufwenberg (2022, September). Belief-dependent motivations and psychological game theory. *Journal of Economic Literature 60*(3), 833–82.

Bénabou, R. (2012). Groupthink: Collective delusions in organizations and markets. *The Review of Economic Studies 80*, rds030.

Benabou, R. and G. Laroque (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *The Quarterly Journal of Economics 107*(3), 921–958.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics 24*(2), 265–272.

Buechel, B., T. Hellmann, and S. Klößner (2015). Opinion dynamics and wisdom under conformity. *Journal of Economic Dynamics and Control 52*, 240–257.

Bursztyn, L., G. Egorov, and S. Fiorin (2020). From extreme to mainstream: The erosion of social norms. *The American Economic Review 110*(11), 3522–48.

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

Caplin, A. and J. Leahy (2004). The supply of information by a concerned expert. *The Economic Journal 114*(497), 487–505.

Che, Y.-K. and N. Kartik (2009). Opinions as incentives. *Journal of Political Economy 117*(5), 815–860.

Dandekar, P., A. Goel, and D. T. Lee (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences 110*(15), 5791–5796.

Deutsch, M. and H. B. Gerard (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology 51*(3), 629.

Dixit, A. K. and J. W. Weibull (2007). Political polarization. *Proceedings of the National Academy of Sciences 104*(18), 7351–7356.

Domínguez, D., F. Juan, S. A. Taing, and P. Molenberghs (2016). Why do some find it hard to disagree? an fmri study. *Frontiers in Human Neuroscience 9*, 718.

Dorison, C. A., J. A. Minson, and T. Rogers (2019). Selective exposure partly relies on faulty affective forecasts. *Cognition 188*, 98–107.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research 23*(1), 123–145.

Ely, J., A. Frankel, and E. Kamenica (2015). Suspense and surprise. *Journal of Political Economy 123*(1), 215–260.

Ely, J. C. and J. Välimäki (2003). Bad reputation. *The Quarterly Journal of Economics 118*(3), 785–814.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, Il: Row, Peterson.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of Political Economy 114*(2), 280–316.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday Anchor Books.

Golman, R., G. Loewenstein, K. O. Moene, and L. Zarri (2016). The preference for belief consonance. *The Journal of Economic Perspectives 30*(3), 165–187.

Golub, B. and M. O. Jackson (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics 127*(3), 1287–1338.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics 24*(3), 461–483.

Jung, W.-O. and Y. K. Kwon (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research 26*(1), 146–153.

Kartik, N., F. X. Lee, and W. Suen (2021). Information validates the prior: A theorem on bayesian updating and applications. *American Economic Review: Insights 3*(2), 165–82.

Kőszegi, B. (2006). Emotional agency. *The Quarterly Journal of Economics 121*(1), 121–155.

Lazarsfeld, P. F. and R. K. Merton (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society 18*(1), 18–66.

Levy, G. (2007). Decision making in committees: Transparency, reputation, and voting rules. *The American Economic Review 97*(1), 150–168.

Levy, G. and R. Razin (2015). Correlation neglect, voting behavior, and information aggregation. *The American Economic Review 105*(4), 1634–45.

Lyons, J., A. E. Sokhey, S. D. McClurg, and D. Seib (2016). Personality, interpersonal disagreement, and electoral information. *The Journal of Politics 78*(3), 806–821.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics 12*(2), 380–391.

Morris, S. (1995). The common prior assumption in economic theory. *Economics & Philosophy 11*(2), 227–253.

Morris, S. (2001). Political correctness. *Journal of Political Economy 109*(2), 231–265.

Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy.* Cambridge University Press.

Newcomb, T. M. (1961). *The acquaintance process.* Holt, Rinehart & Winston.

Ottaviani, M. and P. N. Sørensen (2006a). Reputational cheap talk. *The Rand Journal of Economics 37*(1), 155–175.

Ottaviani, M. and P. N. Sørensen (2006b). The strategy of professional forecasting. *Journal of Financial Economics 81*(2), 441–466.

Pennington, N. and K. L. Winfrey (2021). Engaging in political talk on facebook: Investigating the role of interpersonal goals and cognitive engagement. *Communication Studies 72*(1), 100–114.

Pew Research Center (2018). More now say it's 'stressful' to discuss politics with people they disagree with. *http://www.people-press.org/2018/11/05/more-now-say-its-stressful-to-discuss-politics-with-people-they-disagree-with/*.

Prendergast, C. (1993). A theory of "yes men". *The American Economic Review 83*(4), 757–770.

Prentice, D. A. and D. T. Miller (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology 64*(2), 243.

Sethi, R. and M. Yildiz (2016). Communication with unknown perspectives. *Econometrica 84*(6), 2029–2069.

Sobel, J. (1985). A theory of credibility. *Review of Economic Studies 52*, 557–573.

Sobel, J. (2013). Giving and receiving advice. In M. Acemoglu, D. Arellano and E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress*, pp. 305–341. New York: Cambridge University Press.

Sunstein, C. R. (2007). *Republic. com 2.0.* Princeton University Press.

Visser, B. and O. H. Swank (2007). On committees of experts. *The Quarterly Journal of Economics 122*(1), 337–372.

Yildiz, M. (2004, 02). Waiting to Persuade. *The Quarterly Journal of Economics 119*(1), 223–248.

8 Appendix (omitted proofs)

A.1 Notational conventions

Throughout this appendix, we use the following notation.

Denote by $f_i(\sigma)$ the probability density functions of the distribution of S's signal σ in the eyes of player i, conditional on $\sigma \neq \varnothing$. Denote by $F_i(\sigma)$ the respective cumulative distribution function.

Besides, we define the following functions (where the agreement signal $\widetilde{\sigma}$ is given by (6)):

$$\sigma_l(\eta) \equiv \widetilde{\sigma} - \eta, \tag{A.1}$$
$$\sigma_h(\eta) \equiv \widetilde{\sigma} + \eta, \tag{A.2}$$
$$F_l(\eta) \equiv F_R(\sigma_l(\eta)), \tag{A.3}$$
$$F_h(\eta) \equiv F_R(\sigma_h(\eta)), \tag{A.4}$$
$$f_l(\eta) \equiv f_R(\sigma_l(\eta)), \tag{A.5}$$
$$f_h(\eta) \equiv f_R(\sigma_h(\eta)), \tag{A.6}$$
$$E_l(\eta) \equiv E_R[\sigma|\sigma < \sigma_l(\eta)] = \frac{1}{F_l(\eta)} \int_{-\infty}^{\sigma_l(\eta)} \sigma f_R(\sigma) d\sigma, \tag{A.7}$$
$$E_h(\eta) \equiv E_R[\sigma|\sigma > \sigma_h(\eta)] = \frac{1}{1 - F_h(\eta)} \int_{\sigma_h(\eta)}^{\infty} \sigma f_R(\sigma) d\sigma, \tag{A.8}$$
$$\Delta_0 \equiv |\mu_S - \mu_R|, \tag{A.9}$$
$$\widetilde{\Delta}(\varnothing|\eta) \equiv E_R[\Delta(\sigma)|\sigma \notin [\sigma_l(\eta), \sigma_h(\eta)]]. \tag{A.10}$$

Note, in particular, that $\widetilde{\Delta}(\varnothing|\eta^*)$ is the expected perceived disagreement conditional on non-disclosure in the equilibrium defined by Proposition 1. The argument η in the above functions will be mostly suppressed below for notational convenience.

A.2 Equilibrium disclosure

A.2.1 Preliminaries

Lemma A.1 *For any $x \in \mathbb{R}$ and $i = S, R$ it holds*

$$E_i[\sigma|\sigma < x] = \mu_i - (\gamma_i^2 + \gamma_\varepsilon^2)\frac{f_i(x)}{F_i(x)},$$
$$E_i[\sigma|\sigma > x] = \mu_i + (\gamma_i^2 + \gamma_\varepsilon^2)\frac{f_i(x)}{1 - F_i(x)}.$$

Proof. From player i's perspective, $\sigma \in \mathbb{R}$ is distributed normally with mean μ_i and variance $\gamma_i^2 + \gamma_\varepsilon^2$. Then, using the formula for the mean of the truncated normal distribution, we obtain

$$E_i[\sigma|\sigma < x] = \mu_i - \sqrt{\gamma_i^2 + \gamma_\varepsilon^2} \frac{\phi\left(\frac{x-\mu_i}{\sqrt{\gamma_i^2+\gamma_\varepsilon^2}}\right)}{\Phi\left(\frac{x-\mu_i}{\sqrt{\gamma_i^2+\gamma_\varepsilon^2}}\right)} \tag{A.11}$$
$$= \mu_i - (\gamma_i^2 + \gamma_\varepsilon^2)\frac{f_i(x)}{F_i(x)}, \tag{A.12}$$

where ϕ (Φ) is the pdf (cdf) of the standard normal distribution.

Similarly,

$$
\begin{aligned}
E_i[\sigma|\sigma > x] &= \mu_i + \sqrt{\gamma_i^2 + \gamma_\varepsilon^2}\frac{\phi\left(\frac{x-\mu_i}{\sqrt{\gamma_i^2+\gamma_\varepsilon^2}}\right)}{1 - \Phi\left(\frac{x-\mu_i}{\sqrt{\gamma_i^2+\gamma_\varepsilon^2}}\right)} \qquad \text{(A.13)}\\
&= \mu_i + (\gamma_i^2 + \gamma_\varepsilon^2)\frac{f_i(x)}{1 - F_i(x)}. \qquad \text{(A.14)}
\end{aligned}
$$

■

Lemma A.2 *Let $\gamma_S \neq \gamma_R$. The disclosure strategy in any equilibrium is defined by a finite interval $I(\eta^*) = [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ with $\eta^* > 0$ such that S discloses σ iff $\sigma \in I(\eta^*)$.*

Proof.
Step 1. In this step, we prove that for sufficiently high signals it holds $\Delta(\sigma) > \Delta_0$, i.e., disclosure increases disagreement relative to the prior level of disagreement. Recall that the perceived disagreement is given by $\widetilde{\Delta}(d) = E_R[\Delta(\sigma)\,|d]$, where $\Delta(\sigma) = |E_S[\omega\,|\sigma] - E_R[\omega\,|\sigma]|$. Let us denote

$$
D(\sigma) = E_S[\omega\,|\sigma] - E_R[\omega\,|\sigma], \qquad \text{(A.15)}
$$

so that $\Delta(\sigma) = |D(\sigma)|$. By (4) we have that for any $\sigma \in \mathbb{R}$,

$$
\begin{aligned}
D(\sigma) &= \left(\frac{\frac{1}{\gamma_S^2}}{\frac{1}{\gamma_S^2} + \frac{1}{\gamma_\varepsilon^2}}\right)\mu_S + \left(\frac{\frac{1}{\gamma_\varepsilon^2}}{\frac{1}{\gamma_S^2} + \frac{1}{\gamma_\varepsilon^2}}\right)\sigma \\
&\quad - \left(\frac{\frac{1}{\gamma_R^2}}{\frac{1}{\gamma_R^2} + \frac{1}{\gamma_\varepsilon^2}}\right)\mu_R - \left(\frac{\frac{1}{\gamma_\varepsilon^2}}{\frac{1}{\gamma_R^2} + \frac{1}{\gamma_\varepsilon^2}}\right)\sigma \\
&= \left(\frac{\gamma_S^2}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2 + \gamma_\varepsilon^2}\right)\sigma + \left(\frac{\gamma_\varepsilon^2\mu_S}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2 + \gamma_\varepsilon^2}\right). \qquad \text{(A.16)}
\end{aligned}
$$

Thus, $D(\sigma)$ is a linear function of $\sigma \in \mathbb{R}$, so that it has a unique root in $\mathbb{R}$ (once $\gamma_S \neq \gamma_R$). Consequently, $\Delta(\sigma) = |D(\sigma)|$ is V-shaped in $\sigma \in \mathbb{R}$, with the minimum value of 0 and tending to infinity as σ moves away from its unique root $\widetilde{\sigma}$. It follows immediately that for sufficiently high signals it holds $\Delta(\sigma) > \Delta_0$ (the same also holds for sufficiently low signals).

Step 2. Let us show that any equilibrium under $\gamma_S \neq \gamma_R$ features finite $\eta^* > 0$ such that S discloses σ if and only if $\sigma \in [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ (recall that $\widetilde{\sigma}$ is implicitly given by $D(\widetilde{\sigma}) = 0$).

First, note that full disclosure of all signals (FD) is never an equilibrium under $\gamma_S \neq \gamma_R$. Indeed, in this case $\widetilde{\Delta}(\varnothing) = \Delta_0$, as non-disclosure happens if and only if S is uninformed (which is inferred by R in equilibrium). Then, S would have an incentive to deviate by concealing all signals σ such that $\Delta(\sigma) > \Delta_0$, which exist by Step 1.

Next, let us show that $\widetilde{\Delta}(\varnothing)$ must be finite and strictly positive. The fact that $\widetilde{\Delta}(\varnothing)$ should be finite can also be shown by contradiction. Suppose indeed that $\widetilde{\Delta}(\varnothing)$ is not finite. Then there would exist the FD-equilibrium, as S would always favour disclosing over not disclosing. But the FD-equilibrium cannot exist, as shown above.

The fact that $\widetilde{\Delta}(\varnothing)$ must be strictly positive is proved as follows. Recall that by definition it cannot be strictly negative. Suppose by contradiction that $\widetilde{\Delta}(\varnothing) = 0$. Then, since by (A.16) for any non-empty signal, $\Delta(\sigma) = |D(\sigma)| > 0$ if and only if $\sigma \neq \widetilde{\sigma}$ (under $\gamma_S \neq \gamma_R$), in equilibrium all signals must be concealed other than the signal $\widetilde{\sigma}$. This implies that after non-disclosure, R acknowledges that with a positive probability S holds a signal different from $\widetilde{\sigma}$. But for any such signal σ, given $\gamma_S \neq \gamma_R$, $\Delta(\sigma) > 0$. It follows immediately that $\widetilde{\Delta}(\varnothing) = E_R[\Delta(\sigma)|d = \varnothing] > 0$, which yields a contradiction.

Consider thus the only possible type of equilibrium with disclosure rule featuring a strictly positive and finite $\widetilde{\Delta}(\varnothing)$. In this case, given that, by (A.16), $\Delta(\sigma) = |D(\sigma)|$ as a function of $\sigma \in \mathbb{R}$ is symmetric around $\widetilde{\sigma}$ and linearly decreasing in σ for $\sigma \leq \widetilde{\sigma}$ (where $\Delta(\widetilde{\sigma}) = 0$), there exists a unique $\eta^* > 0$ (for given equilibrium level of $\widetilde{\Delta}(\varnothing)$) such that $\Delta(\widetilde{\sigma} - \eta^*) = \Delta(\widetilde{\sigma} + \eta^*) = \widetilde{\Delta}(\varnothing)$. The latter implies that for any $\sigma \in \mathbb{R}$, $\Delta(\sigma) \leq \widetilde{\Delta}(\varnothing)$ if and only if $\sigma \in [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$. It follows immediately that the equilibrium disclosure set is given by $[\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$. ■

Lemma A.3 *Let $\gamma_S \neq \gamma_R$. A necessary and sufficient condition for $I(\eta) = [\widetilde{\sigma} - \eta, \widetilde{\sigma} + \eta]$ to be an equilibrium disclosure interval at $\eta > 0$ is $\widetilde{\Delta}(\varnothing|\eta) = \Delta(\widetilde{\sigma} + \eta)$.*

Proof. Let us prove the necessity of this condition. Assume that the equilibrium disclosure interval is given by $I(\eta') = [\widetilde{\sigma} - \eta', \widetilde{\sigma} + \eta']$ for some $\eta' > 0$, so that the perceived disagreement conditional on non-disclosure is given by $\widetilde{\Delta}(\varnothing|\eta')$ defined in (A.10). Then, S has incentive to disclose after obtaining a non-empty signal iff $\sigma \in I(\eta')$ or, equivalently,

$$\Delta(\sigma) \leq \widetilde{\Delta}(\varnothing|\eta') \text{ iff } \sigma \in [\widetilde{\sigma} - \eta', \widetilde{\sigma} + \eta'].$$

Since by (A.16), $\Delta(\sigma)$ as a function of $\sigma \in \mathbb{R}$ is symmetrically V-shaped around $\widetilde{\sigma}$ (with $\Delta(\widetilde{\sigma}) = 0$) if $\gamma_S \neq \gamma_R$, this implies that S is indifferent between disclosing and not disclosing at both $\sigma_l = \widetilde{\sigma} - \eta'$ and $\sigma_h = \widetilde{\sigma} + \eta'$, i.e.,

$$\widetilde{\Delta}(\varnothing|\eta') = \Delta(\widetilde{\sigma} - \eta') = \Delta(\widetilde{\sigma} + \eta'). \tag{A.17}$$

Let us show the sufficiency of this condition, i.e., if $\widetilde{\Delta}(\varnothing|\eta) = \Delta(\widetilde{\sigma} + \eta)$ for some $\eta > 0$ while $\gamma_S \neq \gamma_R$, then there exists an equilibrium with disclosure interval $I = [\widetilde{\sigma} - \eta, \widetilde{\sigma} + \eta]$. Assume $\widetilde{\Delta}(\varnothing|\eta') = \Delta(\widetilde{\sigma} + \eta')$ for some $\eta' > 0$ while $\gamma_S \neq \gamma_R$. Since $\Delta(\sigma)$ as a function of $\sigma \in \mathbb{R}$ has a symmetric V-shape around $\widetilde{\sigma}$ if $\gamma_S \neq \gamma_R$, we have that for any non-empty signal σ,

$$\Delta(\sigma) \leq \Delta(\widetilde{\sigma} + \eta') \text{ iff } \sigma \in [\widetilde{\sigma} - \eta', \widetilde{\sigma} + \eta'].$$

This together with assumed $\Delta(\widetilde{\sigma} + \eta') = \widetilde{\Delta}(\varnothing|\eta')$ implies that for any non-empty signal σ,

$$\Delta(\sigma) \leq \widetilde{\Delta}(\varnothing|\eta') \text{ iff } \sigma \in [\widetilde{\sigma} - \eta', \widetilde{\sigma} + \eta'].$$

Consequently, there exists an equilibrium where S discloses σ if and only if $\sigma \in [\widetilde{\sigma} - \eta', \widetilde{\sigma} + \eta']$. ■

Lemma A.4 *Let $\gamma_S \neq \gamma_R$. A necessary and sufficient condition for $I(\eta) = [\widetilde{\sigma} - \eta, \widetilde{\sigma} + \eta]$ to be an equilibrium disclosure interval at $\eta > 0$ is $\tau(\eta) = 0$, where*

$$\tau(\eta) = \varphi \begin{pmatrix} F_l(\sigma_l - \mu_R) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1 - F_h)(\mu_R - \sigma_h) \end{pmatrix} - (1 - \varphi) \left(\sigma_h + \frac{k_2 - \Delta_0}{k_1} \right) \tag{A.18}$$

with

$$
\begin{aligned}
k_1 &= sgn\ (\gamma_S - \gamma_R)\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right), \\
k_2 &= sgn\ (\gamma_S - \gamma_R)\left(\frac{\gamma_\varepsilon^2\mu_S}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right).
\end{aligned}
$$

Proof. In what follows, we show that the necessary and sufficient equilibrium condition stated in Lemma A.3, $\Delta(\widetilde{\sigma}+\eta) = \widetilde{\Delta}(\varnothing|\eta)$, is equivalent to $\tau(\eta) = 0$.

By (A.10), we have

$$
\begin{aligned}
\widetilde{\Delta}(\varnothing|\eta) &= E_R[\Delta(\sigma)|\sigma \notin [\sigma_l, \sigma_h]] \\
&= \Pr[\sigma < \sigma_l]E_R[\Delta(\sigma)\,|\sigma < \sigma_l] + \Pr[\sigma > \sigma_h]E_R[\Delta(\sigma)\,|\sigma > \sigma_h] \\
&\quad + \Pr[\sigma = \varnothing]\Delta_0 \\
&= \frac{\varphi F_l}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\frac{1}{F_l}\int_{-\infty}^{\sigma_l} \Delta(\sigma) f_R(\sigma)d\sigma \\
&\quad + \frac{\varphi(1-F_h)}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\frac{1}{1-F_h}\int_{\sigma_h}^{\infty} \Delta(\sigma) f_R(\sigma)d\sigma \\
&\quad + \frac{1-\varphi}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\Delta_0 \\
&= \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\begin{pmatrix} \varphi\int_{-\infty}^{\sigma_l} \Delta(\sigma) f_R(\sigma)d\sigma \\ +\varphi\int_{\sigma_h}^{\infty} \Delta(\sigma) f_R(\sigma)d\sigma \\ +(1-\varphi)\Delta_0 \end{pmatrix}. \qquad \text{(A.19)}
\end{aligned}
$$

Note that by (A.15) and (A.16), for $\sigma \neq \varnothing$, $\Delta(\sigma)$ can be represented as

$$
\Delta(\sigma) = |D(\sigma)| = \begin{cases} k_1\sigma + k_2 \text{ if } \sigma \geq \widetilde{\sigma}, \\ -k_1\sigma - k_2 \text{ if } \sigma < \widetilde{\sigma}, \end{cases} \qquad \text{(A.20)}
$$

where

$$
k_1 = \text{sgn}\ (\gamma_S - \gamma_R)\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right), \qquad \text{(A.21)}
$$

$$
k_2 = \text{sgn}\ (\gamma_S - \gamma_R)\left(\frac{\gamma_\varepsilon^2\mu_S}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right). \qquad \text{(A.22)}
$$

Substituting this into (A.19) we obtain

$$
\begin{aligned}
\widetilde{\Delta}(\varnothing|\eta) &= \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\begin{pmatrix} -\varphi k_1\int_{-\infty}^{\sigma_l} \sigma f_R(\sigma)d\sigma - \varphi F_l k_2 \\ +\varphi k_1\int_{\sigma_h}^{\infty} \sigma f_R(\sigma)d\sigma + \varphi(1-F_h)k_2 \\ +(1-\varphi)\Delta_0 \end{pmatrix} \\
&= \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\begin{pmatrix} -\varphi k_1 F_l E_l - \varphi F_l k_2 \\ +\varphi k_1(1-F_h)E_h \\ +\varphi(1-F_h)k_2 + (1-\varphi)\Delta_0 \end{pmatrix}. \qquad \text{(A.23)}
\end{aligned}
$$

This together with (A.20) implies that the equilibrium condition stated in Lemma A.3,

$\Delta(\sigma_h) = \widetilde{\Delta}(\varnothing|\eta)$, is equivalent to:

$$0 = \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\begin{pmatrix} -\varphi k_1 F_l E_l - \varphi F_l k_2 + \varphi k_1 (1-F_h) E_h \\ +\varphi(1-F_h)k_2 + (1-\varphi)\Delta_0 \end{pmatrix} \\ -k_1\sigma_h - k_2.$$

Rearranging, we obtain

$$\begin{aligned}
& 0 = \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi}\begin{pmatrix} -\varphi k_1 F_l E_l - \varphi F_l k_2 + \varphi k_1 (1-F_h) E_h \\ +\varphi(1-F_h)k_2 + (1-\varphi)\Delta_0 \end{pmatrix} \\
& -k_1\sigma_h - k_2 \\
\Leftrightarrow\ & 0 = \begin{pmatrix} -\varphi k_1 F_l E_l - \varphi F_l k_2 + \varphi k_1 (1-F_h) E_h \\ +\varphi(1-F_h)k_2 + (1-\varphi)\Delta_0 \end{pmatrix} \\
& -(k_1\sigma_h + k_2)(\varphi(F_l + 1 - F_h) + 1 - \varphi) \\
\Leftrightarrow\ & 0 = (-\varphi k_1 F_l E_l - \varphi F_l k_2 + \varphi k_1 (1-F_h) E_h + \varphi(1-F_h)k_2) \\
& +(k_1\sigma_l + k_2)\varphi F_l - (k_1\sigma_h + k_2)(\varphi(1-F_h)) \\
& -(1-\varphi)(k_1\sigma_h + k_2 - \Delta_0) \\
\Leftrightarrow\ & 0 = \varphi F_l(\sigma_l - E_l) + \varphi(1-F_h)(E_h - \sigma_h) - (1-\varphi)\left(\sigma_h + \frac{k_2 - \Delta_0}{k_1}\right) \\
\Leftrightarrow\ & 0 = \varphi\begin{pmatrix} F_l(\sigma_l - \mu_R) + (\gamma_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1-F_h)(\mu_R - \sigma_h) \end{pmatrix} \\
& -(1-\varphi)\left(\sigma_h + \frac{k_2 - \Delta_0}{k_1}\right),
\end{aligned} \tag{A.24}$$

where the second equivalence follows from $\Delta(\sigma_h) = \Delta(\sigma_l) \Leftrightarrow -(k_1\sigma_l + k_2) = (k_1\sigma_h + k_2)$ by (A.20), and the last equivalence follows by Lemma A.1. Thus, we have shown that the necessary and sufficient equilibrium condition given by Lemma A.3, $\Delta(\sigma_h) = \widetilde{\Delta}(\varnothing|\eta)$, holds if and only if the last equality in (A.24) holds. ■

Lemma A.5 *Let $\gamma_S \neq \gamma_R$. Then, there exists at least one equilibrium.*

Proof. We show that if $\gamma_S \neq \gamma_R$, then there always exists $\eta > 0$ such that $\tau(\eta) = 0$, which is a necessary and sufficient equilibrium condition by Lemma A.4.

Step 1. This shows that $\tau(\eta) < 0$ for sufficiently large η. Indeed, by (A.18),

$$\begin{aligned}
\tau(\eta) &= \varphi\begin{pmatrix} F_l(\widetilde{\sigma} - \eta - \mu_R) + (\gamma_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1-F_h)(\mu_R - \widetilde{\sigma} - \eta) \end{pmatrix} - (1-\varphi)\left(\widetilde{\sigma} + \eta + \frac{k_2 - \Delta_0}{k_1}\right) \\
&= \lambda_1 + \lambda_2,
\end{aligned} \tag{A.25}$$

where

$$\begin{aligned}
\lambda_1 &= \varphi\begin{pmatrix} F_l(\widetilde{\sigma} - \mu_R) + (\gamma_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1-F_h)(\mu_R - \widetilde{\sigma}) \end{pmatrix}, \\
\lambda_2 &= -\eta\left(\varphi(F_l + (1-F_h)) + 1 - \varphi\right) - (1-\varphi)\left(\widetilde{\sigma} + \frac{k_2 - \Delta_0}{k_1}\right).
\end{aligned}$$

If $\eta \to \infty$, then $\sigma_l = \widetilde{\sigma} - \eta \to -\infty$ and $\sigma_h = \widetilde{\sigma} + \eta \to \infty$ so that $f_l \to 0$, $f_h \to 0$, $F_l \to 0$,

$F_h \to 1$. Hence, $\lim_{\eta\to\infty}\lambda_1 = 0$ and $\lim_{\eta\to\infty}\lambda_2 = -\infty$. This implies by (A.25) that $\tau(\eta) < 0$ for sufficiently large η.

Step 2. This shows that there always exists $\eta > 0$ such that $\tau(\eta) = 0$.

Indeed, given that $\sigma_l = \widetilde{\sigma} - \eta$ and $\sigma_h = \widetilde{\sigma} + \eta$ by definition, we have $\lim_{\eta\to 0}\sigma_l = \lim_{\eta\to 0}\sigma_h = \widetilde{\sigma}$. Then, we obtain

$$\begin{aligned}\lim_{\eta\to 0}\tau(\eta) &= \varphi \lim_{\eta\to 0}\left(\begin{array}{c} F_l(\sigma_l-\mu_R)+(\gamma_R^2+\gamma_\varepsilon^2)(f_l+f_h) \\ +(1-F_h)(\mu_R-\sigma_h)\end{array}\right) \\ &\quad -(1-\varphi)\lim_{\eta\to 0}\left(\sigma_h+\frac{k_2-\Delta_0}{k_1}\right). \\ &= \varphi\left((\widetilde{\sigma}-\mu_R)(2F_R(\widetilde{\sigma})-1)+2(\gamma_R^2+\gamma_\varepsilon^2)f_R(\widetilde{\sigma})\right) \\ &\quad -(1-\varphi)\left(\widetilde{\sigma}+\frac{k_2-\Delta_0}{k_1}\right).\end{aligned}$$

The first term on the right-hand side is strictly positive since $F_R(\widetilde{\sigma}) \le 0.5$ if and only if $\widetilde{\sigma} \le \mu_R$. The second term $-(1-\varphi)\left(\widetilde{\sigma}+\frac{k_2-\Delta_0}{k_1}\right)$ is also positive since after substituting for $\widetilde{\sigma}$ from (6) and for k_1 and k_2 from (A.21) and (A.22) we obtain

$$\widetilde{\sigma}+\frac{k_2-\Delta_0}{k_1} = \operatorname{sgn}[\gamma_S-\gamma_R]\frac{\Delta_0(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)}{(\gamma_R^2-\gamma_S^2)\gamma_\varepsilon^2} \le 0.$$

Thus, $\lim_{\eta\to 0}\tau(\eta) > 0$. At the same time, by Step 1, $\tau(\eta) < 0$ for sufficiently large η. Consequently, by the continuity of $\tau(\eta)$ and the intermediate value theorem, there exists at least one $\eta > 0$ such that $\tau(\eta) = 0$. Then, by Lemma A.4 there should exist at least one equilibrium. ■

Lemma A.6 *Let $\gamma_S \neq \gamma_R$. Then, for any $\eta > 0$ it holds $\frac{d\tau(\eta)}{d\eta} = -(1-\varphi(F_h-F_l)) < 0$, where $\tau(\eta)$ is defined in (A.18).*

Proof. By (A.18) we have

$$\tau(\eta) = \varphi\left(\begin{array}{c} F_l(\sigma_l-\mu_R)+(\gamma_R^2+\gamma_\varepsilon^2)(f_l+f_h) \\ +(1-F_h)(\mu_R-\sigma_h)\end{array}\right) - (1-\varphi)\left(\sigma_h+\frac{k_2-\Delta_0}{k_1}\right). \tag{A.26}$$

Taking the derivative with respect to η (noting that $\sigma_l = \widetilde{\sigma}-\eta$, $\sigma_h = \widetilde{\sigma}+\eta$, $f_i = f_R(\sigma_i)$ and $F_i = F_R(\sigma_i)$ are functions of η for $i = l, h$, while $\frac{df_R(\sigma)}{d\sigma} = f_R(\sigma)\frac{\mu_R-\sigma}{\gamma_R^2+\gamma_\varepsilon^2}$ by the properties of the pdf of the normal distribution) we get

$$\begin{aligned}\frac{d\tau(\eta)}{d\eta} &= \varphi\left(\begin{array}{c} -f_l(\sigma_l-\mu_R)-F_l-f_l(\mu_R-\sigma_l) \\ +f_h(\mu_R-\sigma_h)-f_h(\mu_R-\sigma_h)-(1-F_h)\end{array}\right)-(1-\varphi) \\ &= -(1-\varphi(F_h-F_l)) < 0. \end{aligned} \tag{A.27}$$

■

Proof of Proposition 1

Throughout the proof we let $\gamma_S \neq \gamma_R$.

Step 1. The existence of at least one equilibrium characterized by a partial disclosure interval $I = [\widetilde{\sigma}-\eta, \widetilde{\sigma}+\eta]$ for some $\eta > 0$ (and non-existence of any other type of equilibrium) follows by Lemmas A.2 and A.5.

Next, Lemma A.6 implies that $\tau(\eta)$ defined in (A.18) is decreasing in η for $\eta > 0$. Hence, once a positive root of $\tau(\eta)$ exists, it must be unique. Consequently, since $\tau(\eta) = 0$ is a necessary and sufficient condition for $[\widetilde{\sigma} - \eta, \widetilde{\sigma} + \eta]$ to be an equilibrium disclosure interval by Lemma A.4, the equilibrium is unique.

Step 2. This shows that the equilibrium disclosure interval strictly contains the status quo signals $\underline{\sigma}$ and $\overline{\sigma}$. I.e., the disclosure interval $[\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ characterizing the unique equilibrium must be such that $\widetilde{\sigma} - \eta^* < \underline{\sigma}$ and $\overline{\sigma} < \widetilde{\sigma} + \eta^*$, where $\underline{\sigma}, \overline{\sigma}$ are such that $\Delta(\underline{\sigma}) = \Delta(\overline{\sigma}) = |\mu_S - \mu_R|$. This is equivalent to showing that $\widetilde{\Delta}(\varnothing|\eta^*) > \Delta(\underline{\sigma}) = \Delta(\overline{\sigma})$ so that S strictly prefers to disclose $\underline{\sigma}$ and $\overline{\sigma}$ over non-disclosure. The proof is by contradiction.

Suppose by contradiction that $\widetilde{\Delta}(\varnothing|\eta^*) \leq \Delta(\overline{\sigma}) = |\mu_S - \mu_R|$. No disclosure then implies two and only two possible scenarios that both have positive probability. Either σ is such that $\Delta(\sigma) > \widetilde{\Delta}(\varnothing|\eta^*)$ (i.e., S prefers non-disclosure over disclosing such signal). Or, S holds no signal, in which case disagreement is $|\mu_S - \mu_R| \geq \widetilde{\Delta}(\varnothing|\eta^*)$ by assumption. In other words, it must be true that $\widetilde{\Delta}(\varnothing|\eta^*)$ is a linear combination of

$$E\left[\Delta(\sigma) \,\middle|\, \Delta(\sigma) > \widetilde{\Delta}(\varnothing|\eta^*)\right],$$

and of $|\mu_S - \mu_R| \geq \widetilde{\Delta}(\varnothing|\eta^*)$. It follows immediately that $\widetilde{\Delta}(\varnothing|\eta^*) > \widetilde{\Delta}(\varnothing|\eta^*)$, which is a contradiction.

Step 3. This shows that if $\mu_S \neq \mu_R$, the center of the disclosure interval $\widetilde{\sigma}$ is closer to the prior mean of the more confident player (i.e., the player with the smaller prior variance). Assume $\mu_R > \mu_S$ (the proof for $\mu_R < \mu_S$ proceeds analogously and is omitted). Using (6), we obtain

$$\mu_S - \widetilde{\sigma} = \mu_S - \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2) - \mu_R(\gamma_S^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \frac{(\mu_R - \mu_S)(\gamma_S^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2}, \tag{A.28}$$

$$\mu_R - \widetilde{\sigma} = \mu_R - \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2) - \mu_R(\gamma_S^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \frac{(\mu_R - \mu_S)(\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2}. \tag{A.29}$$

Hence, if $\gamma_R > \gamma_S$ we get $\mu_S - \widetilde{\sigma} > 0 \Rightarrow \widetilde{\sigma} < \mu_S < \mu_R$ (where the last inequality is by assumption), while if $\gamma_R < \gamma_S$ we get $\mu_R - \widetilde{\sigma} < 0 \Rightarrow \widetilde{\sigma} > \mu_R > \mu_S$. Thus, $\widetilde{\sigma} \notin [\mu_S, \mu_R]$ while $\widetilde{\sigma}$ is closer to the prior mean of the more confident player. ■

Corollary A.1 *Prior means always lie within the interior of the disclosure interval, i.e.,* $\mu_R \in (\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*)$ *and* $\mu_S \in (\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*)$.

Proof. If $d = \sigma = \mu_R$, then by (A.20) the perceived disagreement is

$$\begin{aligned}\Delta(\mu_R) &= \left|\left(\frac{\gamma_S^2}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2 + \gamma_\varepsilon^2}\right)\mu_R + \left(\frac{\gamma_\varepsilon^2 \mu_S}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2 \mu_R}{\gamma_R^2 + \gamma_\varepsilon^2}\right)\right| \\ &= \left|\frac{\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2}(\mu_S - \mu_R)\right| < |\mu_S - \mu_R|.\end{aligned}$$

Thus,

$$\Delta(\mu_R) < \Delta_0 = \Delta(\underline{\sigma}) = \Delta(\overline{\sigma}),$$

where the equalities are by definition of the status quo signals. Since $\Delta(\sigma)$ is V-shaped around $\widetilde{\sigma}$, this implies that $\mu_R \in (\underline{\sigma}, \overline{\sigma})$. Consequently, by Proposition 1(b), $\mu_R \in (\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*)$.

Similarly,

$$\begin{aligned}\Delta(\mu_S) &= \left|\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\mu_S + \left(\frac{\gamma_\varepsilon^2\mu_S}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right)\right| \\ &= \left|\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}(\mu_S-\mu_R)\right| < |\mu_S-\mu_R|.\end{aligned}$$

The claim that $\mu_S \in (\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*)$ then follows by the analogous argument as for μ_R. ■

A.2.2 Proof of Proposition 2

Let $\gamma_S = \gamma_R$. Then, by (A.20) for any $\sigma \in \mathbb{R}$,

$$\Delta(\sigma) = \left|\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}(\mu_S-\mu_R)\right| \le |\mu_S-\mu_R|. \tag{A.30}$$

Consequently, there exists an equilibrium with full disclosure (as in such equilibrium non-disclosure reveals that S is indeed uninformed so that $\widetilde{\Delta}(\varnothing) = |\mu_S - \mu_R| \ge \Delta(\sigma)$ for any σ).

Let us show that no other equilibrium exists. Assume by contradiction that there exists an equilibrium where some signal σ' is not disclosed. We have

$$\begin{aligned}\widetilde{\Delta}(\varnothing) &= \begin{array}{c}\Pr(\sigma=\varnothing\,|d=\varnothing)\,|\mu_S-\mu_R| \\ +\Pr(\sigma\neq\varnothing\,|d=\varnothing)E_R[\Delta(\sigma)|d=\varnothing,\sigma\neq\varnothing]\end{array} \\ &= \begin{array}{c}\Pr(\sigma=\varnothing\,|d=\varnothing)\,|\mu_S-\mu_R| \\ +\Pr(\sigma\neq\varnothing\,|d=\varnothing)\left|\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}(\mu_S-\mu_R)\right|\end{array} \\ &\ge \left|\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}(\mu_S-\mu_R)\right| = \Delta(\sigma'),\end{aligned} \tag{A.31}$$

where the second and the last equalities are by (A.30). Hence, S has incentive to deviate to disclosure after observing σ', which is a contradiction. ■

A.3 Receiver preferences over senders

A.3.1 Preliminaries

Lemma A.7 *a) Let I' and I'' denote two equilibrium disclosure strategies such that S discloses iff $\sigma \in [\sigma'_l, \sigma'_h]$ and iff $\sigma \in [\sigma''_l, \sigma''_h]$ under I' and I'', respectively. Assume $[\sigma''_l, \sigma''_h]$ is a strict superset of $[\sigma'_l, \sigma'_h]$ such that $\sigma''_l < \sigma'_l < \sigma'_h < \sigma''_h$. Then, I'' is more Blackwell informative than I'.*

b) Full disclosure is more Blackwell informative than a partial disclosure.

Proof. a) Denote by $d_{I'}$ and $d_{I''}$ the value of disclosure d under I' and I'', respectively. Denote also by $g_{I'}(d_{I'})$ and $g_{I''}(d_{I''})$ the corresponding ex ante density functions for $d \neq \varnothing$.

Note that for any $\widehat{\sigma}$ within a disclosure interval $[\sigma_l, \sigma_h]$, it holds in equilibrium

$$\Pr[\sigma \leq \widehat{\sigma}] = \Pr[\sigma \in [\sigma_l, \widehat{\sigma}]] = \varphi \int_{\sigma_l}^{\widehat{\sigma}} f_R(\sigma) d\sigma$$

so that for $I \in \{I', I''\}$ it holds

$$g_I(\widehat{\sigma}) = \frac{\partial \Pr[\sigma \leq \widehat{\sigma}]}{\partial \widehat{\sigma}} = \varphi f_R(\widehat{\sigma}).$$

Clearly, the same applies to the conditional density functions, i.e., $g_I(\widehat{\sigma}|\omega) = \varphi f_R(\widehat{\sigma}|\omega)$. Then, from R's perspective, I'' is sufficient for I' in the sense of Blackwell (1953) since for all possible values of $d_{I'}$, the probability/density of $d_{I'}$ conditional on any $\omega \in \Omega$ can be derived from these probabilities/densities of $d_{I''}$ as follows (i.e., there exists a Markov kernel from the set of signals under disclosure strategy I'' to the set of signals under disclosure strategy I'):

- For $d_{I'} = \varnothing$:

$$\begin{aligned}
\Pr[d_{I'} = \varnothing|\omega] &= \Pr[d_{I''} = \varnothing|\omega] + \varphi \left(\int_{\sigma_l''}^{\sigma_l'} f_R(\sigma|\omega) d\sigma + \int_{\sigma_h'}^{\sigma_h''} f_R(\sigma|\omega) d\sigma \right) \\
&= \Pr[d_{I''} = \varnothing|\omega] + \int_{\sigma_l''}^{\sigma_l'} g_{I''}(\sigma|\omega) d\sigma + \int_{\sigma_h'}^{\sigma_h''} g_{I''}(\sigma|\omega) d\sigma,
\end{aligned}$$

- For $d_{I'} = \sigma \in [\sigma_l', \sigma_h']$:

$$g_{I'}(\sigma|\omega) = \varphi f_R(\sigma|\omega) = g_{I''}(\sigma|\omega).$$

Then, by Blackwell theorem, I'' is more informative than I' (note that the Blackwell theorem is generalized for infinite state spaces, as in our model, by Amershi (1988)).

b) The claim for full disclosure follows by the same arguments as in point a) while setting $\sigma_l'' = -\infty$ and $\sigma_h'' = \infty$. ■

Lemma A.8 *a) Let $I'$and I'' denote two equilibrium disclosure strategies such that S discloses iff $\sigma \in [\sigma_l', \sigma_h']$ and iff $\sigma \in [\sigma_l'', \sigma_h'']$ under I' and I'', respectively. Assume $[\sigma_l'', \sigma_h'']$ is a strict superset of $[\sigma_l', \sigma_h']$ such that $\sigma_l'' < \sigma_l' < \sigma_h' < \sigma_h''$. Then, R's ex ante expected utility is strictly higher under I'' than under I'.*

b) R's ex ante expected utility is strictly higher under full disclosure than under partial disclosure.

Proof.

a) Consider two disclosure intervals $[\sigma_l', \sigma_h']$ and $[\sigma_l'', \sigma_h'']$ such that $\sigma_l'' < \sigma_l' < \sigma_h' < \sigma_h''$. Denote the optimal R's action given disclosure d as

$$a^*(d) = \arg\max_a E_R[u_R(\omega, a)|d] = \arg\max_a E_R[-(\omega - a)^2|d].$$

In particular, denote the optimal action conditional on non-disclosure given disclosure interval $[\sigma_l, \sigma_h]$ as $a^*(\varnothing, \sigma_l, \sigma_h)$. Recall also that S's signal space is given by $(-\infty, \infty) \cup \varnothing$, i.e.,

$\sigma \notin [\sigma_l, \sigma_h] \Leftrightarrow \sigma \in (-\infty, \sigma_l) \cup (\sigma_h, \infty) \cup \varnothing$. Then, R's ex ante expected utility conditional on disclosure interval $[\sigma_l'', \sigma_h'']$ is

$$\begin{aligned}
E_R[u_R|\sigma_l'', \sigma_h''] &= \Pr[\sigma \notin [\sigma_l'', \sigma_h'']] E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l'', \sigma_h''))|\sigma \notin [\sigma_l'', \sigma_h'']\right] \\
&\quad + \varphi \int_{\sigma_l''}^{\sigma_h''} E_R\left[u_R(\omega, a^*(\sigma))|\sigma\right] f_R(\sigma) d\sigma \\
&= \Pr[\sigma \notin [\sigma_l'', \sigma_h'']] E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l'', \sigma_h''))|\sigma \notin [\sigma_l'', \sigma_h'']\right] \\
&\quad + \varphi \int_{\sigma_l''}^{\sigma_l'} E_R\left[u_R(\omega, a^*(\sigma)|\sigma\right] f_R(\sigma) d\sigma \\
&\quad + \varphi \int_{\sigma_h'}^{\sigma_h''} E_R\left[u_R(\omega, a^*(\sigma)|\sigma\right] f_R(\sigma) d\sigma \\
&\quad + \varphi \int_{\sigma_l'}^{\sigma_h'} E_R\left[u_R(\omega, a^*(\sigma))|\sigma\right] f_R(\sigma) d\sigma.
\end{aligned}$$

Analogously, for the disclosure interval $[\sigma_l', \sigma_h']$ we obtain

$$\begin{aligned}
E_R[u_R|\sigma_l', \sigma_h'] &= \Pr[\sigma \notin [\sigma_l', \sigma_h']] E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h'))|\sigma \notin [\sigma_l', \sigma_h']\right] \\
&\quad + \varphi \int_{\sigma_l'}^{\sigma_h'} E_R\left[u_R(\omega, a^*(\sigma))|\sigma\right] f_R(\sigma) d\sigma \\
&= \Pr[\sigma \notin [\sigma_l'', \sigma_h'']] E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h'))|\sigma \notin [\sigma_l'', \sigma_h'']\right] \\
&\quad + \varphi \int_{\sigma_l''}^{\sigma_l'} E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h')|\sigma\right] f_R(\sigma) d\sigma \\
&\quad + \varphi \int_{\sigma_h'}^{\sigma_h''} E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h')|\sigma\right] f_R(\sigma) d\sigma \\
&\quad + \varphi \int_{\sigma_l'}^{\sigma_h'} E_R\left[u_R(\omega, a^*(\sigma))|\sigma\right] f_R(\sigma) d\sigma,
\end{aligned}$$

where the second equality follows by the law of total expectations. Then,

$$E_R[u_R|\sigma_l'', \sigma_h''] - E_R[u_R|\sigma_l', \sigma_h'] = \lambda_1 + \lambda_2 + \lambda_3, \tag{A.32}$$

where

$$\begin{aligned}
\lambda_1 &= \Pr[\sigma \notin [\sigma_l'', \sigma_h'']] \\
&\quad \times \begin{pmatrix} E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l'', \sigma_h''))|\sigma \notin [\sigma_l'', \sigma_h'']\right] \\ -E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h'))|\sigma \notin [\sigma_l'', \sigma_h'']\right] \end{pmatrix}, \\
\lambda_2 &= \varphi \int_{\sigma_l''}^{\sigma_l'} \begin{pmatrix} E_R\left[u_R(\omega, a^*(\sigma)|\sigma\right] \\ -E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h')|\sigma\right] \end{pmatrix} f_R(\sigma) d\sigma, \\
\lambda_3 &= \varphi \int_{\sigma_h'}^{\sigma_h''} \begin{pmatrix} E_R\left[u_R(\omega, a^*(\sigma)|\sigma\right] \\ -E_R\left[u_R(\omega, a^*(\varnothing, \sigma_l', \sigma_h')|\sigma\right] \end{pmatrix} f_R(\sigma) d\sigma.
\end{aligned}$$

We have that $\lambda_1 \geq 0$ since $a^*(\varnothing, \sigma_l'', \sigma_h'') = E[\omega|\sigma \notin [\sigma_l'', \sigma_h'']]$ is the unique maximum of $E_R[u_R(\omega, a)|\sigma \notin [\sigma_l'', \sigma_h'']]$ on $a \in \mathbb{R}$. Analogously, $\lambda_2 > 0$ and $\lambda_3 > 0$ since $a^*(\sigma)$ is the unique maximum of $E_R[u_R(\omega, a)|\sigma]$ on $a \in \mathbb{R}$. The inequalities are strict in the latter cases since, generally, $a^*(\sigma) \neq a^*(\varnothing, \sigma_l', \sigma_h')$ on a positive measure of signals. Given that $\lambda_1 \geq 0$, $\lambda_2 > 0$ and $\lambda_3 > 0$, (A.32) implies that $E_R[u_R|\sigma_l'', \sigma_h''] > E_R[u_R|\sigma_l', \sigma_h']$, i.e., R's ex ante expected utility is strictly higher under $[\sigma_l'', \sigma_h'']$.

b) The fact that R always prefers full disclosure over partial disclosure follows by the same arguments as in point a) while setting $\sigma_l'' = -\infty$ and $\sigma_h'' = \infty$. ■

Lemma A.9 *Let $\gamma_S \neq \gamma_R$. Consider S's priors μ_S', μ_S'' such that μ_S'' is symmetric around μ_R relative to μ_S', i.e., $\mu_S'' = 2\mu_R - \mu_S'$. Denote the equilibrium disclosure intervals under μ_S' and μ_S'' as $[\sigma_l', \sigma_h']$ and $[\sigma_l'', \sigma_h'']$, respectively. Then, $[\sigma_l'', \sigma_h'']$ is symmetric around μ_R relative to $[\sigma_l', \sigma_h']$, i.e., $\sigma_l'' = 2\mu_R - \sigma_h'$ and $\sigma_h'' = 2\mu_R - \sigma_l'$.*

Proof. First, we prove the following claims.

Claim 1: $\Delta(\sigma''|\mu_S'') = \Delta(2\mu_R - \sigma''|\mu_S')$ for any $\sigma'' \in \mathbb{R}$.

Proof. By (A.16), for any signal $\sigma'' \in \mathbb{R}$ it holds

$$\begin{aligned}
\Delta(\sigma''|\mu_S'') &= \left|\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\sigma'' + \left(\frac{\gamma_\varepsilon^2\mu_S''}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right)\right| \\
&= \left|\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\sigma'' + \left(\frac{\gamma_\varepsilon^2(2\mu_R-\mu_S')}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right)\right| \\
&= \left|\begin{array}{c} \left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\sigma'' \\ + \left(\frac{\gamma_\varepsilon^2(2\mu_R-\mu_S')(\gamma_R^2+\gamma_\varepsilon^2) - \gamma_\varepsilon^2\mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)}\right)\end{array}\right|
\end{aligned} \tag{A.33}$$

At the same time,

$$\begin{aligned}
\Delta(2\mu_R - \sigma''|\mu_S') &= \left|\begin{array}{c} \left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)(2\mu_R - \sigma'') \\ + \left(\frac{\gamma_\varepsilon^2\mu_S'}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2\mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right)\end{array}\right| \\
&= \left|\begin{array}{c} -\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\sigma'' \\ + \left(\frac{(\gamma_\varepsilon^2\mu_S' + 2\mu_R\gamma_S^2)(\gamma_R^2+\gamma_\varepsilon^2) - (\gamma_\varepsilon^2\mu_R + 2\mu_R\gamma_R^2)(\gamma_S^2+\gamma_\varepsilon^2)}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)}\right)\end{array}\right| \\
&= \left|\begin{array}{c} -\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)\sigma'' \\ - \left(\frac{\gamma_\varepsilon^2(2\mu_R-\mu_S')(\gamma_R^2+\gamma_\varepsilon^2) - \gamma_\varepsilon^2\mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)}\right)\end{array}\right| \\
&= \Delta(\sigma''|\mu_S''),
\end{aligned}$$

where the last equality is by (A.33).

Claim 2: $\widetilde{\sigma}(\mu_S'') = 2\mu_R - \widetilde{\sigma}(\mu_S')$.

Proof. We have

$$\begin{aligned}
\widetilde{\sigma}(\mu''_S) &= \frac{\mu''_S(\gamma_R^2+\gamma_\varepsilon^2)-\mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{\gamma_R^2-\gamma_S^2} \\
&= \frac{(2\mu_R-\mu'_S)(\gamma_R^2+\gamma_\varepsilon^2)-\mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{\gamma_R^2-\gamma_S^2} \\
&= 2\mu_R-\frac{\mu'_S(\gamma_R^2+\gamma_\varepsilon^2)-\mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{\gamma_R^2-\gamma_S^2} \\
&= 2\mu_R-\widetilde{\sigma}(\mu'_S)
\end{aligned}$$

Claim 3: For any given η, $F_l(\eta,\mu''_S)=1-F_h(\eta,\mu'_S)$ and $F_h(\eta,\mu''_S)=1-F_l(\eta,\mu'_S)$.

Proof. By symmetry of $f_R(\sigma)$ around its mean μ_R, we have

$$f_R(\sigma)=f_R(2\mu_R-\sigma) \tag{A.34}$$

for any σ. Then,

$$\begin{aligned}
F_l(\eta,\mu''_S) &= \int_{-\infty}^{\widetilde{\sigma}(\mu''_S)-\eta} f_R(\sigma)d\sigma \\
&= \int_{-\infty}^{2\mu_R-\widetilde{\sigma}(\mu'_S)-\eta} f_R(2\mu_R-\sigma)d\sigma \\
&= \int_{\widetilde{\sigma}(\mu'_S)+\eta}^{\infty} f_R(\sigma)d\sigma \\
&= 1-F_h(\eta,\mu'_S),
\end{aligned}$$

where the second equality is by Claim 2 and (A.34), and the third equality is obtained by integration by substitution.

Then, $F_l(\eta,\mu'_S)=1-F_h(\eta,\mu''_S)\Leftrightarrow F_h(\eta,\mu''_S)=1-F_l(\eta,\mu'_S)$ follows immediately by symmetry considerations.

Claim 4. $\widetilde{\Delta}(\varnothing|\eta,\mu''_S)=\widetilde{\Delta}(\varnothing|\eta,\mu'_S)$.

Proof. By (A.19)

$$
\begin{aligned}
& \widetilde{\Delta}(\varnothing|\eta, \mu_S'') \\
= & \frac{1}{\varphi(F_l(\mu_S'') + 1 - F_h(\mu_S'')) + 1 - \varphi} \begin{pmatrix} \varphi \int_{-\infty}^{\widetilde{\sigma}(\mu_S'')-\eta} \Delta(\sigma|\mu_S'') f_R(\sigma) d\sigma \\ +\varphi \int_{\widetilde{\sigma}(\mu_S'')+\eta}^{\infty} \Delta(\sigma|\mu_S'') f_R(\sigma) d\sigma \\ +(1-\varphi)\Delta_0(\mu_S'') \end{pmatrix} \\
= & \frac{1}{\varphi(F_l(\mu_S') + 1 - F_h(\mu_S')) + 1 - \varphi} \\
& \times \begin{pmatrix} \varphi \int_{-\infty}^{2\mu_R - \widetilde{\sigma}(\mu_R') - \eta} \Delta(2\mu_R - \sigma|\mu_S') f_R(2\mu_R - \sigma) d\sigma \\ +\varphi \int_{2\mu_R - \widetilde{\sigma}(\mu_S')+\eta}^{\infty} \Delta(2\mu_R - \sigma|\mu_S') f_R(2\mu_R - \sigma) d\sigma \\ +(1-\varphi)\Delta_0(\mu_S') \end{pmatrix} \\
= & \frac{1}{\varphi(F_l(\mu_S') + 1 - F_h(\mu_S')) + 1 - \varphi} \begin{pmatrix} \varphi \int_{\widetilde{\sigma}(\mu_R')+\eta}^{\infty} \Delta(\sigma|\mu_S') f_R(\sigma) d\sigma \\ +\varphi \int_{-\infty}^{\widetilde{\sigma}(\mu_S')-\eta} \Delta(\sigma|\mu_S') f_R(\sigma) d\sigma \\ +(1-\varphi)\Delta_0(\mu_S') \end{pmatrix} \\
= & \widetilde{\Delta}(\varnothing|\eta, \mu_S'),
\end{aligned}
$$

where the second equality follows by Claims 1, 2, 3 and (A.34) (it is also obvious that $\Delta_0(\mu_S'') = \Delta_0(\mu_S')$), and the third equality is derived by integration by substitution.

Claim 5. The equilibrium value of η is the same under both means: $\eta^*(\mu_S'') = \eta^*(\mu_S')$.

Proof. We have

$$
\begin{aligned}
& \widetilde{\Delta}(\varnothing|\eta, \mu_S'') - \Delta(\widetilde{\sigma}(\mu_S'') + \eta|\mu_S'') \\
= & \widetilde{\Delta}(\varnothing|\eta, \mu_S') - \Delta(2\mu_R - \widetilde{\sigma}(\mu_S') + \eta|\mu_S'') \\
= & \widetilde{\Delta}(\varnothing|\eta, \mu_S') - \Delta(\widetilde{\sigma}(\mu_S') - \eta|\mu_S') \\
= & \widetilde{\Delta}(\varnothing|\eta, \mu_S') - \Delta(\widetilde{\sigma}(\mu_S') + \eta|\mu_S'),
\end{aligned}
$$

where the first equality is by Claims 2 and 4, the second equality is by Claim 1, and the third equality is by the symmetry of $\Delta(\sigma)$ around $\widetilde{\sigma}$. Consequently, the equilibrium value of η (solving $\widetilde{\Delta}(\varnothing|\eta) - \Delta(\widetilde{\sigma} + \eta) = 0$ by Lemma A.3, where the solution must be unique by Proposition 1(a)) is the same under μ_S'' and μ_S':

$$
\eta^*(\mu_S'') = \eta^*(\mu_S').
$$

Finally, let us show the claim of the lemma. We have

$$
\begin{aligned}
\sigma_l(\mu_S'') &= \widetilde{\sigma}(\mu_S'') - \eta^*(\mu_S'') \\
&= 2\mu_R - \widetilde{\sigma}(\mu_S') - \eta^*(\mu_S') \\
&= 2\mu_R - \sigma_h(\mu_S'),
\end{aligned}
$$

where the second equality is by Claims 2 and 5. Analogously

$$\begin{aligned}\sigma_h(\mu_S'') &= \widetilde{\sigma}(\mu_S'') + \eta^*(\mu_S'') \\ &= 2\mu_R - \widetilde{\sigma}(\mu_S') + \eta^*(\mu_S') \\ &= 2\mu_R - \sigma_l(\mu_S').\end{aligned}$$

■

Lemma A.10 *Let $\gamma_S \neq \gamma_R$. Consider R's priors μ_R', μ_R'' such that μ_R'' is symmetric around μ_S relative to μ_R', i.e., $\mu_R'' = 2\mu_S - \mu_R'$. Denote the equilibrium disclosure intervals under μ_R' and μ_R'' as $[\sigma_l', \sigma_h']$ and $[\sigma_l'', \sigma_h'']$, respectively. Then, $[\sigma_l'', \sigma_h'']$ is symmetric around μ_S relative to $[\sigma_l', \sigma_h']$, i.e., $\sigma_l'' = 2\mu_S - \sigma_h'$ and $\sigma_h'' = 2\mu_S - \sigma_l'$.*

Proof. The proof proceeds analogously to the proof of Lemma A.9 and is omitted. ■

Lemma A.11 *a) For any parameter $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$, if $\gamma_S \neq \gamma_R$ then*

$$\frac{d\sigma_l(\eta^*)}{d\kappa} = -\frac{1}{1-\varphi(F_h - F_l)} \frac{\partial \widetilde{\tau}(x,\kappa)}{\partial \kappa}_{|x=\sigma_l(\eta^*)},$$

where

$$\widetilde{\tau}(x,\kappa) = \varphi \begin{pmatrix} F_R[x](x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[x]+f_R[2\widetilde{\sigma}-x]) \\ +(1-F_R[2\widetilde{\sigma}-x])(\mu_R - 2\widetilde{\sigma}+x) \end{pmatrix} - (1-\varphi)\left(2\widetilde{\sigma} - x + \frac{k_2-\Delta_0}{k_1}\right).$$

b) For any parameter $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$, if $\gamma_S \neq \gamma_R$ then

$$\frac{d\sigma_h(\eta^*)}{d\kappa} = \frac{1}{1-\varphi(F_h - F_l)} \frac{\partial \widehat{\tau}(x,\kappa)}{\partial \kappa}_{|x=\sigma_h(\eta^*)},$$

where

$$\widehat{\tau}(x,\kappa) = \varphi \begin{pmatrix} F_R[2\widetilde{\sigma}-x](2\widetilde{\sigma}-x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-x]+f_R[x]) \\ +(1-F_R[x])(\mu_R - x) \end{pmatrix} - (1-\varphi)\left(x + \frac{k_2-\Delta_0}{k_1}\right).$$

Proof. a) Consider $\widetilde{\tau}(x,\kappa)$ defined in the lemma for $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ and assume $\gamma_S \neq \gamma_R$. Note that at $x = \sigma_l(\eta^*)$ we obtain

$$\begin{aligned}\widetilde{\tau}(\sigma_l(\eta^*),\kappa) &= \varphi \begin{pmatrix} F_R[\sigma_l(\eta^*)](\sigma_l(\eta^*)-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[\sigma_l(\eta^*)]+f_R[2\widetilde{\sigma}-\sigma_l(\eta^*)]) \\ +(1-F_R[2\widetilde{\sigma}-\sigma_l(\eta^*)])(\mu_R - 2\widetilde{\sigma}+\sigma_l(\eta^*)) \end{pmatrix} \\ &\quad -(1-\varphi)\left(2\widetilde{\sigma} - \sigma_l(\eta^*) + \frac{k_2-\Delta_0}{k_1}\right) \\ &= \varphi \begin{pmatrix} F_R[\sigma_l(\eta^*)](\sigma_l(\eta^*)-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[\sigma_l(\eta^*)]+f_R[\sigma_h(\eta^*)]) \\ +(1-F_R[\sigma_h(\eta^*)])(\mu_R - \sigma_h(\eta^*)) \end{pmatrix} \\ &\quad -(1-\varphi)\left(\sigma_h(\eta^*) + \frac{k_2-\Delta_0}{k_1}\right) \\ &= \tau(\eta^*), \qquad \text{(A.35)}\end{aligned}$$

where the second equality simply follows from $2\widetilde{\sigma} - \sigma_l(\eta) = 2\widetilde{\sigma} - (\widetilde{\sigma} - \eta) = \sigma_h(\eta)$. Since $\tau(\eta^*) = 0$ if $\gamma_S \neq \gamma_R$ by Lemma A.4, (A.35) implies

$$\begin{aligned}
& \widetilde{\tau}(\sigma_l(\eta^*), \kappa) \\
= \; & \varphi \begin{pmatrix} F_R[\sigma_l(\eta^*)](\sigma_l(\eta^*) - \mu_R) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)(f_R[\sigma_l(\eta^*)] + f_R[2\widetilde{\sigma} - \sigma_l(\eta^*)]) \\ +(1 - F_R[2\widetilde{\sigma} - \sigma_l(\eta^*)])(\mu_R - 2\widetilde{\sigma} + \sigma_l(\eta^*)) \end{pmatrix} \\
& -(1-\varphi)\left(2\widetilde{\sigma} - \sigma_l(\eta^*) + \frac{k_2 - \Delta_0}{k_1}\right) \\
= \; & 0.
\end{aligned}$$

Consequently, by the implicit function theorem, for any parameter $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ at the equilibrium value of η^* it holds:

$$\frac{d\sigma_l(\eta^*)}{d\kappa} = -\frac{\partial \widetilde{\tau}(x,\kappa)/\partial\kappa}{\partial \widetilde{\tau}(x,\kappa)/\partial x}_{|x=\sigma_l(\eta^*)}, \tag{A.36}$$

where, in particular, $\frac{\partial\widetilde{\tau}(x,\kappa)}{d\kappa}_{|x=\sigma_l(\eta^*)}$ is the partial derivative of $\widetilde{\tau}(x,\kappa)$ with respect to κ keeping x constant at $\sigma_l(\eta^*)$ (which always exists for any $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ if $\gamma_S \neq \gamma_R$).

Next, consider the denominator in (A.36). Note that for any $x \in \mathbb{R}$,

$$\begin{aligned}
\widetilde{\tau}(x,\kappa) \; = \; & \varphi \begin{pmatrix} F_R[x](x - \mu_R) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)(f_R[x] + f_R[2\widetilde{\sigma} - x]) \\ +(1 - F_R[2\widetilde{\sigma} - x])(\mu_R - 2\widetilde{\sigma} + x) \end{pmatrix} \\
& -(1-\varphi)\left(2\widetilde{\sigma} - x + \frac{k_2 - \Delta_0}{k_1}\right) \\
= \; & \varphi \begin{pmatrix} F_R[\widetilde{\sigma} - (\widetilde{\sigma} - x)](\widetilde{\sigma} - (\widetilde{\sigma} - x) - \mu_R) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)(f_R[\widetilde{\sigma} - (\widetilde{\sigma} - x)] + f_R[\widetilde{\sigma} + (\widetilde{\sigma} - x)]) \\ +(1 - F_R[\widetilde{\sigma} + (\widetilde{\sigma} - x)])(\mu_R - (\widetilde{\sigma} + (\widetilde{\sigma} - x))) \end{pmatrix} \\
& -(1-\varphi)\left(\widetilde{\sigma} + (\widetilde{\sigma} - x) + \frac{k_2 - \Delta_0}{k_1}\right) \\
= \; & \tau(\widetilde{\sigma} - x).
\end{aligned} \tag{A.37}$$

It follows that $\frac{\partial\widetilde{\tau}(x,\kappa)}{\partial x} = \frac{d\tau(\widetilde{\sigma}-x)}{dx} = -\frac{d\tau(\eta)}{d\eta}_{|\eta=\widetilde{\sigma}-x}$. This implies that at $x = \sigma_l(\eta^*) = \widetilde{\sigma} - \eta^*$ we have

$$\frac{\partial\widetilde{\tau}(x,\kappa)}{\partial x}_{|x=\sigma_l(\eta^*)} = -\frac{d\tau(\eta^*)}{d\eta^*} = 1 - \varphi(F_h - F_l),$$

where the last equality is by Lemma A.6. Substituting this into (A.36) we finally obtain

$$\frac{d\sigma_l(\eta^*)}{d\kappa} = -\frac{1}{1 - \varphi(F_h - F_l)} \frac{\partial\widetilde{\tau}(x,\kappa)}{\partial\kappa}_{|x=\sigma_l(\eta^*)}.$$

b) Consider $\widehat{\tau}(x,\kappa)$ defined in the lemma for $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ and assume $\gamma_S \neq \gamma_R$. As in (A.35), it is easy to verify that $\widehat{\tau}(\sigma_h(\eta^*), \kappa) = \tau(\eta^*) = 0$, where the last equality is by

Lemma A.4. This implies

$$\begin{aligned}
\widehat{\tau}(\sigma_h(\eta^*),\kappa) &= \varphi\left(\begin{array}{c} F_R[2\widetilde{\sigma}-\sigma_h(\eta^*)](2\widetilde{\sigma}-\sigma_h(\eta^*)-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-\sigma_h(\eta^*)]+f_R[\sigma_h(\eta^*)]) \\ +(1-F_R[\sigma_h(\eta^*)])(\mu_R-\sigma_h(\eta^*)) \end{array}\right) \\
&\quad -(1-\varphi)\left(\sigma_h(\eta^*)+\frac{k_2-\Delta_0}{k_1}\right) \\
&= 0.
\end{aligned}$$

Consequently, by the implicit function theorem, for any parameter $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ at the equilibrium value of η^* it holds:

$$\frac{d\sigma_h(\eta^*)}{d\kappa} = -\frac{\partial\widehat{\tau}(x,\kappa)/\partial\kappa}{\partial\widehat{\tau}(x,\kappa)/\partial x}_{|x=\sigma_h(\eta^*)}, \tag{A.38}$$

where, in particular, $\frac{\partial\widehat{\tau}(x,\kappa)}{\partial\kappa}_{|x=\sigma_h(\eta^*)}$ is the partial derivative of $\widehat{\tau}(x,\kappa)$ with respect to κ keeping x constant at $\sigma_h(\eta^*)$ (which always exists for any $\kappa \in \{\mu_S, \gamma_S, \mu_R, \gamma_R\}$ if $\gamma_S \neq \gamma_R$).

Next, consider the denominator in (A.38). Note that for any $x \in \mathbb{R}$,

$$\begin{aligned}
\widehat{\tau}(x,\kappa) &= \varphi\left(\begin{array}{c} F_R[2\widetilde{\sigma}-x](2\widetilde{\sigma}-x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-x]+f_R[x]) \\ +(1-F_R[x])(\mu_R-x) \end{array}\right) \\
&\quad -(1-\varphi)\left(x+\frac{k_2-\Delta_0}{k_1}\right) \\
&= \varphi\left(\begin{array}{c} F_R[\widetilde{\sigma}-(x-\widetilde{\sigma})](\widetilde{\sigma}-(x-\widetilde{\sigma})-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[\widetilde{\sigma}-(x-\widetilde{\sigma})]+f_R[\widetilde{\sigma}+(x-\widetilde{\sigma})]) \\ +(1-F_R[\widetilde{\sigma}+(x-\widetilde{\sigma})])(\mu_R-(\widetilde{\sigma}+(x-\widetilde{\sigma}))) \end{array}\right) \\
&\quad -(1-\varphi)\left(\widetilde{\sigma}+(x-\widetilde{\sigma})+\frac{k_2-\Delta_0}{k_1}\right) \\
&= \tau(x-\widetilde{\sigma}).
\end{aligned} \tag{A.39}$$

It follows that $\frac{\partial\widehat{\tau}(x,\kappa)}{\partial x} = \frac{d\tau(x-\widetilde{\sigma})}{dx} = \frac{d\tau(\eta)}{d\eta}_{|\eta=x-\widetilde{\sigma}}$. This implies that at $x = \sigma_h(\eta^*) = \widetilde{\sigma}+\eta^*$ we have

$$\frac{\partial\widehat{\tau}(x,\kappa)}{\partial x}_{|x=\sigma_h(\eta^*)} = \frac{d\tau(\eta^*)}{d\eta^*} = -\left(1-\varphi(F_h-F_l)\right),$$

where the last equality is by Lemma A.6. Substituting this into (A.38) we finally obtain

$$\frac{d\sigma_h(\eta^*)}{d\kappa} = \frac{1}{1-\varphi(F_h-F_l)}\frac{\partial\widehat{\tau}(x,\kappa)}{\partial\kappa}_{|x=\sigma_h(\eta^*)}.$$

■

Lemma A.12 *Let $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Fix the receiver's prior (μ_R, γ_R) as well as the sender's prior confidence $1/\gamma_S^2$, with $\gamma_R \neq \gamma_S$. Then, keeping the order of μ_R and μ_S fixed, the lower (upper) bound of the equilibrium disclosure interval strictly decreases (increases) when prior disagreement $|\mu_R - \mu_S|$ strictly increases.*

Proof.

Step 1. This shows the claim of the lemma for the case $\mu_S \geq \mu_R$. Herewith, we separately consider two possible parameter cases: $\gamma_S > \gamma_R$ and $\gamma_S < \gamma_R$.

Case 1: $\gamma_S > \gamma_R$.

Claim 1: *A sufficient condition for the claim of Lemma A.12 under* $\mu_S \geq \mu_R$ *and* $\gamma_S > \gamma_R$ *is*

$$\frac{d\sigma_h(\eta^*)}{d\mu_S} > 0 \text{ at any } \mu_S \geq \mu_R.$$

Proof. Assume $\frac{d\sigma_h(\eta^*)}{d\mu_S} > 0$ at any $\mu_S \geq \mu_R$. Then, given that $\sigma_l(\eta^*) = \widetilde{\sigma} - \eta^* = 2\widetilde{\sigma} - \sigma_h(\eta^*)$, we also have $\frac{d\sigma_l(\eta^*)}{d\mu_S} = 2\frac{d\widetilde{\sigma}}{d\mu_S} - \frac{d\sigma_h(\eta^*)}{d\mu_S} < 0$ at any $\mu_S \geq \mu_R$ (since $\frac{d\widetilde{\sigma}}{d\mu_S} = \frac{\gamma_R^2+\gamma_\varepsilon^2}{\gamma_R^2-\gamma_S^2} < 0$ under $\gamma_S > \gamma_R$). Thus, the lower (upper) bound of the equilibrium disclosure interval strictly decreases (increases) when prior disagreement $|\mu_R - \mu_S|$ strictly increases.

Claim 2: *If* $\gamma_S > \gamma_R$ *and* $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, *then* $\frac{d\sigma_h(\eta^*)}{d\mu_S} > 0$ *at any* $\mu_S \geq \mu_R$.

Proof. By Lemma A.11,

$$\frac{d\sigma_h(\eta^*)}{d\mu_S} = \frac{1}{1-\varphi(F_h - F_l)} \frac{\partial\widehat{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_h(\eta^*)}, \tag{A.40}$$

where

$$\begin{aligned}\widehat{\tau}(x,\mu_S) &= \varphi\begin{pmatrix} F_R[2\widetilde{\sigma}-x](2\widetilde{\sigma}-x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-x]+f_R[x]) \\ +(1-F_R[x])(\mu_R-x)\end{pmatrix} \\ &\quad -(1-\varphi)\left(x+\frac{k_2-\Delta_0}{k_1}\right).\end{aligned} \tag{A.41}$$

Consider $\frac{\partial\widehat{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_h(\eta^*)}$. Taking the partial derivative of the right-hand side of (A.41) with respect to μ_S at $x = \sigma_h(\eta^*)$ (treating x as constant), and subsequently substituting $2\widetilde{\sigma} - \sigma_h(\eta^*) = \widetilde{\sigma} - \eta^* = \sigma_l(\eta^*)$ yields:

$$\begin{aligned}\frac{\partial\widehat{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_h(\eta^*)} &= \varphi\begin{pmatrix} 2f_l\frac{d\widetilde{\sigma}}{d\mu_S}(\sigma_l(\eta^*)-\mu_R)+2F_l\frac{d\widetilde{\sigma}}{d\mu_S} \\ +2(\gamma_R^2+\gamma_\varepsilon^2)\frac{df_l}{d\sigma_l}\frac{d\widetilde{\sigma}}{d\mu_S}\end{pmatrix} \\ &\quad -(1-\varphi)\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\mu_S}.\end{aligned} \tag{A.42}$$

By the properties of the pdf of the normal distribution,

$$\frac{df_l}{d\sigma_l} = \frac{df_R(\sigma_l)}{d\sigma_l} = \frac{\mu_R-\sigma_l}{\gamma_R^2+\gamma_\varepsilon^2}f_l. \tag{A.43}$$

Substituting this into (A.42) yields

$$\frac{\partial\widehat{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_h(\eta^*)} = 2\varphi F_l\frac{d\widetilde{\sigma}}{d\mu_S} - (1-\varphi)\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\mu_S}.$$

Taking the derivatives on the right-hand side and simplifying, we eventually obtain:

$$\frac{\partial \widehat{\tau}(x,\mu_S)}{\partial \mu_S}_{|x=\sigma_h(\eta^*)} = \frac{(\gamma_R^2+\gamma_\varepsilon^2)\left(\gamma_S^2(1-\varphi)-2F_l\gamma_\varepsilon^2\varphi\right)}{(\gamma_S^2-\gamma_R^2)\gamma_\varepsilon^2}. \tag{A.44}$$

The sign of this expression corresponds to the sign of the nominator. Since the nominator decreases in F_l while $F_l = F_R(\sigma_l) < F_R(\mu_R) = 0.5$ by Corollary A.1, we obtain

$$\gamma_S^2(1-\varphi) - 2F_l\gamma_\varepsilon^2\varphi > \gamma_S^2(1-\varphi) - \gamma_\varepsilon^2\varphi. \tag{A.45}$$

Assume $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Then, for any $\mu_S \geq \mu_R$ we have

$$\begin{aligned} & \gamma_S^2(1-\varphi) - \gamma_\varepsilon^2\varphi \geq 0 \\ \Rightarrow\ & \gamma_S^2(1-\varphi) - 2F_l\gamma_\varepsilon^2\varphi > 0 \\ \Leftrightarrow\ & \frac{\partial \widehat{\tau}(x,\mu_S)}{\partial \mu_S}_{|x=\sigma_h(\eta^*)} > 0, \end{aligned}$$

where the second line follows by (A.45), and the third line follows by (A.44). Thus, $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ is a sufficient condition for $\frac{\partial \widehat{\tau}(x,\mu_S)}{\partial \mu_S}_{|x=\sigma_h(\eta^*)} > 0$ at any $\mu_S \geq \mu_R$. It follows by (A.40) that the same is true for $\frac{d\sigma_h(\eta^*)}{d\mu_S} > 0$ at any $\mu_S \geq \mu_R$, which proves Claim 2.

Claim 2 together with Claim 1 lead to the claim of Lemma *A.12* for the case $\gamma_S > \gamma_R$ and $\mu_S \geq \mu_R$.

Case 2: $\gamma_S < \gamma_R$.

The proof proceeds analogously to Case 1.

Claim 1: A sufficient condition for the claim of Lemma A.12 under $\mu_S \geq \mu_R$ *and* $\gamma_S < \gamma_R$ *is*

$$\frac{d\sigma_l(\eta^*)}{d\mu_S} < 0 \textit{ at any } \mu_S \geq \mu_R.$$

Proof. Assume $\frac{d\sigma_l(\eta^*)}{d\mu_S} < 0$ at any $\mu_S \geq \mu_R$. Then, given that $\sigma_h(\eta^*) = \widetilde{\sigma} + \eta^* = 2\widetilde{\sigma} - \sigma_l(\eta^*)$, we also have $\frac{d\sigma_h(\eta^*)}{d\mu_S} = 2\frac{d\widetilde{\sigma}}{d\mu_S} - \frac{d\sigma_l(\eta^*)}{d\mu_S} > 0$ at any $\mu_S \geq \mu_R$ (since $\frac{d\widetilde{\sigma}}{d\mu_S} = \frac{\gamma_R^2+\gamma_\varepsilon^2}{\gamma_R^2-\gamma_S^2} > 0$ under $\gamma_S < \gamma_R$). Thus, the lower (upper) bound of the equilibrium disclosure interval strictly decreases (increases) when prior disagreement $|\mu_R - \mu_S|$ strictly increases.

Claim 2: If $\gamma_S < \gamma_R$ *and* $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$*, then* $\frac{d\sigma_l(\eta^*)}{d\mu_S} < 0$ *at any* $\mu_S \geq \mu_R$.

Proof. By Lemma A.11,

$$\frac{d\sigma_l(\eta^*)}{d\mu_S} = -\frac{1}{1-\varphi(F_h - F_l)} \frac{\partial \widehat{\tau}(x,\mu_S)}{\partial \mu_S}_{|x=\sigma_l(\eta^*)}, \tag{A.46}$$

where

$$\widetilde{\tau}(x,\mu_S) = \varphi \begin{pmatrix} F_R[x](x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[x]+f_R[2\widetilde{\sigma}-x]) \\ +(1-F_R[2\widetilde{\sigma}-x])(\mu_R-2\widetilde{\sigma}+x) \end{pmatrix} -(1-\varphi)\left(2\widetilde{\sigma}-x+\frac{k_2-\Delta_0}{k_1}\right). \tag{A.47}$$

Consider $\frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)}$. Taking the partial derivative of the right-hand side of (A.47) with respect to μ_S at $x=\sigma_l(\eta^*)$ (treating x as constant), and subsequently substituting $2\widetilde{\sigma}-\sigma_l(\eta^*)=\widetilde{\sigma}+\eta^*=\sigma_h(\eta^*)$ yields:

$$\frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)} = \varphi \begin{pmatrix} 2(\gamma_R^2+\gamma_\varepsilon^2)\frac{df_h}{d\sigma_h}\frac{d\widetilde{\sigma}}{d\mu_S} \\ -2f_h\frac{d\widetilde{\sigma}}{d\mu_S}(\mu_R-\sigma_h(\eta^*)) \\ -2(1-F_h)\frac{d\widetilde{\sigma}}{d\mu_S} \end{pmatrix} -(1-\varphi)\left(2\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\mu_S}\right). \tag{A.48}$$

By the properties of the pdf of the normal distribution,

$$\frac{df_h}{d\sigma_h}=\frac{df_R(\sigma_h)}{d\sigma_h}=\frac{\mu_R-\sigma_h}{\gamma_R^2+\gamma_\varepsilon^2}f_h. \tag{A.49}$$

Substituting this into (A.48) yields

$$\frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)} = -2\varphi(1-F_h)\frac{d\widetilde{\sigma}}{d\mu_S}-(1-\varphi)\left(2\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\mu_S}\right).$$

Taking the derivatives on the right-hand side and simplifying, we eventually obtain:

$$\frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)} = \frac{(\gamma_R^2+\gamma_\varepsilon^2)\left(\gamma_S^2(1-\varphi)-2(1-F_h)\gamma_\varepsilon^2\varphi\right)}{(\gamma_R^2-\gamma_S^2)\gamma_\varepsilon^2}. \tag{A.50}$$

The sign of this expression corresponds to the sign of the nominator. Since the nominator increases in F_h while $F_h=F_R(\sigma_h)>F_R(\mu_R)=0.5$ by Corollary A.1, we obtain

$$\gamma_S^2(1-\varphi)-2(1-F_h)\gamma_\varepsilon^2\varphi>\gamma_S^2(1-\varphi)-\gamma_\varepsilon^2\varphi. \tag{A.51}$$

Assume $\frac{\gamma_S^2}{\gamma_\varepsilon^2}\geq\frac{\varphi}{1-\varphi}$. Then, for any $\mu_S\geq\mu_R$ we have

$$\begin{aligned} & \gamma_S^2(1-\varphi)-\gamma_\varepsilon^2\varphi\geq 0 \\ \Rightarrow\ & \gamma_S^2(1-\varphi)-2(1-F_h)\gamma_\varepsilon^2\varphi>0 \\ \Leftrightarrow\ & \frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)}>0, \end{aligned}$$

where the second line follows by (A.51), and the third line follows by (A.50). Thus, $\frac{\gamma_S^2}{\gamma_\varepsilon^2}\geq\frac{\varphi}{1-\varphi}$ is a sufficient condition for $\frac{\partial\widetilde{\tau}(x,\mu_S)}{\partial\mu_S}_{|x=\sigma_l(\eta^*)}>0$ at any $\mu_S\geq\mu_R$. It follows by (A.46) that the same is true for $\frac{d\sigma_l(\eta^*)}{d\mu_S} < 0$ at any $\mu_S \geq \mu_R$, which proves Claim 2.

Claim 2 together with Claim 1 lead to the claim of the lemma for the case $\gamma_S < \gamma_R$ and $\mu_S \geq \mu_R$.

Step 2. Let us show the claim of Lemma *A.12* for the remaining case $\mu_S < \mu_R$, given that the claim holds for $\mu_S \geq \mu_R$ as shown in Step 1.

Let $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Denote the lower and upper bounds of the disclosure interval given μ_S as $\sigma_l[\mu_S]$ and $\sigma_h[\mu_S]$, respectively. Denote $\mu'_S = 2\mu_R - \mu_S$ so that μ'_S and μ_S are equidistant to μ_R. Then, by Lemma A.9, it holds

$$\sigma_l[\mu_S] = 2\mu_R - \sigma_h[\mu'_S].$$

Consequently,

$$\frac{d\sigma_l[\mu_S]}{d\mu_S} = -\frac{d\sigma_h[\mu'_S]}{d\mu'_S}\frac{d\mu'_S}{d\mu_S} = \frac{d\sigma_h[\mu'_S]}{d\mu'_S} > 0,$$

where the inequality follows by the fact that $\mu'_S > \mu_R$ (due to $\mu_S < \mu_R$ by assumption), while for any $\mu'_S \geq \mu_R$ it holds $\frac{d\sigma_h[\mu'_S]}{d\mu'_S} > 0$ (if $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$) as shown in Step 1.

By the same argument, $\sigma_h[\mu_S] = 2\mu_R - \sigma_l[\mu'_S]$ so that

$$\frac{d\sigma_h[\mu_S]}{d\mu_S} = -\frac{d\sigma_l[\mu'_S]}{d\mu'_S}\frac{d\mu'_S}{d\mu_S} = \frac{d\sigma_l[\mu'_S]}{d\mu'_S} < 0.$$

Thus, for any $\mu_S < \mu_R$ and $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, $\sigma_l[\mu_S]$ ($\sigma_h[\mu_S]$) strictly increases (decreases) as μ_S increases, i.e., as the prior disagreement $|\mu_S - \mu_R|$ decreases. This is equivalent to the claim of the lemma in the considered case. ■

Lemma A.13 *Let $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Fix the receiver's prior (μ_R, γ_R) as well as the sender's prior opinion μ_S, with $\mu_S \neq \mu_R$. Then, if $\gamma_S \neq \gamma_R$ and the order of γ_R and γ_S is fixed, the lower (upper) bound of the equilibrium disclosure interval strictly increases (decreases) when the absolute difference $|\gamma_R - \gamma_S|$ strictly increases.*

Proof.

Step 1. This shows the claim of the lemma for the case $\mu_S > \mu_R$. Herewith, we separately consider two possible parameter cases: $\gamma_S > \gamma_R$ and $\gamma_S < \gamma_R$.

Case 1: $\gamma_S > \gamma_R$.

Claim 1: *A sufficient condition for the claim of Lemma A.13 under $\gamma_S > \gamma_R$ and $\mu_S > \mu_R$ is*

$$\frac{d\sigma_h(\eta^*)}{d\gamma_S} < 0 \text{ at any } \gamma_S > \gamma_R.$$

Proof. Assume $\frac{d\sigma_h(\eta^*)}{d\gamma_S} < 0$ at any $\gamma_S > \gamma_R$. Then, given that $\sigma_l(\eta^*) = \widetilde{\sigma} - \eta^* = 2\widetilde{\sigma} - \sigma_h(\eta^*)$, we also have $\frac{d\sigma_l(\eta^*)}{d\gamma_S} = 2\frac{d\widetilde{\sigma}}{d\gamma_S} - \frac{d\sigma_h(\eta^*)}{d\gamma_S} > 0$ at any $\gamma_S > \gamma_R$ (since $\frac{d\widetilde{\sigma}}{d\gamma_S} = \frac{2\gamma_S(\gamma_R^2+\gamma_\varepsilon^2)(\mu_S-\mu_R)}{(\gamma_S^2-\gamma_R^2)^2} > 0$ under the current parameter assumptions). Thus, the lower (upper) bound of the equilibrium disclosure interval strictly increases (decreases) when the absolute difference $|\gamma_R - \gamma_S|$ strictly increases.

Claim 2: *If* $\mu_S > \mu_R$ *and* $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, *then* $\frac{d\sigma_h(\eta^*)}{d\gamma_S} < 0$ *at any* $\gamma_S > \gamma_R$.

Proof. By Lemma A.11,

$$\frac{d\sigma_h(\eta^*)}{d\gamma_S} = \frac{1}{1-\varphi(F_h - F_l)} \frac{\partial \widehat{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)}, \tag{A.52}$$

where

$$\begin{aligned} \widehat{\tau}(x,\gamma_S) &= \varphi \begin{pmatrix} F_R[2\widetilde{\sigma} - x](2\widetilde{\sigma} - x - \mu_R) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)(f_R[2\widetilde{\sigma} - x] + f_R[x]) \\ +(1 - F_R[x])(\mu_R - x) \end{pmatrix} \\ & \quad -(1-\varphi)\left(x + \frac{k_2 - \Delta_0}{k_1}\right). \end{aligned} \tag{A.53}$$

Consider $\frac{\partial \widehat{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)}$. Taking the partial derivative of the right-hand side of (A.53) with respect to γ_S at $x = \sigma_h(\eta^*)$ (treating x as constant), and subsequently substituting $2\widetilde{\sigma} - \sigma_h(\eta^*) = \widetilde{\sigma} - \eta^* = \sigma_l(\eta^*)$ yields:

$$\begin{aligned} \frac{\partial \widehat{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)} &= \varphi \begin{pmatrix} 2f_l \frac{d\widetilde{\sigma}}{d\gamma_S}(\sigma_l(\eta^*) - \mu_R) + 2F_l \frac{d\widetilde{\sigma}}{d\gamma_S} \\ +2(\gamma_R^2 + \gamma_\varepsilon^2)\frac{df_l}{d\sigma_l}\frac{d\widetilde{\sigma}}{d\gamma_S} \end{pmatrix} \\ & \quad -(1-\varphi)\frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \gamma_S} \end{aligned} \tag{A.54}$$

By the properties of the pdf of the normal distribution,

$$\frac{df_l}{d\sigma_l} = \frac{df_R(\sigma_l)}{d\sigma_l} = \frac{\mu_R - \sigma_l}{\gamma_R^2 + \gamma_\varepsilon^2} f_l. \tag{A.55}$$

Substituting this into (A.54) yields

$$\frac{\partial \widehat{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)} = 2\varphi F_l \frac{d\widetilde{\sigma}}{d\gamma_S} - (1-\varphi)\frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \gamma_S}.$$

Taking the derivatives on the right-hand side and simplifying, we eventually obtain:

$$\frac{\partial \widehat{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)} = \frac{2\gamma_S(\gamma_R^2 + \gamma_\varepsilon^2)(\mu_S - \mu_R)(2F_l\gamma_\varepsilon^2\varphi - \gamma_R^2(1-\varphi))}{(\gamma_S^2 - \gamma_R^2)^2\gamma_\varepsilon^2}. \tag{A.56}$$

The sign of this expression corresponds to the sign of the term $2F_l\gamma_\varepsilon^2\varphi - \gamma_R^2(1-\varphi)$. Since it increases in F_l while $F_l = F_R(\sigma_l) < F_R(\mu_R) = 0.5$ by Corollary A.1, we obtain

$$2F_l\gamma_\varepsilon^2\varphi - \gamma_R^2(1-\varphi) < \gamma_\varepsilon^2\varphi - \gamma_R^2(1-\varphi). \tag{A.57}$$

Assume $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Then, for any $\gamma_S > \gamma_R$ we have

$$\begin{aligned} & \gamma_\varepsilon^2 \varphi - \gamma_R^2(1-\varphi) \leq 0 \\ \Rightarrow \quad & 2F_l \gamma_\varepsilon^2 \varphi - \gamma_R^2(1-\varphi) < 0 \\ \Leftrightarrow \quad & \frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)} < 0, \end{aligned}$$

where the second line follows by (A.57), and the third line follows by (A.56). Thus, $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ is a sufficient condition for $\frac{\partial \widetilde{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_h(\eta^*)} < 0$ at any $\gamma_S > \gamma_R$. It follows by (A.52) that the same is true for $\frac{d\sigma_h(\eta^*)}{d\gamma_S} < 0$ at any $\gamma_S > \gamma_R$, which proves Claim 2.

Claim 2 together with Claim 1 lead to the claim of Lemma A.13 for the case $\gamma_S > \gamma_R$ and $\mu_S > \mu_R$.

Case 2: $\gamma_S < \gamma_R$.

The proof proceeds analogously to Case 1.

Claim 1: A sufficient condition for the claim of Lemma A.13 under $\gamma_S < \gamma_R$ and $\mu_S > \mu_R$ is

$$\frac{d\sigma_l(\eta^*)}{d\gamma_S} < 0 \text{ at any } \gamma_S < \gamma_R.$$

Proof. Assume $\frac{d\sigma_l(\eta^*)}{d\gamma_S} < 0$ at any $\gamma_S < \gamma_R$. Then, given that $\sigma_h(\eta^*) = \widetilde{\sigma} + \eta^* = 2\widetilde{\sigma} - \sigma_l(\eta^*)$, we also have $\frac{d\sigma_h(\eta^*)}{d\gamma_S} = 2\frac{d\widetilde{\sigma}}{d\gamma_S} - \frac{d\sigma_l(\eta^*)}{d\gamma_S} > 0$ at any $\gamma_S < \gamma_R$ (since $\frac{d\widetilde{\sigma}}{d\gamma_S} = \frac{2\gamma_S(\gamma_R^2+\gamma_\varepsilon^2)(\mu_S-\mu_R)}{(\gamma_S^2-\gamma_R^2)^2} > 0$ under the current parameter assumptions). Thus, the lower (upper) bound of the equilibrium disclosure interval strictly decreases (increases) when the absolute difference $|\gamma_R - \gamma_S|$ strictly decreases.

Claim 2: If $\mu_S > \mu_R$ and $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, then $\frac{d\sigma_l(\eta^*)}{d\gamma_S} < 0$ at any $\gamma_S < \gamma_R$.

Proof. By Lemma A.11,

$$\frac{d\sigma_l(\eta^*)}{d\gamma_S} = -\frac{1}{1-\varphi(F_h - F_l)} \frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)}, \tag{A.58}$$

where

$$\begin{aligned} \widetilde{\tau}(x, \gamma_S) \quad = \quad & \varphi \begin{pmatrix} F_R[x](x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[x] + f_R[2\widetilde{\sigma} - x]) \\ +(1 - F_R[2\widetilde{\sigma} - x])(\mu_R - 2\widetilde{\sigma} + x) \end{pmatrix} \\ & -(1-\varphi)\left(2\widetilde{\sigma} - x + \frac{k_2 - \Delta_0}{k_1}\right). \end{aligned} \tag{A.59}$$

Consider $\frac{\partial \widetilde{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)}$. Taking the partial derivative of the right-hand side of (A.59) with respect to γ_S at $x = \sigma_l(\eta^*)$ (treating x as constant), and subsequently substituting

$2\widetilde{\sigma} - \sigma_l(\eta^*) = \widetilde{\sigma} + \eta^* = \sigma_h(\eta^*)$ yields:

$$\frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)} = \varphi \begin{pmatrix} -2f_h \frac{d\widetilde{\sigma}}{d\gamma_S}(\mu_R - \sigma_h(\eta^*)) \\ -2(1-F_h)\frac{d\widetilde{\sigma}}{d\gamma_S} \\ +2(\gamma_R^2 + \gamma_\varepsilon^2)\frac{df_h}{d\sigma_h}\frac{d\widetilde{\sigma}}{d\gamma_S} \end{pmatrix} - (1-\varphi)\left(2\frac{d\widetilde{\sigma}}{d\gamma_S} + \frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \gamma_S}\right). \tag{A.60}$$

By the properties of the pdf of the normal distribution,

$$\frac{df_h}{d\sigma_h} = \frac{df_R(\sigma_h)}{d\sigma_h} = \frac{\mu_R - \sigma_h}{\gamma_R^2 + \gamma_\varepsilon^2} f_h. \tag{A.61}$$

Substituting this into (A.60) yields

$$\frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)} = -2\varphi(1-F_h)\frac{d\widetilde{\sigma}}{d\gamma_S} - (1-\varphi)\left(2\frac{d\widetilde{\sigma}}{d\gamma_S} + \frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \gamma_S}\right).$$

Taking the derivatives on the right-hand side and simplifying, we eventually obtain:

$$\frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)} = \frac{2\gamma_S(\gamma_R^2 + \gamma_\varepsilon^2)(\mu_S - \mu_R)(\gamma_R^2(1-\varphi) - 2(1-F_h)\gamma_\varepsilon^2\varphi)}{(\gamma_R^2 - \gamma_S^2)^2\gamma_\varepsilon^2}. \tag{A.62}$$

The sign of this expression corresponds to the sign of the term $\gamma_R^2(1-\varphi) - 2(1-F_h)\gamma_\varepsilon^2\varphi$. Since it increases in F_h while $F_h = F_R(\sigma_h) > F_R(\mu_R) = 0.5$ by Corollary A.1, we obtain

$$\gamma_R^2(1-\varphi) - 2(1-F_h)\gamma_\varepsilon^2\varphi > \gamma_R^2(1-\varphi) - \gamma_\varepsilon^2\varphi. \tag{A.63}$$

Assume $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Then, for any $\gamma_S < \gamma_R$ we have

$$\begin{aligned} & \gamma_R^2(1-\varphi) - \gamma_\varepsilon^2\varphi \geq 0 \\ \Rightarrow\ & \gamma_R^2(1-\varphi) - 2(1-F_h)\gamma_\varepsilon^2\varphi > 0 \\ \Leftrightarrow\ & \frac{\partial \widetilde{\tau}(x, \gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)} > 0, \end{aligned}$$

where the second line follows by (A.63), and the third line follows by (A.62). Thus, $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ is a sufficient condition for $\frac{\partial \widetilde{\tau}(x,\gamma_S)}{\partial \gamma_S}_{|x=\sigma_l(\eta^*)} > 0$ at any $\gamma_S < \gamma_R$. It follows by (A.58) that the same is true for $\frac{d\sigma_l(\eta^*)}{d\gamma_S} < 0$ at any $\gamma_S < \gamma_R$, which proves Claim 2.

Claim 2 together with Claim 1 lead to the claim of Lemma A.13 for the case $\gamma_S < \gamma_R$ and $\mu_S > \mu_R$.

Step 2. Let us show the claim of Lemma A.13 for the remaining case $\mu_S < \mu_R$, given that the claim holds for $\mu_S > \mu_R$ as shown in Step 1.

Let $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$. Denote the lower and upper bounds of the disclosure interval given μ_S and γ_S as $\sigma_l[\mu_S, \gamma_S]$ and $\sigma_h[\mu_S, \gamma_S]$, respectively. Denote $\mu_S' = 2\mu_R - \mu_S$ so that μ_S' and μ_S

are equidistant to μ_R. Then, by Lemma A.9, it holds

$$\sigma_l[\mu_S, \gamma_S] = 2\mu_R - \sigma_h[\mu'_S, \gamma_S].$$

Consequently,

$$\frac{d\sigma_l[\mu_S, \gamma_S]}{d\gamma_S} = -\frac{d\sigma_h[\mu'_S, \gamma_S]}{d\gamma_S} \gtrless 0 \text{ if } \gamma_S \gtrless \gamma_R,$$

where the inequality follows by the fact that $\mu'_S > \mu_R$ (due to $\mu_S < \mu_R$ by assumption), while for any $\mu'_S > \mu_R$ it holds $\frac{d\sigma_h[\mu'_S,\gamma_S]}{d\gamma_S} \lessgtr 0$ if $\gamma_S \gtrless \gamma_R$ (under $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$) as shown in Step 1.

By the same argument, $\sigma_h[\mu_S, \gamma_S] = 2\mu_R - \sigma_l[\mu'_S, \gamma_S]$ so that

$$\frac{d\sigma_h[\mu_S, \gamma_S]}{d\gamma_S} = -\frac{d\sigma_l[\mu'_S, \gamma_S]}{d\gamma_S} \lessgtr 0 \text{ if } \gamma_S \gtrless \gamma_R.$$

Thus, for any $\mu_S < \mu_R$ and $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$, $\sigma_l[\mu_S]$ ($\sigma_h[\mu_S]$) strictly increases (decreases) as the absolute difference $|\gamma_S - \gamma_R|$ increases. This is equivalent to the claim of the lemma in the considered case. ■

A.3.2 Proof of Proposition 3

Step 1. Let us show the claim of the proposition if μ_S changes in such way that the order of μ_S and μ_R remains fixed.

Let $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and $\gamma_S \neq \gamma_R$. Then, by Lemma A.12, if μ_S changes in such way that the order of μ_R and μ_S is unchanged while $|\mu_R - \mu_S|$ strictly increases, then the equilibrium disclosure interval strictly expands on both sides (i.e., σ_l strictly decreases while σ_h strictly increases). It follows by Lemma A.7 that S's disclosure strategy becomes more Blackwell informative in this case. Further, by Lemma A.8, R's ex ante expected utility strictly increases.

Step 2. It is left to show that claim b) of the proposition regarding R's ex ante expected utility holds independently of the order of μ_S and μ_R. In other words, we need to show that if $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and $\gamma_S \neq \gamma_R$, then for any μ'_S, μ''_S it holds $E[u_R|\mu'_S] \lessgtr E[u_R|\mu''_S]$ if $|\mu_R - \mu'_S| \lessgtr |\mu_R - \mu''_S|$.

Without loss of generality, assume $\mu'_S < \mu''_S$. The claim for the fixed order of μ_S and μ_R, i.e., for the cases $\mu'_S < \mu''_S \leq \mu_R$ and $\mu_R \leq \mu'_S < \mu''_S$, was shown in Step 1. Consider the last case $\mu'_S < \mu_R < \mu''_S$ under assumptions $\frac{\gamma_S^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and $\gamma_S \neq \gamma_R$.

Denote $\widehat{\mu}'_S = 2\mu_R - \mu'_S$ so that $\widehat{\mu}'_S$ and μ'_S are equidistant to μ_R. By Lemma A.9, the disclosure intervals under $\widehat{\mu}'_S$ and μ'_S are symmetric around μ_R. Then, by symmetry considerations, R obtains the same ex ante expected loss $E_R[(a-\omega)^2]$ under these symmetric disclosure intervals:

$$E[u_R|\mu'_S] = E[u_R|\widehat{\mu}'_S]. \tag{A.64}$$

Since, by assumption, $\mu'_S < \mu_R \Leftrightarrow \widehat{\mu}'_S > \mu_R$ and $\mu''_S > \mu_R$ (i.e., $\widehat{\mu}'_S$ and μ''_S are both on the same side of μ_R), by Step 1 it holds

$$E[u_R|\widehat{\mu}'_S] \lessgtr E[u_R|\mu''_S] \text{ if } |\mu_R - \widehat{\mu}'_S| \lessgtr |\mu_R - \mu''_S| \Leftrightarrow |\mu_R - \mu'_S| \lessgtr |\mu_R - \mu''_S|.$$

Since $E[u_R|\widehat{\mu}'_S] = E[u_R|\mu'_S]$ by (A.64), this implies

$$E[u_R|\mu'_S] \lesseqgtr E[u_R|\mu''_S] \text{ if } |\mu_R - \mu'_S| \lesseqgtr |\mu_R - \mu''_S|.$$

■

A.3.3 Proof of Proposition 4, Point a)

Let $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and $\mu_S \neq \mu_R$. Then, claim a) of Proposition 4 is equivalent to showing that if γ_S changes to some γ'_S such that the order of S's and R's confidences remains unchanged while $|\gamma_S - \gamma_R| < |\gamma'_S - \gamma_R|$, then the equilibrium disclosure strategy becomes less Blackwell informative and the receiver's expected utility strictly decreases. Below we show this claim for two parameter cases depending on whether γ_S is equal to γ_R.

Case 1. $\gamma_S \neq \gamma_R$.

Let $\frac{\gamma_R^2}{\gamma_\varepsilon^2} \geq \frac{\varphi}{1-\varphi}$ and $\mu_S \neq \mu_R$. Then, by Lemma A.13, the equilibrium disclosure interval under γ'_S is a strict subset of the equilibrium disclosure interval under γ_S. Consequently, S's disclosure strategy is less Blackwell informative and R's ex ante expected utility is strictly lower under γ'_S than under γ_S by Lemmas A.7 and A.8, respectively.

Case 2. $\gamma_S = \gamma_R$.

By Propositions 1 and 2, there is full disclosure under $\gamma_S = \gamma_R$ and a partial disclosure under $\gamma'_S \neq \gamma_R$. Consequently, the disclosure strategy is less Blackwell informative while R obtains a strictly lower ex ante expected utility under γ'_S than under γ_S by Lemmas A.7(b) and A.8(b), respectively. ■

A.3.4 Proof of Proposition 4, Point b)

Without loss of generality, normalize $\mu_R = 0$. Consider arbitrary R's prior $(0, \widetilde{\gamma}_R)$ and S's prior $(\widetilde{\mu}_S, \widetilde{\gamma}_S)$ such that $\widetilde{\mu}_S > 0$ and $\widetilde{\gamma}_S \in (0, \widetilde{\gamma}_R)$. In Steps 1-3, we are going to show that R's expected utility is strictly lower under $(\widetilde{\mu}_S, \widetilde{\gamma}_S)$ than under $(\widetilde{\mu}_S, \sqrt{2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2})$ (i.e., under larger prior variance that is equidistant to $\widetilde{\gamma}_R^2$), which is equivalent to the claim of the proposition for the case $\widetilde{\mu}_S > 0$. The remaining case $\widetilde{\mu}_S < 0$ is considered in Step 4 (recall that $\widetilde{\mu}_S \neq \mu_R$ by assumption).

Step 1. Recall that the value of the 0-disagreement signal under assumed priors (once μ_R is normalized to 0) is:

$$\widetilde{\sigma}(\widetilde{\mu}_S, \widetilde{\gamma}_S) = \frac{\widetilde{\mu}_S(\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2)}{\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2} > 0.$$

Fixing $\widetilde{\gamma}_R$ and γ_ε there is a continuum of S's priors for which the corresponding 0-disagreement signal is the same as under $(\widetilde{\mu}_S, \widetilde{\gamma}_S)$. Formally, there is a function

$$\begin{aligned} g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S) \quad : \quad &= \left(\widetilde{\gamma}_R^2 - \frac{\left(\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2\right)\mu_S}{\widetilde{\sigma}(\widetilde{\mu}_S, \widetilde{\gamma}_S)}\right)^{\frac{1}{2}} \\ &= \left(\widetilde{\gamma}_R^2 - \frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2\right)\right)^{\frac{1}{2}} \end{aligned}$$

such that for any $\mu_S \leq \widetilde{\mu}_S$ it holds true that $\widetilde{\sigma}(\mu_S, \gamma_S) = \widetilde{\sigma}(\widetilde{\mu}_S, \widetilde{\gamma}_S)$ if and only if $\gamma_S = g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S)$.

Note the following properties of the function $g()$. It is continuous and differentiable in μ_S for $\mu_S \leq \widetilde{\mu}_S \neq 0$ (where the inequality holds since $\widetilde{\mu}_S > \mu_R = 0$ by assumption). Besides, $g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S) > 0$ for any $\mu_S \leq \widetilde{\mu}_S$ if $\widetilde{\mu}_S > 0$ and $\widetilde{\gamma}_S < \widetilde{\gamma}_R$. Further, for $\mu_S = -\widetilde{\mu}_S$ we have

$$\begin{aligned} g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S) &= \left(\widetilde{\gamma}_R^2 - \frac{-\widetilde{\mu}_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2\right)\right)^{\frac{1}{2}} \\ &= (2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2)^{\frac{1}{2}}, \end{aligned} \tag{A.65}$$

which is in turn equivalent to stating:

$$g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S)^2 - \widetilde{\gamma}_R^2 = \widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2.$$

In words, S's priors $(\widetilde{\mu}_S, \widetilde{\gamma}_S)$ and $(-\widetilde{\mu}_S, g(-\widetilde{\mu}_S))$ have variances that are equidistant from R's prior variance $\widetilde{\gamma}_R^2$.

To conduct our proof, we shall consider three S's priors $A = (\widetilde{\mu}_S, \widetilde{\gamma}_S)$, $B = (-\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$ and $C = (\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S)) = (\widetilde{\mu}_S, \left(2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2\right)^{\frac{1}{2}})$, where the latter equality is by (A.65). We shall show the following. First, when switching from A to B, the equilibrium expected utility of R strictly increases because the disclosure interval becomes a superset of the original disclosure interval. Second, when switching from B to C, the equilibrium expected utility of R remains fixed. It follows by transitivity that when moving from A to C, the equilibrium expected utility of R strictly increases. Given that A and C have the same means while variances are equidistant from $\widetilde{\gamma}_R^2$, this will prove the claim of the proposition for the case $\mu_S \geq 0$.

Step 2. We here consider the move from $A = (\widetilde{\mu}_S, \widetilde{\gamma}_S)$ to $B = (-\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$. We need to show that the disclosure interval under B is a superset of the disclosure interval under A.

By Lemma A.4, given priors (μ_S, γ_S) and $(\widetilde{\mu}_R, \widetilde{\gamma}_R)$ it holds in equilibrium

$$\begin{aligned} & \varphi \begin{pmatrix} F_l(\sigma_l - \widetilde{\mu}_R) + (\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1 - F_h)(\widetilde{\mu}_R - \sigma_h) \end{pmatrix} - (1 - \varphi)\left(\sigma_h + \frac{k_2 - \Delta_0}{k_1}\right) = 0 \\ \Leftrightarrow \;\; & \varphi F_l(\sigma_l - E_l) + \varphi(1 - F_h)(E_h - \sigma_h) - (1 - \varphi)\left(\sigma_h + \frac{k_2 - \Delta_0}{k_1}\right) = 0, \end{aligned} \tag{A.66}$$

where the equivalence is by Lemma A.1. Denote

$$\begin{aligned} & \chi(\mu_S, \gamma_S) \\ = \;& \frac{k_2(\mu_S, \gamma_S, \widetilde{\mu}_R, \widetilde{\gamma}_R) - \Delta_0(\mu_S, \widetilde{\mu}_R)}{k_1(\mu_S, \gamma_S, \widetilde{\mu}_R, \widetilde{\gamma}_R)} \\ = \;& \text{sgn}\,(\gamma_S - \widetilde{\gamma}_R)\frac{1}{\frac{\gamma_S^2}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\widetilde{\gamma}_R^2}{\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2}}\left(\text{sgn}\,(\gamma_S - \widetilde{\gamma}_R)\frac{\gamma_\varepsilon^2 \mu_S}{\gamma_S^2 + \gamma_\varepsilon^2} - |\mu_S|\right) \end{aligned} \tag{A.67}$$

Then, the equilibrium condition (A.66) given priors (μ_S, γ_S) and $(\widetilde{\mu}_R, \widetilde{\gamma}_R)$ can be rewritten as

$$\varphi F_l(\sigma_l - E_l) + \varphi(1 - F_h)(E_h - \sigma_h) - (1 - \varphi)\left(\sigma_h + \chi(\mu_S, \gamma_S)\right) = 0. \tag{A.68}$$

Consider any $\mu_S \le \widetilde{\mu}_S$ and $\gamma_S = g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$. Differentiating both sides of (A.68) with respect to μ_S (while setting $\gamma_S = g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$, we obtain, noting that $E_l \equiv E_R[\sigma|\sigma \le \sigma_l]$ and $E_h \equiv E_R[\sigma|\sigma \ge \sigma_h]$:

$$\begin{aligned}
& \varphi\left(f_l\left(\frac{d\widetilde{\sigma}}{d\mu_S}-\frac{d\eta}{d\mu_S}\right)(\sigma_l - E_l) - f_h\left(\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{d\eta}{d\mu_S}\right)(E_h-\sigma_h)\right) \\
& +\varphi\left(\begin{array}{c} F_l\left[\left(\frac{d\widetilde{\sigma}}{d\mu_S}-\frac{d\eta}{d\mu_S}\right)-\frac{dE_l}{d\sigma_l}\left(\frac{d\widetilde{\sigma}}{d\mu_S}-\frac{d\eta}{d\mu_S}\right)\right] \\ +(1-F_h)\left[\frac{dE_h}{d\sigma_h}\left(\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{d\eta}{d\mu_S}\right)-\left(\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{d\eta}{d\mu_S}\right)\right]\end{array}\right) \\
& -(1-\varphi)\left[\left(\frac{d\widetilde{\sigma}}{d\mu_S}+\frac{d\eta}{d\mu_S}\right)+\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S}\right] \\
= \;& 0
\end{aligned}$$

Note that $\frac{d\widetilde{\sigma}}{d\mu_S} = 0$ by construction (since γ_S is set to be equal to $g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$. Hence, we can reorganize the above inequality as:

$$\frac{d\eta}{d\mu_S} = \frac{(1-\varphi)\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S}}{\varphi\left(\begin{array}{c} -f_l(\sigma_l - E_l) - f_h(E_h - \sigma_h) \\ -F_l\left(1-\frac{dE_l}{d\sigma_l}\right)-(1-F_h)\left(1-\frac{dE_h}{d\sigma_h}\right)\end{array}\right)-(1-\varphi)}. \tag{A.69}$$

Consider $\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S}$. If $\mu_S \ge 0$, then $g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S) = \left(\widetilde{\gamma}_R^2 - \frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2\right)\right)^{\frac{1}{2}} \le \widetilde{\gamma}_R$. Then, by (A.67) we obtain

$$\begin{aligned}
\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S} &= \frac{d}{d\mu_S}\left(-\frac{-\frac{\gamma_\varepsilon^2\mu_S}{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)+\gamma_\varepsilon^2}-\mu_S}{\frac{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)}{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)+\gamma_\varepsilon^2}-\frac{\widetilde{\gamma}_R^2}{\widetilde{\gamma}_R^2+\gamma_\varepsilon^2}}\right) \\
&= 1+\frac{\widetilde{\gamma}_R^2}{\gamma_\epsilon^2} > 0.
\end{aligned} \tag{A.70}$$

In the other case, if $\mu_S < 0$, then $g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S) = \left(\widetilde{\gamma}_R^2 - \frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2\right)\right)^{\frac{1}{2}} > \widetilde{\gamma}_R$. Then, by (A.67) we obtain

$$\begin{aligned}
\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S} &= \frac{d}{d\mu_S}\left(\frac{\frac{\gamma_\varepsilon^2\mu_S}{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)+\gamma_\varepsilon^2}+\mu_S}{\frac{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)}{\left(\widetilde{\gamma}_R^2-\frac{\mu_S}{\widetilde{\mu}_S}\left(\widetilde{\gamma}_R^2-\widetilde{\gamma}_S^2\right)\right)+\gamma_\varepsilon^2}-\frac{\widetilde{\gamma}_R^2}{\widetilde{\gamma}_R^2+\gamma_\varepsilon^2}}\right) \\
&= 1+\frac{\widetilde{\gamma}_R^2}{\gamma_\epsilon^2} > 0.
\end{aligned} \tag{A.71}$$

Coming back to (A.69), note that its denominator is always strictly negative (in particular, $\frac{dE_l}{d\sigma_l} < 1$ and $\frac{dE_h}{d\sigma_h} < 1$ by (A.11), (A.13) and the standard properties of the inverse Mills ratio). This together with $\frac{d\chi(\mu_S, g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S))}{d\mu_S} > 0$ as shown above implies that (A.69) is strictly negative. It follows that $\frac{d\eta}{d\mu_S} < 0$ for any $\mu_S \le \widetilde{\mu}_S$ and $\gamma_S = g(\mu_S,\widetilde{\mu}_S,\widetilde{\gamma}_S)$. Thus, if we take any sender prior $(\mu_S, g(\mu_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$ satisfying $\mu_S \leq \widetilde{\mu}_S$, including the prior B, the implied equilibrium features a disclosure interval that is a strict superset of the disclosure interval under A (noting that the 0-disagreement signal $\widetilde{\sigma}$ stays the same by construction). It follows by Lemma A.8 that R strictly prefers S with prior B over S with prior A.

Step 3. Next, consider a move from $B = (-\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$ to $C = (\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S))$. Note that B and C share the same prior variance and have means which are symmetric around μ_R. Hence, the disclosure intervals under B and C are symmetric relative to R's prior mean $\mu_R = 0$ by Lemma A.9. Then, by symmetry considerations, R derives the same ex ante expected utility $E_R[-(a-\omega)^2]$ from both.

Thus, we have shown that for any $\widetilde{\mu}_S > 0$ and $\widetilde{\gamma}_S \in (0, \widetilde{\gamma}_R)$, moving from $A = (\widetilde{\mu}_S, \widetilde{\gamma}_S)$ to $C = (\widetilde{\mu}_S, g(-\widetilde{\mu}_S, \widetilde{\mu}_S, \widetilde{\gamma}_S)) = (\widetilde{\mu}_S, (2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2)^{\frac{1}{2}})$ strictly increases R's expected utility.

Step 4. Consider the remaining case $\widetilde{\mu}_S < 0$ and $\widetilde{\gamma}_S < \widetilde{\gamma}_R$, given that the claim for $\widetilde{\mu}_S > 0$ has been shown in the previous steps.

Note that for any S's prior (μ_S, γ_S) the disclosure intervals under (μ_S, γ_S) and $(-\mu_S, \gamma_S)$ are symmetric relative to R's prior mean $\mu_R = 0$ by Lemma A.9. Then, by symmetry considerations, R derives the same expected utility under both S's priors. Consequently, for any $\widetilde{\mu}_S < 0$ and $\widetilde{\gamma}_S < \widetilde{\gamma}_R$ it holds

$$\begin{aligned} E_R[u_R|\widetilde{\mu}_S, \widetilde{\gamma}_S] &= E_R[u_R| - \widetilde{\mu}_S, \widetilde{\gamma}_S] < E_R[u_R| - \widetilde{\mu}_S, (2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2)^{\frac{1}{2}}] \\ &= E_R[u_R|\widetilde{\mu}_S, (2\widetilde{\gamma}_R^2 - \widetilde{\gamma}_S^2)^{\frac{1}{2}}], \end{aligned}$$

where the inequality is by Steps 1-3 since $-\widetilde{\mu}_S > 0$ by assumption. Thus, we again showed that shifting from $\widetilde{\gamma}_S^2 < \widetilde{\gamma}_R^2$ to a new S's prior variance that is equidistant from $\widetilde{\gamma}_R$ strictly increases R's expected utility. ■

A.3.5 Proof of Proposition 5

Step 1. This proves point a) of the Proposition. By Lemma A.4, given priors (μ_S, γ_S) and (μ_R, γ_R) it holds in equilibrium

$$\varphi \begin{pmatrix} F_l(\sigma_l - \mu_R) + (\gamma_R^2 + \gamma_\varepsilon^2)(f_l + f_h) \\ +(1 - F_h)(\mu_R - \sigma_h) \end{pmatrix} - (1-\varphi)\left(\sigma_h + \frac{k_2 - \Delta_0}{k_1}\right) = 0. \quad \text{(A.72)}$$

Rewriting the left-hand explicitly in terms of η we have

$$\begin{aligned} &\varphi \begin{pmatrix} F_R[\widetilde{\sigma} - \eta](\widetilde{\sigma} - \eta - \mu_R) \\ +(\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2)(f_R[\widetilde{\sigma} - \eta] + f_R[\widetilde{\sigma} + \eta]) \\ +(1 - F_R[\widetilde{\sigma} + \eta])(\mu_R - \widetilde{\sigma} - \eta) \end{pmatrix} \\ &\quad -(1-\varphi)\left(\widetilde{\sigma} + \eta + \frac{k_2 - \Delta_0}{k_1}\right) = 0. \end{aligned} \quad \text{(A.73)}$$

Further, note that for $\mu_S = \mu_R$

$$\begin{aligned}
\frac{k_2 - \Delta_0}{k_1} &= \frac{\text{sgn}\ (\gamma_S - \gamma_R)\left(\frac{\gamma_\varepsilon^2 \mu_R}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2 \mu_R}{\gamma_R^2+\gamma_\varepsilon^2}\right)}{\text{sgn}\ (\gamma_S - \gamma_R)\left(\frac{\gamma_S^2}{\gamma_S^2+\gamma_\varepsilon^2} - \frac{\gamma_R^2}{\gamma_R^2+\gamma_\varepsilon^2}\right)} = -\mu_R, \\
\widetilde{\sigma} &= \frac{\mu_R(\gamma_R^2+\gamma_\varepsilon^2) - \mu_R(\gamma_S^2+\gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \mu_R.
\end{aligned}$$

Substituting this into the equilibrium condition (A.73) yields

$$\varphi\left(\begin{array}{c} -F_R[\mu_R - \eta]\eta \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[\mu_R-\eta] + f_R[\mu_R+\eta]) \\ -(1-F_R[\mu_R+\eta])\eta \end{array}\right) - (1-\varphi)\eta = 0.$$

Note that the left-hand side does not depend on γ_S. Consequently, the equilibrium solution for η (unique by Proposition 1(a)) does not depend on γ_S. It follows that the equilibrium disclosure interval (and hence R's ex ante expected utility), $[\widetilde{\sigma}-\eta^*, \widetilde{\sigma}+\eta^*] = [\mu_R-\eta^*, \mu_R+\eta^*]$, does not depend on γ_S.

Step 2. Let us show point b) of the proposition. By Propositions 2 and 1(a), $\gamma_S = \gamma_R$ ($\gamma_S \neq \gamma_R$) implies full (partial) disclosure in equilibrium. Consequently, R's expected utility is strictly higher under $\gamma_S = \gamma_R$ by Lemma A.8(b). ■

A.4 Sender preferences over receivers

A.4.1 Preliminaries

Lemma A.14 *Let $\mu_S \geq \mu_R$. Then,* $\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} = \begin{cases} -\frac{\gamma_S^2(1-\varphi)+\gamma_\varepsilon^2(1-\varphi(F_h+F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)} & \text{if } \gamma_S > \gamma_R \\ -\frac{\gamma_S^2(1-\varphi)+\gamma_\varepsilon^2(1-\varphi(2-F_h-F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)} & \text{if } \gamma_S < \gamma_R \end{cases}.$

Proof. Since in equilibrium $\widetilde{\Delta}(\varnothing|\eta^*) = \Delta(\sigma_h)$ by Lemma A.3, we have (given that $\frac{dk_1}{d\mu_R} = 0$)

$$\begin{aligned}
\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} &= \frac{d\Delta(\sigma_h)}{d\mu_R} = \frac{d(k_1\sigma_h + k_2)}{d\mu_R} \\
&= k_1 \frac{d\sigma_h}{d\mu_R} + \frac{dk_2}{d\mu_R} \\
&= k_1\left(\frac{1}{1-\varphi(F_h - F_l)} \frac{\partial\widehat{\tau}(x,\mu_R)}{\partial\mu_R}_{|x=\sigma_h(\eta^*)}\right) + \frac{dk_2}{d\mu_R}, \qquad \text{(A.74)}
\end{aligned}$$

where the second equality is by (A.20), and the last equality is by Lemma A.11.

Let us compute $\frac{\partial\widehat{\tau}(x,\mu_R)}{\partial\mu_R}_{|x=\sigma_h(\eta^*)}$. By definition in Lemma A.11 we have

$$\widehat{\tau}(x,\mu_R) = \varphi\left(\begin{array}{c} F_R[2\widetilde{\sigma} - x](2\widetilde{\sigma} - x - \mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-x] + f_R[x]) \\ +(1-F_R[x])(\mu_R - x) \end{array}\right) - (1-\varphi)\left(x + \frac{k_2-\Delta_0}{k_1}\right).$$

Taking the partial derivative of the right-hand side with respect to μ_R at $x = \sigma_h(\eta^*)$ (treating

x as constant), and subsequently substituting $2\widetilde{\sigma} - \sigma_h(\eta^*) = \widetilde{\sigma} - \eta^* = \sigma_l(\eta^*)$ yields:

$$\begin{aligned}
& \frac{\partial \widehat{\tau}(x, \mu_R)}{\partial \mu_R}_{|x=\sigma_h(\eta^*)} \\
= \; & \varphi \begin{pmatrix} \left(\frac{\partial F_l}{\partial \mu_R} + 2 f_l \frac{d\widetilde{\sigma}}{d\mu_R}\right)(\sigma_l(\eta^*) - \mu_R) + F_l\left(2\frac{d\widetilde{\sigma}}{d\mu_R} - 1\right) \\ +(\gamma_R^2 + \gamma_\varepsilon^2)\left(\frac{\partial f_l}{\partial \mu_R} + 2\frac{df_l}{d\sigma_l}\frac{d\widetilde{\sigma}}{d\mu_R} + \frac{\partial f_h}{\partial \mu_R}\right) \\ -\frac{\partial F_h}{\partial \mu_R}(\mu_R - \sigma_h(\eta^*)) + (1 - F_h) \end{pmatrix} \\
& -(1-\varphi)\frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \mu_R}, \qquad \text{(A.75)}
\end{aligned}$$

where $\frac{\partial f_k}{\partial \mu_R} = \frac{\partial f_R(\sigma_k, \mu_R)}{\partial \mu_R}$ and $\frac{\partial F_k}{\partial \mu_R} = \frac{\partial F_R(\sigma_k, \mu_R)}{\partial \mu_R}$ are the partial derivatives of, respectively, $f_R(\sigma_k, \mu_R)$ and $F_R(\sigma_k, \mu_R)$ with respect to μ_R keeping σ_k constant, for $k = l, h$. By the properties of the cdf and pdf of the normal distribution,

$$\frac{\partial F_l}{\partial \mu_R} = -f_l, \qquad \text{(A.76)}$$

$$\frac{\partial F_h}{\partial \mu_R} = -f_h, \qquad \text{(A.77)}$$

$$\frac{\partial f_l}{\partial \mu_R} = -f_l \frac{\mu_R - \sigma_l}{\gamma_R^2 + \gamma_\varepsilon^2}, \qquad \text{(A.78)}$$

$$\frac{\partial f_h}{\partial \mu_R} = -f_h \frac{\mu_R - \sigma_h}{\gamma_R^2 + \gamma_\varepsilon^2}, \qquad \text{(A.79)}$$

$$\frac{df_l}{d\sigma_l} = f_l \frac{\mu_R - \sigma_l}{\gamma_R^2 + \gamma_\varepsilon^2}. \qquad \text{(A.80)}$$

Substituting the above expressions into (A.75) yields

$$\frac{\partial \widehat{\tau}(x, \mu_R)}{\partial \mu_R}_{|x=\sigma_h(\eta^*)} = \varphi\left(F_l\left(2\frac{d\widetilde{\sigma}}{d\mu_R} - 1\right) + (1 - F_h)\right) - (1-\varphi)\frac{\partial \frac{k_2 - \Delta_0}{k_1}}{\partial \mu_R}. \qquad \text{(A.81)}$$

Substituting this into (A.74), taking the derivatives and simplifying, we finally get

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} = \begin{cases} -\frac{\gamma_S^2(1-\varphi)+\gamma_\varepsilon^2(1-\varphi(F_h+F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)} \text{ if } \gamma_S > \gamma_R \\ -\frac{\gamma_S^2(1-\varphi)+\gamma_\varepsilon^2(1-\varphi(2-F_h-F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)} \text{ if } \gamma_S < \gamma_R \end{cases}.$$

■

Lemma A.15 *Let $\mu_S \geq \mu_R = 0$ and $\gamma_S \neq \gamma_R$. Then,*

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} = sgn(\gamma_R - \gamma_S)\frac{(f_h + f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h\varphi + F_l\varphi)}$$

Proof: Since in equilibrium $\widetilde{\Delta}(\varnothing|\eta^*) = \Delta(\sigma_h)$ by Lemma A.3, we have (noting that $\frac{dk_2}{d\gamma_R} = 0$

as far as μ_R is normalized to 0)

$$\begin{aligned}
\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} &= \frac{d\Delta(\sigma_h)}{d\gamma_R} = \frac{d(k_1\sigma_h + k_2)}{d\gamma_R} \\
&= k_1\frac{d\sigma_h}{d\gamma_R} + \frac{dk_1}{d\gamma_R}\sigma_h \\
&= k_1\left(\frac{1}{1-\varphi(F_h - F_l)}\frac{\partial\widehat{\tau}(x,\gamma_R)}{\partial\gamma_R}_{|x=\sigma_h(\eta^*)}\right) + \frac{dk_1}{d\gamma_R}\sigma_h, \qquad \text{(A.82)}
\end{aligned}$$

where the second equality is by (A.20), and the last equality is by Lemma A.11.

Let us compute $\frac{\partial\widehat{\tau}(x,\gamma_R)}{\partial\gamma_R}_{|x=\sigma_h(\eta^*)}$. By definition in Lemma A.11 we have

$$\begin{aligned}
\widehat{\tau}(x,\gamma_R) &= \varphi\begin{pmatrix} F_R[2\widetilde{\sigma}-x](2\widetilde{\sigma}-x-\mu_R) \\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_R[2\widetilde{\sigma}-x]+f_R[x]) \\ +(1-F_R[x])(\mu_R - x)\end{pmatrix} \\
&\quad -(1-\varphi)\left(x + \frac{k_2-\Delta_0}{k_1}\right). \qquad \text{(A.83)}
\end{aligned}$$

Taking the partial derivative of the right-hand side with respect to γ_R at $x = \sigma_h(\eta^*)$ (treating x as constant), and subsequently substituting $2\widetilde{\sigma} - \sigma_h(\eta^*) = \widetilde{\sigma} - \eta^* = \sigma_l(\eta^*)$ yields:

$$\begin{aligned}
&\frac{\partial\widehat{\tau}(x,\gamma_R)}{\partial\gamma_R}_{|x=\sigma_h(\eta^*)} \\
&= \varphi\begin{pmatrix} \left(\frac{\partial F_l}{\partial\gamma_R} + 2f_l\frac{d\widetilde{\sigma}}{d\gamma_R}\right)(\sigma_l(\eta^*)-\mu_R) + 2F_l\frac{d\widetilde{\sigma}}{d\gamma_R} \\ +(\gamma_R^2+\gamma_\varepsilon^2)\left(\frac{\partial f_l}{\partial\gamma_R} + 2\frac{df_l}{d\sigma_l}\frac{d\widetilde{\sigma}}{d\gamma_R} + \frac{\partial f_h}{\partial\gamma_R}\right) \\ +2\gamma_R(f_l+f_h) - \frac{\partial F_h}{\partial\gamma_R}(\mu_R - \sigma_h(\eta^*))\end{pmatrix} \\
&\quad -(1-\varphi)\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\gamma_R}, \qquad \text{(A.84)}
\end{aligned}$$

where $\frac{\partial f_k}{\partial\gamma_R} = \frac{\partial f_R(\sigma_k,\gamma_R)}{\partial\gamma_R}$ and $\frac{\partial F_k}{\partial\gamma_R} = \frac{\partial F_R(\sigma_k,\gamma_R)}{\partial\gamma_R}$ are the partial derivatives of, respectively, $f_R(\sigma_k,\gamma_R)$ and $F_R(\sigma_k,\gamma_R)$ with respect to γ_R keeping σ_k constant, for $k = l,h$. By the properties of the cdf and pdf of the normal distribution (using $\sigma_l = 2\widetilde{\sigma} - \sigma_h$),

$$\begin{aligned}
\frac{\partial F_l}{\partial\gamma_R} &= -f_l\frac{\gamma_R}{\gamma_R^2+\gamma_\varepsilon^2}\sigma_l, \qquad \text{(A.85)} \\
\frac{\partial F_h}{\partial\gamma_R} &= -f_h\frac{\gamma_R}{\gamma_R^2+\gamma_\varepsilon^2}\sigma_h, \qquad \text{(A.86)} \\
\frac{\partial f_l}{\partial\gamma_R} &= -f_l\frac{\gamma_R(\gamma_R^2+\gamma_\varepsilon^2-\sigma_l^2)}{(\gamma_R^2+\gamma_\varepsilon^2)^2}, \qquad \text{(A.87)} \\
\frac{\partial f_h}{\partial\gamma_R} &= -f_h\frac{\gamma_R(\gamma_R^2+\gamma_\varepsilon^2-\sigma_h^2)}{(\gamma_R^2+\gamma_\varepsilon^2)^2}, \qquad \text{(A.88)} \\
\frac{df_l}{d\sigma_l} &= -f_l\frac{2\widetilde{\sigma}-\sigma_h}{\gamma_R^2+\gamma_\varepsilon^2}. \qquad \text{(A.89)}
\end{aligned}$$

Substituting these expressions into (A.84) yields

$$\frac{\partial\widehat{\tau}(x,\gamma_R)}{\partial\gamma_R}_{|x=\sigma_h(\eta^*)} = \varphi\left(2F_l\frac{d\widetilde{\sigma}}{d\gamma_R}+\gamma_R\left(f_l+f_h\right)\right)-(1-\varphi)\frac{\partial\frac{k_2-\Delta_0}{k_1}}{\partial\gamma_R}. \tag{A.90}$$

Next, to get an expression for σ_h, note that by Lemma A.4 in equilibrium it must hold:

$$\begin{aligned} &\tau(\eta^*)=0\\ \Leftrightarrow\quad &\varphi\left(\begin{array}{c} F_l(2\widetilde{\sigma}-\sigma_h-\mu_R)+(\gamma_R^2+\gamma_\varepsilon^2)(f_l+f_h)\\ +(1-F_h)(\mu_R-\sigma_h)\end{array}\right)-(1-\varphi)\left(\sigma_h+\frac{k_2-\Delta_0}{k_1}\right)=0\\ \Leftrightarrow\quad &\sigma_h=\frac{\varphi}{1-F_h\varphi+F_l\varphi}\left(\begin{array}{c}-\frac{1-\varphi}{\varphi}\frac{k_2-\Delta_0}{k_1}+(1-F_h)\mu_R-F_l\mu_R\\ +(\gamma_R^2+\gamma_\varepsilon^2)(f_l+f_h)+2F_l\widetilde{\sigma}\end{array}\right). \end{aligned} \tag{A.91}$$

Substituting this together with (A.90) into (A.82), taking the derivatives and simplifying, we finally get

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R}=\mathrm{sgn}(\gamma_R-\gamma_S)\frac{(f_h+f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)}. \tag{A.92}$$

■

Lemma A.16 *Let $\mu_S\geq\mu_R$ and $\gamma_S>\gamma_R$. Then, the probability of non-disclosure from S's ex ante perspective is strictly larger than the probability of non-disclosure from R's ex ante prospective.*

Proof. Denote the probabilities of non-disclosure in equilibrium from S's and R's ex ante perspectives, respectively, as $P_S^\varnothing$ and $P_R^\varnothing$. Given that disclosure happens in equilibrium if and only if $\sigma\in[\widetilde{\sigma}-\eta^*,\widetilde{\sigma}+\eta^*]$ by Proposition 1(b), we obtain

$$\begin{aligned} P_S^\varnothing &= \Pr_S[d=\varnothing]=\varphi F_S(\widetilde{\sigma}-\eta^*)+\varphi(1-F_S(\widetilde{\sigma}+\eta^*))+1-\varphi,\\ P_R^\varnothing &= \Pr_R[d=\varnothing]=\varphi F_R(\widetilde{\sigma}-\eta^*)+\varphi(1-F_R(\widetilde{\sigma}+\eta^*))+1-\varphi. \end{aligned}$$

Thus,

$$P_S^\varnothing-P_R^\varnothing=\varphi(F_S(\widetilde{\sigma}-\eta^*)-F_R(\widetilde{\sigma}-\eta^*)+F_R(\widetilde{\sigma}+\eta^*)-F_S(\widetilde{\sigma}+\eta^*))$$

Define function

$$\Delta P^\varnothing(\mu_S,\gamma_S,\mu_R,\gamma_R,s_1,s_2)=\varphi(F_S(s_1)-F_R(s_1)+F_R(s_2)-F_S(s_2)).$$

Hence, $\Delta P^\varnothing(\mu_S,\gamma_S,\mu_R,\gamma_R,\widetilde{\sigma}-\eta^*,\widetilde{\sigma}+\eta^*)=P_S^\varnothing-P_R^\varnothing$.

Let us show that $\Delta P^\varnothing(\mu_S,\gamma_S,\mu_R,\gamma_R,s_1,s_2)>0$ for any s_1 and s_2 such that

$$s_1\leq 2\mu_R-s_2<\mu_R\leq\mu_S<s_2. \tag{A.93}$$

Indeed, for any given s_1 and s_2 satisfying (A.93) we have

$$\begin{aligned} \frac{\partial\Delta P^\varnothing(\mu_S,\gamma_S,\mu_R,\gamma_R,s_1,s_2)}{\partial\mu_S} &= \varphi\left(-f_S(s_1)+f_S(s_2)\right)\\ &\geq \varphi\left(-f_S(2\mu_S-s_2)+f_S(s_2)\right)\\ &= 0, \end{aligned} \tag{A.94}$$

where the inequality follows from $s_1 \leq 2\mu_R - s_2 \leq 2\mu_S - s_2 < \mu_S$ by (A.93). Besides, for any given s_1 and s_2 satisfying (A.93) we have

$$\begin{aligned} & \frac{\partial \Delta P^{\varnothing}(\mu_S, \gamma_S, \mu_R, \gamma_R, s_1, s_2)}{\partial \gamma_S} \\ = \; & \varphi\left(f_S(s_1)\frac{\gamma_S(\mu_S - s_1)}{\gamma_S^2 + \gamma_\varepsilon^2} - f_S(s_2)\frac{\gamma_S(\mu_S - s_2)}{\gamma_S^2 + \gamma_\varepsilon^2}\right) > 0, \end{aligned} \tag{A.95}$$

where the inequality is due to $s_1 < \mu_S < s_2$ by (A.93).

Let us now fix some arbitrary priors $\widetilde{\gamma}_S > \widetilde{\gamma}_R$ and $\widetilde{\mu}_S \geq \widetilde{\mu}_R$ and any s_1 and s_2 such that

$$s_1 \leq 2\widetilde{\mu}_R - s_2 < \widetilde{\mu}_R \leq \widetilde{\mu}_S < s_2. \tag{A.96}$$

We then have

$$\begin{aligned} \Delta P^{\varnothing}(\widetilde{\mu}_S, \widetilde{\gamma}_S, \widetilde{\mu}_R, \widetilde{\gamma}_R, s_1, s_2) \; & > \; \Delta P^{\varnothing}(\widetilde{\mu}_S, \widetilde{\gamma}_R, \widetilde{\mu}_R, \widetilde{\gamma}_R, s_1, s_2) \\ & \geq \; \Delta P^{\varnothing}(\widetilde{\mu}_R, \widetilde{\gamma}_R, \widetilde{\mu}_R, \widetilde{\gamma}_R, s_1, s_2) = 0, \end{aligned} \tag{A.97}$$

where the first inequality is by (A.95) and (A.96), and the second inequality is by (A.94) and the fact that condition (A.93) holds for any $\mu_S \in [\widetilde{\mu}_R, \widetilde{\mu}_S]$ by assumption (A.96). Thus, (A.97) implies that for any given $\widetilde{\gamma}_S > \widetilde{\gamma}_R$ and $\widetilde{\mu}_S \geq \widetilde{\mu}_R$ it holds

$$\begin{aligned} \Delta P^{\varnothing}(\widetilde{\mu}_S, \widetilde{\gamma}_S, \widetilde{\mu}_R, \widetilde{\gamma}_R, s_1, s_2) \; & > \; 0 \text{ for any} \\ (s_1, s_2) \; & : \; s_1 \leq 2\widetilde{\mu}_R - s_2 < \widetilde{\mu}_R \leq \widetilde{\mu}_S < s_2. \end{aligned} \tag{A.98}$$

Finally, for any given $\widetilde{\gamma}_S > \widetilde{\gamma}_R$ and $\widetilde{\mu}_S \geq \widetilde{\mu}_R$, it holds that $s_1 = \widetilde{\sigma} - \eta^*$ and $s_2 = \widetilde{\sigma} + \eta^*$ satisfy (A.93) since

$$\widetilde{\sigma} - \eta^* \leq 2\widetilde{\mu}_R - (\widetilde{\sigma} + \eta^*) < \widetilde{\mu}_R \leq \widetilde{\mu}_S < \widetilde{\sigma} + \eta^*,$$

where the first equality is due to $\widetilde{\sigma} - \widetilde{\mu}_R = -\frac{(\widetilde{\gamma}_R^2 + \gamma_\varepsilon^2)(\widetilde{\mu}_S - \widetilde{\mu}_R)}{\widetilde{\gamma}_S^2 - \widetilde{\gamma}_R^2} \leq 0$ for $\widetilde{\mu}_S \geq \widetilde{\mu}_R$ and $\widetilde{\gamma}_S > \widetilde{\gamma}_R$, and the second and fourth equalities hold by Corollary A.1. Consequently, $P_S^{\varnothing} - P_R^{\varnothing} = \Delta P^{\varnothing}(\widetilde{\mu}_S, \widetilde{\gamma}_S, \widetilde{\mu}_R, \widetilde{\gamma}_R, \widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*) > 0$ by (A.98). ■

A.4.2 Proof of Proposition 6

In what follows, we show that for any given R's prior means denoted as μ'_R and μ''_R and fixed R's prior variance, S prefers R with the prior mean closer to μ_S.

Step 1. This proves the claim of the proposition for $\mu_S \geq \max\{\mu'_R, \mu''_R\}$. In this case, the claim is equivalent to showing that the expected perceived disagreement from S's ex ante perspective $E_S[\widetilde{\Delta}(d)]$ strictly decreases with μ_R for any $\mu_R \leq \mu_S$ (so that S prefers an R with a prior mean closer to μ_S). In turn, here we separately consider three possible parameter cases: $\gamma_S > \gamma_R$, $\gamma_S < \gamma_R$ and $\gamma_S = \gamma_R$. In what follows, the argument d in $E_S[\widetilde{\Delta}(d)]$ is suppressed for notational simplicity.

Case 1: $\gamma_S > \gamma_R$.

As before, denote by η^* the value of η corresponding to the equilibrium disclosure interval

$[\widetilde{\sigma}-\eta^*,\widetilde{\sigma}+\eta^*]$. Denote also

$$P_S^{\varnothing}=\Pr_S[d=\varnothing]=\varphi F_S(\widetilde{\sigma}-\eta^*)+\varphi(1-F_S(\widetilde{\sigma}+\eta^*))+1-\varphi$$

(the probability of non-disclosure from S' ex ante perspective). Then, we have:

$$\begin{aligned}
& E_S[\widetilde{\Delta}] \\
= \; & E_S[E_R[\Delta(\sigma)\,|d]] \\
= \; & P_S^{\varnothing}\widetilde{\Delta}(\varnothing|\eta^*)+(1-P_S^{\varnothing})\frac{1}{F_S(\widetilde{\sigma}+\eta^*)-F_S(\widetilde{\sigma}-\eta^*)}\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*}\Delta(\sigma)f_S(\sigma)d\sigma \\
= \; & (\varphi F_S(\widetilde{\sigma}-\eta^*)+\varphi(1-F_S(\widetilde{\sigma}+\eta^*))+1-\varphi)\widetilde{\Delta}(\varnothing|\eta^*) \\
& +\varphi\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*}\Delta(\sigma)f_S(\sigma)d\sigma \\
= \; & P_S^{\varnothing}\widetilde{\Delta}(\varnothing|\eta^*)+\varphi\begin{pmatrix}\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}-(k_1\sigma+k_2)f_S(\sigma)d\sigma \\ +\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}(k_1\sigma+k_2)f_S(\sigma)d\sigma\end{pmatrix},
\end{aligned} \tag{A.99}$$

where the last equality is by (A.20).

We need to show that $\frac{dE_S[\widetilde{\Delta}]}{d\mu_R}<0$ for any $\mu_R\leq\mu_S$. Taking the derivative, we obtain

$$\begin{aligned}
& \frac{dE_S[\widetilde{\Delta}]}{d\mu_R} \\
= \; & \frac{dP_S^{\varnothing}}{d\mu_R}\widetilde{\Delta}(\varnothing|\eta^*)+P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} \\
& +\varphi\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}f_S(\sigma)d\sigma-\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}f_S(\sigma)d\sigma\right) \\
& +\varphi\frac{d\widetilde{\sigma}}{d\mu_R}\Delta(\widetilde{\sigma})f_S(\widetilde{\sigma})-\varphi\left(\frac{d\widetilde{\sigma}}{d\mu_R}-\frac{d\eta^*}{d\mu_R}\right)\Delta(\widetilde{\sigma}-\eta^*)f_S(\widetilde{\sigma}-\eta^*) \\
& +\varphi\left(\frac{d\widetilde{\sigma}}{d\mu_R}+\frac{d\eta^*}{d\mu_R}\right)\Delta(\widetilde{\sigma}+\eta^*)f_S(\widetilde{\sigma}+\eta^*)-\varphi\frac{d\widetilde{\sigma}}{d\mu_R}\Delta(\widetilde{\sigma})f_S(\widetilde{\sigma}) \\
= \; & \varphi\begin{pmatrix} f_S(\widetilde{\sigma}-\eta^*)\left(\frac{d\widetilde{\sigma}}{d\mu_R}-\frac{d\eta^*}{d\mu_R}\right) \\ -f_S(\widetilde{\sigma}+\eta^*)\left(\frac{d\widetilde{\sigma}}{d\mu_R}+\frac{d\eta^*}{d\mu_R}\right)\end{pmatrix}\widetilde{\Delta}(\varnothing|\eta^*) \\
& +P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} \\
& +\varphi\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}f_S(\sigma)d\sigma-\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}f_S(\sigma)d\sigma\right) \\
& -\varphi\Delta(\widetilde{\sigma}+\eta^*)\begin{pmatrix}\left(\frac{d\widetilde{\sigma}}{d\mu_R}-\frac{d\eta^*}{d\mu_R}\right)f_S(\widetilde{\sigma}-\eta^*) \\ -\left(\frac{d\widetilde{\sigma}}{d\mu_R}+\frac{d\eta^*}{d\mu_R}\right)f_S(\widetilde{\sigma}+\eta^*)\end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= \varphi\left(\widetilde{\Delta}(\varnothing|\eta^*) - \Delta(\widetilde{\sigma}+\eta^*)\right) \\
&\quad \times \left(\begin{array}{c} \left(\frac{d\widetilde{\sigma}}{d\mu_R} - \frac{d\eta^*}{d\mu_R}\right) f_S(\widetilde{\sigma}-\eta^*) \\ -\left(\frac{d\widetilde{\sigma}}{d\mu_R} + \frac{d\eta^*}{d\mu_R}\right) f_S(\widetilde{\sigma}+\eta^*) \end{array}\right) \\
&\quad + P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} + \varphi\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} f_S(\sigma)d\sigma\right) \\
&= P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} + \varphi\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} f_S(\sigma)d\sigma\right), \qquad \text{(A.100)}
\end{aligned}
$$

where the second equality follows by $\Delta(\widetilde{\sigma}+\eta^*) = \Delta(\widetilde{\sigma}-\eta^*)$, and the last equality follows due to $\widetilde{\Delta}(\varnothing|\eta^*) - \Delta(\widetilde{\sigma}+\eta^*) = 0$ by Lemma A.3.

Next, by Lemma A.14 we have

$$
\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} = -\frac{\gamma_S^2(1-\varphi) + \gamma_\varepsilon^2(1-\varphi(F_h+F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)}. \qquad \text{(A.101)}
$$

Note that $F_h + F_l \leq 1$ since

$$
F_h = F_R(\widetilde{\sigma}+\eta^*) \leq F_R(\mu_R+\eta^*) = 1 - F_R(\mu_R-\eta^*) \leq 1 - F_R(\widetilde{\sigma}-\eta^*) = 1 - F_l,
$$

where the inequalities are due to $\widetilde{\sigma} \leq \mu_R$ in the considered parameter case by (A.29), and the second equality follows by the symmetry of f_R around μ_R. Consequently, by (A.101)

$$
\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} < 0. \qquad \text{(A.102)}
$$

Next, consider the following term on the right-hand side of (A.100):

$$
\begin{aligned}
&\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} f_S(\sigma)d\sigma\right) \\
&= -\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(2\widetilde{\sigma}-\sigma)d\sigma\right) \\
&= -\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \left(f_S(\sigma) - f_S(2\widetilde{\sigma}-\sigma)\right)d\sigma \leq 0, \qquad \text{(A.103)}
\end{aligned}
$$

where the first equality is by integration by substitution, and the inequality follows from $f_S(\sigma) \geq f_S(2\widetilde{\sigma}-\sigma)$ for any $\sigma \geq \widetilde{\sigma}$ due to $\widetilde{\sigma} \leq \mu_S$ in the current parameter case by (A.28). Combining (A.103) and (A.102) with (A.100) we obtain $\frac{dE_S[\widetilde{\Delta}]}{d\mu_R} < 0$.

Case 2: $\gamma_S < \gamma_R$.

From (A.100) we have:

$$
\frac{dE_S[\widetilde{\Delta}]}{d\mu_R} = P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} + \varphi\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} f_S(\sigma)d\sigma\right). \qquad \text{(A.104)}
$$

Next, by Lemma A.14 we have

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} = -\frac{\gamma_S^2(1-\varphi)+\gamma_\varepsilon^2(1-\varphi(2-F_h-F_l))}{(\gamma_S^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)}. \tag{A.105}$$

Note that

$$\begin{aligned} 2-F_h-F_l &\leq 1 \Leftrightarrow \\ 1-F_h &\leq F_l. \end{aligned}$$

At the same time,

$$1-F_h = 1-F_R(\widetilde{\sigma}+\eta^*) \leq 1-F_R(\mu_R+\eta^*) = F_R(\mu_R-\eta^*) \leq F_R(\widetilde{\sigma}-\eta^*) = F_l,$$

where the inequalities are due to $\widetilde{\sigma} \geq \mu_R$ in the considered parameter case by (A.29), and the second equality follows by the symmetry of f_R around μ_R. Thus, $2-F_h-F_l \leq 1$ so that by (A.105)

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\mu_R} < 0. \tag{A.106}$$

Next, consider the following term on the right-hand side of (A.104):

$$\begin{aligned} &\frac{dk_2}{d\mu_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} f_S(\sigma)d\sigma\right) \\ = &\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} f_S(2\widetilde{\sigma}-\sigma)d\sigma\right) \\ = &\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2}\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \left(f_S(\sigma)-f_S(2\widetilde{\sigma}-\sigma)\right)d\sigma \leq 0, \end{aligned} \tag{A.107}$$

where the first equality is by integration by substitution, and the inequality follows from $f_S(\sigma) \leq f_S(2\widetilde{\sigma}-\sigma)$ for any $\sigma \geq \widetilde{\sigma}$ due to $\widetilde{\sigma} \geq \mu_S$ in the current parameter case by (A.28). Combining (A.107) and (A.106) with (A.104) we obtain $\frac{dE_S[\widetilde{\Delta}]}{d\mu_R} < 0$.

Case 3: $\gamma_S = \gamma_R$.

In this case, (A.20) implies that at any signal σ the disagreement function is equal to the constant $\frac{\gamma_\varepsilon^2(\mu_S-\mu_R)}{\gamma_R^2+\gamma_\varepsilon^2}$. Moreover, by Proposition 2 full disclosure is the only equilibrium (so that the perceived disagreement conditional on no disclosure is equal to the prior disagreement $\mu_S-\mu_R$). Hence,

$$E_S[\widetilde{\Delta}|\gamma_S=\gamma_R] = (1-\varphi)(\mu_S-\mu_R)+\varphi\frac{\gamma_\varepsilon^2(\mu_S-\mu_R)}{\gamma_R^2+\gamma_\varepsilon^2} \tag{A.108}$$

and

$$\frac{dE_S[\widetilde{\Delta}|\gamma_S=\gamma_R]}{d\mu_R} = -(1-\dot{\phi})-\varphi\frac{\gamma_\varepsilon^2}{\gamma_R^2+\gamma_\varepsilon^2} < 0.$$

Step 2. Let us generalize the claim of the proposition for arbitrary $\mu_R^{'}$ and $\mu_R^{''}$ (where $\mu_R^{''}$ is closer to μ_S without loss of generality), given that the claim holds for $\mu_R^{'} < \mu_R^{''} \leq \mu_S$ by Step 1.

Consider first the case $\mu_R^{'} > \mu_R^{''} \geq \mu_S$. Denote function $M(\mu_R) = 2\mu_S - \mu_R$, so that $M(\mu_R)$ and μ_R are symmetric to each other relative to μ_S. By Lemma A.10, the equilibrium disclosure intervals under $M(\mu_R)$ and μ_R are then also symmetric to each other relative to μ_S for any μ_R. Consequently, from S's perspective the expected perceived disagreement must be equivalent between μ_R and $M(\mu_R)$ (as the values of R's prior mean), i.e.,

$$E_S[\widetilde{\Delta}|\mu_R] = E_S[\widetilde{\Delta}|M(\mu_R)] \text{ for any } \mu_R \in \mathbb{R} \tag{A.109}$$

(a more formal derivation can be obtained by similar steps as in the proof of Lemma A.9). Hence,

$$\begin{aligned} & E_S[\widetilde{\Delta}|\mu'_R] - E_S[\widetilde{\Delta}|\mu''_R] \\ = \;& E_S[\widetilde{\Delta}|M(\mu'_R)] - E_S[\widetilde{\Delta}|M(\mu''_R)] > 0, \end{aligned}$$

where the inequality follows by

$$\begin{aligned} & \mu_R^{'} > \mu_R^{''} \geq \mu_S \\ \Leftrightarrow \;& 2\mu_S - \mu'_R < 2\mu_S - \mu''_R \leq 2\mu_S - \mu_S \\ \Leftrightarrow \;& M(\mu'_R) < M(\mu''_R) \leq \mu_S \end{aligned}$$

and Step 1. Thus, S prefers μ''_R (the closer mean) over μ'_R in terms of the expected perceived disagreement.

Finally, consider the last possible case where two R's means are on the opposite sides of μ_S: $\mu''_R > \mu_S > \mu'_R$. As μ''_R is assumed to be closer to μ_S than μ'_R (without loss of generality), we have $|\mu'_R - \mu_S| > |\mu''_R - \mu_S|$. Then, by (A.109),

$$E_S[\widetilde{\Delta}|\mu'_R] - E_S[\widetilde{\Delta}|\mu''_R] = E_S[\widetilde{\Delta}|\mu'_R] - E_S[\widetilde{\Delta}|M(\mu''_R)]. \tag{A.110}$$

At the same time, by assumption

$$\begin{aligned} |\mu_R^{'} - \mu_S| &> |\mu''_R - \mu_S| \Leftrightarrow \\ \mu_S - \mu_R^{'} &> \mu''_R - \mu_S \Leftrightarrow \\ \mu'_R &< 2\mu_S - \mu''_R \Leftrightarrow \\ \mu'_R &< M(\mu''_R), \end{aligned}$$

where the first equivalence is due to $\mu''_R > \mu_S > \mu'_R$ by assumption. Note also that $\mu_S > 2\mu_S - \mu''_R = M(\mu''_R)$ due to $\mu''_R > \mu_S$ by assumption. Consequently, by $\mu_S > M(\mu''_R) > \mu'_R$ and Step 1 we obtain $E_S[\widetilde{\Delta}|\mu'_R] > E_S[\widetilde{\Delta}|M(\mu''_R)]$, which together with (A.110) implies $E_S[\widetilde{\Delta}|\mu'_R] > E_S[\widetilde{\Delta}|\mu''_R]$, i.e., S prefers μ''_R (the closer mean) over μ'_R in terms of the expected perceived disagreement. ■

A.4.3 Proof of Proposition 7

Without loss of generality, throughout the proof we assume $\mu_S \geq \mu_R$ and normalize μ_R to 0. In what follows, the argument d in $E_S[\widetilde{\Delta}(d)]$ is suppressed for notational simplicity.

Claim 1. *The expected perceived disagreement from S's ex ante perspective $E_S[\widetilde{\Delta}]$ strictly decreases in γ_R if $\gamma_R < \gamma_S$.*

Proof: By (A.99),

$$E_S[\widetilde{\Delta}] = P_S^{\varnothing}\widetilde{\Delta}(\varnothing|\eta^*) + \varphi\begin{pmatrix} \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} -(k_1\sigma + k_2)f_S(\sigma)d\sigma \\ +\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}(k_1\sigma + k_2)f_S(\sigma)d\sigma \end{pmatrix}.$$

We have (noting that $\frac{dk_2}{d\gamma_R} = 0$ since μ_R is normalized to 0):

$$\begin{aligned}
&\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} \\
= \; & \frac{dP_S^{\varnothing}}{d\gamma_R}\widetilde{\Delta}(\varnothing|\eta^*) + P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} \\
& +\varphi\frac{dk_1}{d\gamma_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right) \\
& +\varphi\frac{d\widetilde{\sigma}}{d\gamma_R}\Delta(\widetilde{\sigma})f_S(\widetilde{\sigma}) - \varphi\left(\frac{d\widetilde{\sigma}}{d\gamma_R} - \frac{d\eta^*}{d\gamma_R}\right)\Delta(\widetilde{\sigma}-\eta^*)f_S(\widetilde{\sigma}-\eta^*) \\
& +\varphi\left(\frac{d\widetilde{\sigma}}{d\gamma_R} + \frac{d\eta^*}{d\gamma_R}\right)\Delta(\widetilde{\sigma}+\eta^*)f_S(\widetilde{\sigma}+\eta^*) - \varphi\frac{d\widetilde{\sigma}}{d\gamma_R}\Delta(\widetilde{\sigma})f_S(\widetilde{\sigma}) \\
= \; & \varphi\begin{pmatrix} f_S(\widetilde{\sigma}-\eta^*)\left(\frac{d\widetilde{\sigma}}{d\gamma_R} - \frac{d\eta^*}{d\gamma_R}\right) \\ -f_S(\widetilde{\sigma}+\eta^*)\left(\frac{d\widetilde{\sigma}}{d\gamma_R} + \frac{d\eta^*}{d\gamma_R}\right) \end{pmatrix}\widetilde{\Delta}(\varnothing|\eta^*) \\
& +P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} \\
& +\varphi\frac{dk_1}{d\gamma_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right) \\
& -\varphi\Delta(\widetilde{\sigma}+\eta^*)\begin{pmatrix} \left(\frac{d\widetilde{\sigma}}{d\gamma_R} - \frac{d\eta^*}{d\gamma_R}\right)f_S(\widetilde{\sigma}-\eta^*) \\ -\left(\frac{d\widetilde{\sigma}}{d\gamma_R} + \frac{d\eta^*}{d\gamma_R}\right)f_S(\widetilde{\sigma}+\eta^*) \end{pmatrix} \\
= \; & \varphi\left(\widetilde{\Delta}(\varnothing|\eta^*) - \Delta(\widetilde{\sigma}+\eta^*)\right) \\
& \times\begin{pmatrix} \left(\frac{d\widetilde{\sigma}}{d\gamma_R} - \frac{d\eta^*}{d\gamma_R}\right)f_S(\widetilde{\sigma}-\eta^*) \\ -\left(\frac{d\widetilde{\sigma}}{d\gamma_R} + \frac{d\eta^*}{d\gamma_R}\right)f_S(\widetilde{\sigma}+\eta^*) \end{pmatrix} \\
& +P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} + \varphi\frac{dk_1}{d\gamma_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right)
\end{aligned}$$

$$= P_S^{\varnothing}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} + \varphi\frac{dk_1}{d\gamma_R}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right), \tag{A.111}$$

where the second equality follows by $\Delta(\widetilde{\sigma}+\eta^*) = \Delta(\widetilde{\sigma}-\eta^*)$, and the last equality follows due to $\widetilde{\Delta}(\varnothing|\eta^*) - \Delta(\widetilde{\sigma}+\eta^*) = 0$ by Lemma A.3.

Further, by Lemma A.15 for $\gamma_R < \gamma_S$ it holds

$$\begin{aligned}\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} &= -\frac{(f_h+f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)(1-F_h\varphi+F_l\varphi)} \\ &= -\frac{(f_h+f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)P_R^\varnothing},\end{aligned} \tag{A.112}$$

where $P_R^\varnothing = (1-F_h\varphi+F_l\varphi)$ is the ex ante probability of non-disclosure from R's perspective. Substituting it back to (A.111) and taking the derivative $\frac{dk_1}{d\gamma_R}$ we obtain

$$\begin{aligned}\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} &= -\frac{P_S^\varnothing}{P_R^\varnothing}\frac{(f_h+f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)} \\ &\quad -\varphi\frac{2\gamma_R\gamma_\varepsilon^2}{(\gamma_R^2+\gamma_\varepsilon^2)^2}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right).\end{aligned} \tag{A.113}$$

The subsequent analysis distinguishes two cases: $\widetilde{\sigma}+\eta^* \geq -\widetilde{\sigma}$ and $\widetilde{\sigma}+\eta^* < -\widetilde{\sigma}$.

Case 1: $\widetilde{\sigma}+\eta^* \geq -\widetilde{\sigma}$.

Consider the second term in (A.113). We have:

$$\begin{aligned}&\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma \\ =\ &\int_{\widetilde{\sigma}}^{-\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma + \left(\int_{-\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right).\end{aligned} \tag{A.114}$$

Since, $\widetilde{\sigma}+\eta^* \geq -\widetilde{\sigma}$ by assumption and $\widetilde{\sigma} \leq 0 \leq -\widetilde{\sigma}$ in the considered parameter case, the term in brackets in (A.114) is strictly positive:

$$\int_{-\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma > 0. \tag{A.115}$$

Next, using integration by substitution we obtain:

$$\begin{aligned}\int_{\widetilde{\sigma}}^{-\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma &= \int_{\widetilde{\sigma}}^{0}\sigma f_S(\sigma)d\sigma + \int_{0}^{-\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma \\ &= \int_{0}^{-\widetilde{\sigma}}(-x)f_S(-x)dx + \int_{0}^{-\widetilde{\sigma}}x f_S(x)dx \\ &= \int_{0}^{-\widetilde{\sigma}}x(f_S(x)-f_S(-x))dx \geq 0,\end{aligned} \tag{A.116}$$

where the inequality follows from the fact that $\mu_S \geq \mu_R = 0$ so that $f_S(x) - f_S(-x) \geq 0$ for any $x \geq 0$ (recall that $0 \leq -\widetilde{\sigma}$). (A.116) and (A.115) finally imply that the whole right-hand side of (A.114) is strictly positive, which together with (A.113) yields $\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} < 0$ for $\gamma_R < \gamma_S$. Thus, Claim 1 is proven for the case $\widetilde{\sigma}+\eta^* \geq -\widetilde{\sigma}$.

Case 2: $\widetilde{\sigma}+\eta^* < -\widetilde{\sigma}$.

The proof of Claim 1 for this case proceeds in three steps.

Step 1. First, let us decompose $\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R}$ in two components, denoted further as A and B, as follows. From (A.113) we have

$$\begin{aligned}\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} &= -\frac{P_S^{\varnothing}}{P_R^{\varnothing}}\frac{(f_h+f_l)\gamma_R\gamma_\varepsilon^2(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2+\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)} \\ &\quad -\varphi\frac{2\gamma_R\gamma_\varepsilon^2}{(\gamma_R^2+\gamma_\varepsilon^2)^2}\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma-\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right). \end{aligned} \tag{A.117}$$

Note that

$$\begin{aligned} f_h &= \frac{1-F_h}{\gamma_R^2+\gamma_\varepsilon^2}(\gamma_R^2+\gamma_\varepsilon^2)\frac{f_h}{1-F_h} \\ &= \frac{1-F_h}{\gamma_R^2+\gamma_\varepsilon^2}E_R[\sigma|\sigma\geq\widetilde{\sigma}+\eta^*] \\ &= \frac{1}{\gamma_R^2+\gamma_\varepsilon^2}\int_{\widetilde{\sigma}+\eta^*}^{\infty}\sigma f_R(\sigma)d\sigma \\ &= \frac{1}{\gamma_R^2+\gamma_\varepsilon^2}\left(\int_{-\infty}^{\infty}\sigma f_R(\sigma)d\sigma-\int_{-\infty}^{\widetilde{\sigma}+\eta^*}\sigma f_R(\sigma)d\sigma\right) \\ &= -\frac{1}{\gamma_R^2+\gamma_\varepsilon^2}\int_{-\infty}^{\widetilde{\sigma}+\eta^*}\sigma f_R(\sigma)d\sigma, \end{aligned}$$

where the second equality is by (A.14), and the last equality follows from $\int_{-\infty}^{\infty}\sigma f_R(\sigma)d\sigma = \mu_R = 0$. After substituting this into (A.117) and rearranging we obtain:

$$\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = \varphi\frac{\gamma_R\gamma_\varepsilon^2}{(\gamma_R^2+\gamma_\varepsilon^2)^2}\left(\begin{array}{c} -\frac{P_S^{\varnothing}}{P_R^{\varnothing}}\frac{f_l(\gamma_R^2+\gamma_\varepsilon^2)(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)}{\gamma_S^2+\gamma_\varepsilon^2} \\ +\frac{P_S^{\varnothing}}{P_R^{\varnothing}}\frac{\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2}{\gamma_S^2+\gamma_\varepsilon^2}\int_{-\infty}^{\widetilde{\sigma}+\eta^*}\sigma f_R(\sigma)d\sigma \\ -2\left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}\sigma f_S(\sigma)d\sigma-\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}}\sigma f_S(\sigma)d\sigma\right)\end{array}\right).$$

In turn, the right-hand side can now be decomposed as follows:

$$\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = \varphi\frac{\gamma_R\gamma_\varepsilon^2}{(\gamma_R^2+\gamma_\varepsilon^2)^2}(A+B) \tag{A.118}$$

where

$$\begin{aligned}
A &= -\frac{P_S^{\varnothing}}{P_R^{\varnothing}} \frac{f_l(\gamma_R^2+\gamma_\varepsilon^2)(\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2)}{\gamma_S^2+\gamma_\varepsilon^2} \\
&\quad +\frac{P_S^{\varnothing}}{P_R^{\varnothing}} \frac{\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2}{\gamma_S^2+\gamma_\varepsilon^2} \int_{-\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*} \sigma f_R(\sigma) d\sigma \\
&\quad -2\int_{-\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma + 2\int_{\widetilde{\sigma}-\eta^*}^{3\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma, \qquad \text{(A.119)}\\
B &= \frac{P_S^{\varnothing}}{P_R^{\varnothing}} \frac{\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2}{\gamma_S^2+\gamma_\varepsilon^2} \int_{-\infty}^{-\widetilde{\sigma}-\eta^*} \sigma f_R(\sigma) d\sigma \\
&\quad -2\int_{\widetilde{\sigma}}^{-\widetilde{\sigma}-\eta^*} \sigma f_S(\sigma) d\sigma + 2\int_{3\widetilde{\sigma}+\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \qquad \text{(A.120)}
\end{aligned}$$

In Steps 2 and 3, we show that both A and B are strictly negative, respectively.

Step 2. Let us show that $A < 0$. The first term is clearly strictly negative. The second term is 0, since $\int_{-\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*} \sigma f_R(\sigma) d\sigma = 0$ (as $f_R(\sigma)$ is symmetric around $\mu_R = 0$).

Consider the third term of A. Using integration by substitution we obtain

$$\begin{aligned}
\int_{-\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma &= \int_{-\widetilde{\sigma}-\eta^*}^{0} \sigma f_S(\sigma) d\sigma + \int_{0}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma \\
&= \int_{0}^{\widetilde{\sigma}+\eta^*} (-x) f_S(-x) dx + \int_{0}^{\widetilde{\sigma}+\eta^*} x f_S(x) dx \\
&= \int_{0}^{\widetilde{\sigma}+\eta^*} x(f_S(x) - f_S(-x)) dx \geq 0, \qquad \text{(A.121)}
\end{aligned}$$

where the inequality follows from the fact that $\mu_S \geq \mu_R = 0$ so that $f_S(x) - f_S(-x) \geq 0$ for any $x \geq 0$. This implies that the third term of A, $-2\int_{-\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma$, is negative.

Finally, consider the last term of A, $2\int_{\widetilde{\sigma}-\eta^*}^{3\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma$. We have

$$\widetilde{\sigma} - \eta^* < 3\widetilde{\sigma} + \eta^* < 0, \qquad \text{(A.122)}$$

where the first inequality is by $3\widetilde{\sigma} + \eta^* - (\widetilde{\sigma} - \eta^*) = 2(\widetilde{\sigma} + \eta^*) > 0$, which in turn holds since $\widetilde{\sigma} + \eta^* > \mu_R = 0$ by Corollary A.1 The second inequality is by $3\widetilde{\sigma} + \eta^* = \widetilde{\sigma} + (2\widetilde{\sigma} + \eta^*) < 0$, which in turn holds since $\widetilde{\sigma} = \frac{\mu_S(\gamma_R^2+\gamma_\varepsilon^2)}{\gamma_R^2-\gamma_S^2} \leq 0$ (as $\gamma_R < \gamma_S$ by assumption) while $2\widetilde{\sigma} + \eta^* < 0 \Leftrightarrow \widetilde{\sigma} + \eta^* < -\widetilde{\sigma}$ by assumption of Case 2. (A.122) immediately implies that the last term of A, $2\int_{\widetilde{\sigma}-\eta^*}^{3\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma$, is strictly negative.

Thus, all terms of A are negative or 0 (with at least the first and the last terms strictly negative), which implies $A < 0$.

Step 3. Let us show that B defined in (A.120) is strictly negative.
Recall

$$\begin{aligned}
B &= \frac{P_S^{\varnothing}}{P_R^{\varnothing}} \frac{\gamma_S^2+\gamma_R^2+2\gamma_\varepsilon^2}{\gamma_S^2+\gamma_\varepsilon^2} \int_{-\infty}^{-\widetilde{\sigma}-\eta^*} \sigma f_R(\sigma) d\sigma \\
&\quad -2\int_{\widetilde{\sigma}}^{-\widetilde{\sigma}-\eta^*} \sigma f_S(\sigma) d\sigma + 2\int_{3\widetilde{\sigma}+\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma.
\end{aligned}$$

Consider the first term. Note that $\frac{P_S^{\varnothing}}{P_R^{\varnothing}} > 1$ by Lemma A.16, while $\int_{-\infty}^{-\widetilde{\sigma}-\eta^*} \sigma f_R(\sigma) d\sigma < 0$ due to $-\widetilde{\sigma} - \eta^* = -(\widetilde{\sigma} + \eta^*) < \mu_R = 0$ by Lemma A.1. This implies

$$
\begin{aligned}
B &< \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \int_{-\infty}^{-\widetilde{\sigma}-\eta^*} \sigma f_R(\sigma) d\sigma \\
&\quad -2 \int_{\widetilde{\sigma}}^{-\widetilde{\sigma}-\eta^*} \sigma f_S(\sigma) d\sigma + 2 \int_{3\widetilde{\sigma}+\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma. \qquad (A.123)
\end{aligned}
$$

Next, define function

$$
\begin{aligned}
\Upsilon(z) &= \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \int_{-\infty}^{z} \sigma f_R(\sigma) d\sigma \\
&\quad -2 \int_{\widetilde{\sigma}}^{z} \sigma f_S(\sigma) d\sigma + 2 \int_{2\widetilde{\sigma}-z}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma. \qquad (A.124)
\end{aligned}
$$

Then, (A.123) can be rewritten as

$$
B < \Upsilon(-\widetilde{\sigma} - \eta^*). \qquad (A.125)
$$

Below, we show that $\Upsilon(z) < 0$ for any $z \in (\widetilde{\sigma}, 0)$ that will be sufficient to show that $B < 0$ in the considered case ($\widetilde{\sigma} + \eta^* < -\widetilde{\sigma} \Leftrightarrow -\widetilde{\sigma} - \eta^* > \widetilde{\sigma}$).

Rearranging, we obtain

$$
\begin{aligned}
\Upsilon(z) &= \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \int_{-\infty}^{z} \sigma f_R(\sigma) d\sigma - 2 \int_{\widetilde{\sigma}}^{z} \sigma f_S(\sigma) d\sigma + 2 \int_{2\widetilde{\sigma}-z}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \\
&= \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \int_{-\infty}^{z} \sigma f_R(\sigma) d\sigma - 2 \left(\int_{-\infty}^{z} \sigma f_S(\sigma) d\sigma - \int_{-\infty}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right) \\
&\quad +2 \left(\int_{-\infty}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma - \int_{-\infty}^{2\widetilde{\sigma}-z} \sigma f_S(\sigma) d\sigma \right) \\
&= \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \left(F_R(z) E_R[\sigma | \sigma \le z] \right) \\
&\quad -2 \left(F_S(z) E_S[\sigma | \sigma \le z] - F_S(\widetilde{\sigma}) E_S[\sigma | \sigma \le \widetilde{\sigma}] \right) \\
&\quad +2 \left(F_S(\widetilde{\sigma}) E_S[\sigma | \sigma \le \widetilde{\sigma}] - F_S(2\widetilde{\sigma} - z) E_S[\sigma | \sigma \le 2\widetilde{\sigma} - z] \right) \\
&= \frac{\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2} \left(-(\gamma_R^2 + \gamma_\varepsilon^2) f_R(z) \right) \\
&\quad -2 \left(\mu_S F_S(z) - (\gamma_S^2 + \gamma_\varepsilon^2) f_S(z) - \mu_S F_S(\widetilde{\sigma}) + (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma}) \right) \\
&\quad +2 \left(\mu_S F_S(\widetilde{\sigma}) - (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma}) - \mu_S F_S(2\widetilde{\sigma} - z) + (\gamma_S^2 + \gamma_\varepsilon^2) f_S(2\widetilde{\sigma} - z) \right),
\end{aligned}
$$

where the last equality follows from Lemma A.1. Substituting explicit expressions for F_R, F_S, f_R and f_S and simplifying we obtain

$$\Upsilon(z) = \frac{1}{\sqrt{2\pi}\widehat{\gamma}_S^2}\left(\begin{array}{c} 2\widehat{\gamma}_S^3 e^{-\frac{(z-\mu_S)^2}{2\widehat{\gamma}_S^2}} - 4\widehat{\gamma}_S^3 e^{-\frac{\widehat{\gamma}_S^2\mu_S^2}{2(\widehat{\gamma}_S^2-\widehat{\gamma}_R^2)^2}} \\ +2\widehat{\gamma}_S^3 e^{-\frac{\left(z+\frac{(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2)\mu_S}{\widehat{\gamma}_S^2-\widehat{\gamma}_R^2}\right)^2}{2\widehat{\gamma}_S^2}} - e^{-\frac{z^2}{2\widehat{\gamma}_R^2}}\widehat{\gamma}_R\left(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2\right)\end{array}\right)$$
$$-\mu_S\left(\begin{array}{c} -2\,\mathrm{Erfc}\left[\frac{\widehat{\gamma}_S\mu_S}{\sqrt{2}(\widehat{\gamma}_S^2-\widehat{\gamma}_R^2)}\right] + \mathrm{Erfc}\left[\frac{-z+\mu_S}{\sqrt{2}\widehat{\gamma}_S}\right] \\ +\,\mathrm{Erfc}\left[\frac{z+\frac{(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2)\mu_S}{\widehat{\gamma}_S^2-\widehat{\gamma}_R^2}}{\sqrt{2}\widehat{\gamma}_S}\right]\end{array}\right), \tag{A.126}$$

where $\widehat{\gamma}_S = \sqrt{\gamma_S^2+\gamma_\varepsilon^2}$ and $\widehat{\gamma}_R = \sqrt{\gamma_R^2+\gamma_\varepsilon^2}$. Taking the derivative wrt z and rearranging we obtain

$$\frac{\partial\Upsilon(z)}{\partial z} = \frac{1}{\sqrt{2\pi}\widehat{\gamma}_S^2}\left(\frac{2e^{-\frac{\left(z+\frac{(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2)\mu_S}{\widehat{\gamma}_S^2-\widehat{\gamma}_R^2}\right)^2}{2\widehat{\gamma}_S^2}}\widehat{\gamma}_S}{\widehat{\gamma}_S^2-\widehat{\gamma}_R^2}\lambda_1(z) + ze^{-\frac{(z-\mu_S)^2}{2\widehat{\gamma}_S^2}}\lambda_2(z)\right), \tag{A.127}$$

where

$$\begin{aligned}\lambda_1(z) &= z(\widehat{\gamma}_R^2-\widehat{\gamma}_S^2) - 2\widehat{\gamma}_R^2\mu_S, \\ \lambda_2(z) &= \frac{e^{\frac{1}{2}\left(-\frac{z^2}{\widehat{\gamma}_R^2}+\frac{(z-\mu_S)^2}{\widehat{\gamma}_S^2}\right)}(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2)}{\widehat{\gamma}_R} - 2\widehat{\gamma}_S.\end{aligned}$$

Let us show that $\lambda_1(z) < 0$ and $\lambda_2(z) > 0$ for $z \in (\widetilde{\sigma}, 0)$. Consider $\lambda_1(z)$. Since $\gamma_S > \gamma_R$ by assumption, it is decreasing in z. Hence, $\lambda_1(z)$ reaches its maximum on $(\widetilde{\sigma}, 0)$ at $z = \widetilde{\sigma}$, i.e.,

$$\begin{aligned}\forall z &\in (\widetilde{\sigma},0) : \lambda_1(z) < \lambda_1(\widetilde{\sigma}) = \frac{\mu_S\widehat{\gamma}_R^2}{\widehat{\gamma}_R^2-\widehat{\gamma}_S^2}(\widehat{\gamma}_R^2-\widehat{\gamma}_S^2) - 2\widehat{\gamma}_R^2\mu_S \\ &= -\widehat{\gamma}_R^2\mu_S \le -\widehat{\gamma}_R^2\mu_R = 0.\end{aligned} \tag{A.128}$$

It follows that $\lambda_1(z) < 0$ for any $z \in (\widetilde{\sigma}, 0)$.

Consider $\lambda_2(z)$. We have

$$\frac{\partial\lambda_2(z)}{\partial z} = \frac{e^{\frac{1}{2}\left(-\frac{z^2}{\widehat{\gamma}_R^2}+\frac{(z-\mu_S)^2}{\widehat{\gamma}_S^2}\right)}(\widehat{\gamma}_S^2+\widehat{\gamma}_R^2)(z(\widehat{\gamma}_R^2-\widehat{\gamma}_S^2)-2\widehat{\gamma}_R^2\mu_S)}{\widehat{\gamma}_S^2\widehat{\gamma}_R^3} < 0,$$

since $z(\widehat{\gamma}_R^2-\widehat{\gamma}_S^2) - 2\widehat{\gamma}_R^2\mu_S = \lambda_1(z) < 0$ for any $z \in (\widetilde{\sigma}, 0)$ as shown above. Since $\lambda_2(z)$ is decreasing in z, it reaches its minimum on $(\widetilde{\sigma}, 0)$ at $z = 0$, i.e.,

$$\begin{aligned}\forall z &\in (\widetilde{\sigma}, 0) : \lambda_2(z) > \lambda_2(0) = \frac{e^{\frac{1}{2}\frac{\mu_S^2}{\widehat{\gamma}_S^2}} \left(\widehat{\gamma}_S^2 + \widehat{\gamma}_R^2\right)}{\widehat{\gamma}_R} - 2\widehat{\gamma}_S \\ &\geq \frac{\widehat{\gamma}_S^2 + \widehat{\gamma}_R^2}{\widehat{\gamma}_R} - 2\widehat{\gamma}_S = \frac{\left(\widehat{\gamma}_S - \widehat{\gamma}_R\right)^2}{\widehat{\gamma}_R} > 0.\end{aligned}$$

It follows that $\lambda_2(z) > 0$ for $z \in (\widetilde{\sigma}, 0)$. Consequently, the term $ze^{-\frac{(z-\mu_S)^2}{2\widehat{\gamma}_S^2}} \lambda_2(z) < 0$ for $z \in (\widetilde{\sigma}, 0)$. This together with $\lambda_1(z) < 0$ by (A.128) implies by (A.127) that $\frac{\partial \Upsilon(z)}{\partial z} < 0$. Hence, $\Upsilon(z)$ reaches its maximum on $z \in (\widetilde{\sigma}, 0)$ at $z = \widetilde{\sigma}$, i.e.,

$$\forall z \in (\widetilde{\sigma}, 0) : \Upsilon(z) < \Upsilon(\widetilde{\sigma}). \tag{A.129}$$

Substituting $\widetilde{\sigma} = \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \frac{\mu_S \widehat{\gamma}_R^2}{\widehat{\gamma}_R^2 - \widehat{\gamma}_S^2}$ into (A.126) and simplifying we obtain

$$\Upsilon(\widetilde{\sigma}) = -\frac{e^{-\frac{\widehat{\gamma}_R^2 \mu_S^2}{2\left(\widehat{\gamma}_S^2 - \widehat{\gamma}_R^2\right)^2}} \widehat{\gamma}_R \left(\widehat{\gamma}_S^2 + \widehat{\gamma}_R^2\right)}{\sqrt{2\pi}\widehat{\gamma}_S^2} < 0. \tag{A.130}$$

This together with (A.129) implies

$$\forall z \in (\widetilde{\sigma}, 0) : \Upsilon(z) < 0. \tag{A.131}$$

Now, recall that we consider the case $\widetilde{\sigma} + \eta^* < -\widetilde{\sigma}$, which together with $\widetilde{\sigma} + \eta^* > \mu_R = 0$ (by Corollary A.1) implies $-\widetilde{\sigma} - \eta^* \in (\widetilde{\sigma}, 0)$. This together with (A.131) yields $\Upsilon(-\widetilde{\sigma} - \eta^*) < 0$. This together with (A.125) finally leads to $B < 0$.

Overall, we have shown that both A and B from the right-hand side of (A.118) are strictly negative. Consequently, in the considered case of $\widetilde{\sigma} + \eta^* < -\widetilde{\sigma}$, we obtain $\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} < 0$. As this has also been shown for the other possible case $\widetilde{\sigma} + \eta^* \geq -\widetilde{\sigma}$, the proof of Claim 1 is complete.

Claim 2. *The expected perceived disagreement from S's ex ante perspective $E_S[\widetilde{\Delta}]$ is continuous in γ_R at any $\gamma_R \in (0, \infty)$.*

Proof. We show the continuity of $E_S[\widetilde{\Delta}]$ at a given value of γ_R in two separate cases: $\gamma_R \neq \gamma_S$ and $\gamma_R = \gamma_S$.

Case 1. $\gamma_R \neq \gamma_S$.

Recall that by (A.99) for $\gamma_R \neq \gamma_S$ we have

$$E_S[\widetilde{\Delta}] = (\varphi F_S(\sigma_l) + \varphi(1 - F_S(\sigma_h)) + 1 - \varphi)\, \widetilde{\Delta}(\varnothing|\eta^*) + \varphi \begin{pmatrix} \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} -(k_1\sigma + k_2) f_S(\sigma) d\sigma \\ + \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} (k_1\sigma + k_2) f_S(\sigma) d\sigma \end{pmatrix}. \tag{A.132}$$

Note that if $\gamma_R \neq \gamma_S$, then $\widetilde{\Delta}(\varnothing|\eta^*)$ is differentiable and hence continuous in γ_R by Lemma A.15. $\widetilde{\sigma}$, k_1 and k_2 are also continuous in γ_R as far as $\gamma_R \neq \gamma_S$ by (6), (A.21) and (A.22), respectively. Finally, since in equilibrium $\widetilde{\Delta}(\varnothing|\eta^*) = \Delta(\sigma_h) = k_1\sigma_h + k_2$ by Lemma A.3 and (A.20), we have

$$\sigma_h = \frac{\widetilde{\Delta}(\varnothing|\eta^*) - k_2}{k_1}. \tag{A.133}$$

Since $k_1 \neq 0$ and $\widetilde{\Delta}(\varnothing|\eta^*)$, k_1 and k_2 are all continuous in γ_R as far as $\gamma_R \neq \gamma_S$ as mentioned above, it follows from (A.133) that both σ_h and $\sigma_l = 2\widetilde{\sigma} - \sigma_h$ are also continuous in γ_R if $\gamma_R \neq \gamma_S$.

Thus, $\widetilde{\Delta}(\varnothing|\eta^*)$, $\widetilde{\sigma}$, k_1, k_2, σ_l and σ_h are all continuous in γ_R at any $\gamma_R \neq \gamma_S$. Consequently, the same is true for the right-hand side of (A.132) and thus $E_S[\widetilde{\Delta}]$.

Case 2. $\gamma_R = \gamma_S$.

We need to show that $E_S[\widetilde{\Delta}]$ is continuous in γ_R at $\gamma_R = \gamma_S$, i.e.,

$$\lim_{\gamma_R \to \gamma_S^-} E_S[\widetilde{\Delta}|\gamma_R] = \lim_{\gamma_R \to \gamma_S^+} E_S[\widetilde{\Delta}|\gamma_R] = E_S[\widetilde{\Delta}|\gamma_R = \gamma_S].$$

The proof proceeds by two steps corresponding to the cases $\mu_S > \mu_R$ and $\mu_S = \mu_R$ (recall that assumption $\mu_S \geq \mu_R$ is without loss of generality).

Step 1. In this step, we prove Claim 2 for the case $\mu_S > \mu_R$.

Claim 2.1: If $\mu_S > \mu_R$, *then* $\lim_{\gamma_R \to \gamma_S^\pm} \sigma_l = -\infty$, $\lim_{\gamma_R \to \gamma_S^\pm} \sigma_h = \infty$, $\lim_{\gamma_R \to \gamma_S^-} \widetilde{\sigma} = -\infty$, $\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} = \infty$, $\lim_{\gamma_R \to \gamma_S^\pm} \eta^* = \infty$.

Proof. By definition of the upper status quo signal $\overline{\sigma}$ and (A.20), we have (recall that μ_R is normalized to 0):

$$\begin{aligned}
\Delta(\overline{\sigma}) &= \Delta_0 \Leftrightarrow \\
k_1\overline{\sigma} + k_2 &= \Delta_0 \Leftrightarrow \\
\overline{\sigma} &= \frac{\Delta_0 - k_2}{k_1} = \begin{cases} \frac{\gamma_S^2(\gamma_R^2+\gamma_\varepsilon^2)\mu_S}{(\gamma_S^2-\gamma_R^2)\gamma_\varepsilon^2} & \text{if } \gamma_S > \gamma_R \\ \frac{(\gamma_S^2+2\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)\mu_S}{(\gamma_R^2-\gamma_S^2)\gamma_\varepsilon^2} & \text{if } \gamma_S < \gamma_R \end{cases},
\end{aligned} \tag{A.134}$$

so that

$$\lim_{\gamma_R \to \gamma_S^\pm} \overline{\sigma} = \infty \tag{A.135}$$

as $\mu_S > \mu_R = 0$ by assumption.

Analogously, for the lower status quo signal we have

$$\begin{aligned}
\Delta(\underline{\sigma}) &= \Delta_0 \Leftrightarrow \\
-k_1\underline{\sigma} - k_2 &= \Delta_0 \Leftrightarrow \\
\underline{\sigma} &= \frac{\Delta_0 + k_2}{-k_1} = \begin{cases} \frac{(\gamma_S^2+2\gamma_\varepsilon^2)(\gamma_R^2+\gamma_\varepsilon^2)\mu_S}{(\gamma_R^2-\gamma_S^2)\gamma_\varepsilon^2} & \text{if } \gamma_S > \gamma_R \\ \frac{\gamma_S^2(\gamma_R^2+\gamma_\varepsilon^2)\mu_S}{(\gamma_S^2-\gamma_R^2)\gamma_\varepsilon^2} & \text{if } \gamma_S < \gamma_R \end{cases},
\end{aligned} \tag{A.136}$$

so that

$$\lim_{\gamma_R \to \gamma_S^\pm} \underline{\sigma} = -\infty \tag{A.137}$$

as $\mu_S > \mu_R = 0$ by assumption. Further, by (6)

$$\lim_{\gamma_R \to \gamma_S^-} \widetilde{\sigma} = \lim_{\gamma_R \to \gamma_S^-} \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = -\infty, \tag{A.138}$$

$$\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} = \lim_{\gamma_R \to \gamma_S^+} \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \infty. \tag{A.139}$$

Since the status quo signals lie within the equilibrium disclosure interval $[\sigma_l, \sigma_h]$ by Proposition 1(b), (A.135) and (A.137) yield $\lim_{\gamma_R \to \gamma_S^\pm} \sigma_h = \infty$ and $\lim_{\gamma_R \to \gamma_S^\pm} \sigma_l = -\infty$. At the same time, since $\widetilde{\sigma}$ converges to either positive or negative infinity as shown above, it must hold that $\lim_{\gamma_R \to \gamma_S^\pm} \eta^* = \infty$ (as otherwise either $\sigma_l = \widetilde{\sigma} - \eta^* \to -\infty$ or $\sigma_h = \widetilde{\sigma} + \eta^* \to \infty$ would yield a contradiction).

Claim 2.2: If $\mu_S > \mu_R$, then for $i = S, R$: $\lim_{\gamma_R \to \gamma_S^\pm} f_i(\widetilde{\sigma} - \eta^*) = \lim_{\gamma_R \to \gamma_S^\pm} f_i(\widetilde{\sigma} + \eta^*) = \lim_{\gamma_R \to \gamma_S^\pm} F_i(\widetilde{\sigma} - \eta^*) = \lim_{\gamma_R \to \gamma_S^\pm}(1 - F_i(\widetilde{\sigma} + \eta^*)) = 0, \lim_{\gamma_R \to \gamma_S^-} F_i(\widetilde{\sigma}) = 0, \lim_{\gamma_R \to \gamma_S^+} F_i(\widetilde{\sigma}) = 1.$

Proof. Follows immediately from Claim 2.1.

Claim 2.3: If $\mu_S > \mu_R$, then $\lim_{\gamma_R \to \gamma_S^\pm} \widetilde{\Delta}(\varnothing|\eta^*) = \Delta_0$.

Proof. By (A.23) and Lemma A.1 (recall that μ_R is normalized to 0),

$$\begin{aligned} & \widetilde{\Delta}(\varnothing|\eta) \\ = \; & \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi} \\ & \times \begin{pmatrix} \varphi k_1(\gamma_R^2 + \gamma_\varepsilon^2) f_l - \varphi F_l k_2 + \varphi k_1(\gamma_R^2 + \gamma_\varepsilon^2) f_h \\ + \varphi(1 - F_h) k_2 + (1 - \varphi)\Delta_0 \end{pmatrix}. \end{aligned} \tag{A.140}$$

Then, by Claim 2.2 we obtain (given that $\lim_{\gamma_R \to \gamma_S^\pm} k_1 = 0$ and $\lim_{\gamma_R \to \gamma_S^\pm} k_2 =$sgn $(\gamma_S - \gamma_R)\frac{\gamma_\varepsilon^2}{\gamma_S^2 + \gamma_\varepsilon^2}\Delta_0$, i.e., these limits are bounded):

$$\lim_{\gamma_R \to \gamma_S^\pm} \widetilde{\Delta}(\varnothing|\eta^*) = \Delta_0.$$

Claim 2.4. If $\mu_S > \mu_R$, then $\lim_{\gamma_R \to \gamma_S^\pm} E_S[\widetilde{\Delta}|\gamma_R] = E_S[\widetilde{\Delta}|\gamma_R = \gamma_S]$.

Proof. By (A.99), for $\gamma_R \neq \gamma_S$ we have

$$\begin{aligned}
E_S[\widetilde{\Delta}] &= P_S^{\varnothing}\widetilde{\Delta}(\varnothing|\eta^*) + \varphi\begin{pmatrix} \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} -(k_1\sigma + k_2) f_S(\sigma)d\sigma \\ + \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*}(k_1\sigma + k_2) f_S(\sigma)d\sigma \end{pmatrix} \\
&= P_S^{\varnothing}\widetilde{\Delta}(\varnothing|\eta^*) \\
&\quad +\varphi k_1 \begin{pmatrix} \int_{-\infty}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma)d\sigma - \int_{-\infty}^{\widetilde{\sigma}} \sigma f_S(\sigma)d\sigma \\ -\left(\int_{-\infty}^{\widetilde{\sigma}} \sigma f_S(\sigma)d\sigma - \int_{-\infty}^{\widetilde{\sigma}-\eta^*} \sigma f_S(\sigma)d\sigma\right) \end{pmatrix} \\
&\quad +\varphi k_2 \begin{pmatrix} \int_{-\infty}^{\widetilde{\sigma}+\eta^*} f_S(\sigma)d\sigma - \int_{-\infty}^{\widetilde{\sigma}} f_S(\sigma)d\sigma \\ -\left(\int_{-\infty}^{\widetilde{\sigma}} f_S(\sigma)d\sigma - \int_{-\infty}^{\widetilde{\sigma}-\eta^*} f_S(\sigma)d\sigma\right) \end{pmatrix} \\
&= (\varphi F_S(\widetilde{\sigma}-\eta^*) + \varphi(1 - F_S(\widetilde{\sigma}+\eta^*) + 1 - \varphi)\widetilde{\Delta}(\varnothing|\eta^*) \\
&\quad +\varphi k_1 \begin{pmatrix} F_S(\widetilde{\sigma}+\eta^*)E_S[\sigma|\sigma \leq \widetilde{\sigma}+\eta^*] - 2F_S(\widetilde{\sigma})E_S[\sigma|\sigma \leq \widetilde{\sigma}] \\ +F_S(\widetilde{\sigma}-\eta^*)E_S[\sigma|\sigma \leq \widetilde{\sigma}-\eta^*] \end{pmatrix} \\
&\quad +\varphi k_2 \left(F_S(\widetilde{\sigma}+\eta^*) - 2F_S(\widetilde{\sigma}) + F_S(\widetilde{\sigma}-\eta^*)\right) \\
&= (\varphi F_S(\widetilde{\sigma}-\eta^*) + \varphi(1 - F_S(\widetilde{\sigma}+\eta^*) + 1 - \varphi)\widetilde{\Delta}(\varnothing|\eta^*) \\
&\quad +\varphi k_1 \begin{pmatrix} F_S(\widetilde{\sigma}+\eta^*)\mu_S - (\gamma_S^2+\gamma_\varepsilon^2) f_S(\widetilde{\sigma}+\eta^*) \\ -2F_S(\widetilde{\sigma})\mu_S + 2(\gamma_S^2+\gamma_\varepsilon^2) f_S(\widetilde{\sigma}) \\ +F_S(\widetilde{\sigma}-\eta^*)\mu_S - (\gamma_S^2+\gamma_\varepsilon^2) f_S(\widetilde{\sigma}-\eta^*) \end{pmatrix} \\
&\quad +\varphi k_2 \left(F_S(\widetilde{\sigma}+\eta^*) - 2F_S(\widetilde{\sigma}) + F_S(\widetilde{\sigma}-\eta^*)\right), \qquad \text{(A.141)}
\end{aligned}$$

where the last equality follows by Lemma A.1. Then, taking the limit and using Claims 2.2 and 2.3 we obtain (given that $\lim_{\gamma_R\to\gamma_S^\pm} k_1 = 0$ and $\lim_{\gamma_R\to\gamma_S^\pm} k_2 = \text{sgn}\,(\gamma_S - \gamma_R)\frac{\gamma_\varepsilon^2}{\gamma_S^2+\gamma_\varepsilon^2}\Delta_0$):

$$\begin{aligned}
\lim_{\gamma_R\to\gamma_S^\pm} E_S[\widetilde{\Delta}] &= (1-\varphi)\Delta_0 + \varphi\frac{\gamma_\varepsilon^2\mu_S}{\gamma_S^2+\gamma_\varepsilon^2} \\
&= E_S[\widetilde{\Delta}|\gamma_R = \gamma_S],
\end{aligned}$$

where the last equality is by (A.108) (recall that μ_R is normalized to 0).

Claim 2.4 implies that for $\mu_S > \mu_R$, $E_S[\widetilde{\Delta}|\gamma_R]$ is continuous in γ_R at $\gamma_R = \gamma_S$.

Step 2. In this step, we prove the continuity of $E_S[\widetilde{\Delta}|\gamma_R]$ in γ_R at $\gamma_R = \gamma_S$ if $\mu_S = \mu_R = 0$.

Claim 2.5. If $\mu_S = \mu_R = 0$, *then* $\lim_{\gamma_R\to\gamma_S^\pm}\widetilde{\Delta}(\varnothing|\eta^*) = 0$.

Proof. By (A.140), if $\mu_S = \mu_R = 0$ then

$$\begin{aligned}
&\widetilde{\Delta}(\varnothing|\eta) \\
&= \frac{1}{\varphi(F_l + 1 - F_h) + 1 - \varphi} \\
&\quad \times \begin{pmatrix} \varphi k_1(\gamma_R^2+\gamma_\varepsilon^2) f_l - \varphi F_l k_2 + \varphi k_1(\gamma_R^2+\gamma_\varepsilon^2) f_h \\ +\varphi(1-F_h)k_2 + (1-\varphi)\Delta_0 \end{pmatrix} \\
&= \varphi k_1(\gamma_R^2+\gamma_\varepsilon^2)\frac{f_l + f_h}{\varphi(F_l + 1 - F_h) + 1 - \varphi},
\end{aligned}$$

where the last equality follows from $k_2 = \text{sgn}\,(\gamma_S - \gamma_R)\left(\frac{\gamma_\varepsilon^2 \mu_S}{\gamma_S^2 + \gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2 \mu_R}{\gamma_R^2 + \gamma_\varepsilon^2}\right) = 0$ and $\Delta_0 = \mu_S - \mu_R = 0$ by assumption. Then, since $\lim_{\gamma_R \to \gamma_S^{\pm}} k_1 = 0$, we obtain

$$\mu_S = \mu_R = 0: \lim_{\gamma_R \to \gamma_S^{\pm}} \widetilde{\Delta}(\varnothing|\eta^*) = 0.$$

Claim 2.6. If $\mu_S = \mu_R = 0$, *then* $\lim_{\gamma_R \to \gamma_S^{\pm}} E_S[\widetilde{\Delta}] = E_S[\widetilde{\Delta}|\mu_R = \mu_S, \gamma_R = \gamma_S]$.

Proof. By (A.141), for $\gamma_S \neq \gamma_R$ we have

$$\begin{aligned} E_S[\widetilde{\Delta}] \;=\; & (\varphi F_S(\widetilde{\sigma} - \eta^*) + \varphi(1 - F_S(\widetilde{\sigma} + \eta^*)) + 1 - \varphi)\widetilde{\Delta}(\varnothing|\eta^*) \\ & + \varphi k_1 \begin{pmatrix} F_S(\widetilde{\sigma} + \eta^*)\mu_S - (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma} + \eta^*) \\ -2F_S(\widetilde{\sigma})\mu_S + 2(\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma}) \\ +F_S(\widetilde{\sigma} - \eta^*)\mu_S - (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma} - \eta^*) \end{pmatrix} \\ & + \varphi k_2 \left(F_S(\widetilde{\sigma} + \eta^*) - 2F_S(\widetilde{\sigma}) + F_S(\widetilde{\sigma} - \eta^*)\right). \end{aligned} \tag{A.142}$$

Since for $\mu_S = \mu_R = 0$ we have $\lim_{\gamma_R \to \gamma_S^{\pm}} k_1 = 0$, $k_2 = 0$ and $\lim_{\gamma_R \to \gamma_S^{\pm}} \widetilde{\Delta}(\varnothing|\eta^*) = 0$ (by Claim 2.5), the above expression implies

$$\lim_{\gamma_R \to \gamma_S^{\pm}} E_S[\widetilde{\Delta}] = 0 = E_S[\widetilde{\Delta}|\mu_R = \mu_S, \gamma_R = \gamma_S],$$

where the last equality follows from the fact that under identical priors $\Delta(\sigma) = 0$ for any σ.

Thus, Claim 2.6 implies that for $\mu_S = \mu_R$, $E_S[\widetilde{\Delta}|\gamma_R]$ is again continuous in γ_R at $\gamma_R = \gamma_S$. This completes the proof of Claim 2 for the case $\gamma_R = \gamma_S$.

Claim 3. *The expected perceived disagreement from S's ex ante perspective* $E_S[\widetilde{\Delta}]$ *strictly decreases in* γ_R *if* $\gamma_R > \gamma_S$, $\mu_S > \mu_R$ *and* γ_R *is sufficiently close to* γ_S.

Proof. Assume $\gamma_R > \gamma_S$ and $\mu_S > \mu_R$. Note that the derivations in (A.111) hold independently from whether $\gamma_S > \gamma_R$ or not, thus we still have

$$\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = P_S^{\varnothing} \frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} + \varphi \frac{dk_1}{d\gamma_R} \left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma} + \eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma} - \eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right). \tag{A.143}$$

By Lemma A.15 for $\gamma_R > \gamma_S$ we have

$$\frac{d\widetilde{\Delta}(\varnothing|\eta^*)}{d\gamma_R} = \frac{(f_h + f_l)\gamma_R \gamma_\varepsilon^2 (\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h \varphi + F_l \varphi)}. \tag{A.144}$$

Substituting it back to (A.143) and taking the derivative $\frac{dk_1}{d\gamma_R}$ in the considered parameter case we obtain

$$\begin{aligned} \frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} \;=\; & P_S^{\varnothing} \frac{(f_h + f_l)\gamma_R \gamma_\varepsilon^2 (\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h \varphi + F_l \varphi)} \\ & + \varphi \frac{2\gamma_R \gamma_\varepsilon^2}{(\gamma_R^2 + \gamma_\varepsilon^2)^2} \left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma} + \eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma} - \eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right). \end{aligned} \tag{A.145}$$

Consider the limits of each term as γ_R goes to γ_S from the right. For the first term, we have

$$\lim_{\gamma_R \to \gamma_S^+} \frac{(f_h + f_l)\gamma_R \gamma_\varepsilon^2 (\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h\varphi + F_l\varphi)} = 0, \tag{A.146}$$

since $\lim_{\gamma_R \to \gamma_S^+} f_l = \lim_{\gamma_R \to \gamma_S^+} f_h = 0$ by Claim 2.2 from the proof of Claim 2 (given that $\mu_S > \mu_R$ by assumption).

Consider the second term of (A.145). First, we have

$$\lim_{\gamma_R \to \gamma_S^+} \left(-\int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right) = -\int_{-\infty}^{\infty} \sigma f_S(\sigma) d\sigma = -\mu_S < 0, \tag{A.147}$$

as $\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} - \eta^* = -\infty$ and $\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} = \infty$ by Claim 2.1 from the proof of Claim 2 (given that $\mu_S > \mu_R$ by assumption). Next,

$$\begin{aligned}
& \lim_{\gamma_R \to \gamma_S^+} \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma \\
= & \lim_{\gamma_R \to \gamma_S^+} \left(\int_{-\infty}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{-\infty}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right) \\
= & \lim_{\gamma_R \to \gamma_S^+} \left(F_S(\widetilde{\sigma}+\eta^*) E_S[\sigma | \sigma \le \widetilde{\sigma}+\eta^*] - F_S(\widetilde{\sigma}) E_S[\sigma | \sigma \le \widetilde{\sigma}] \right) \\
= & \lim_{\gamma_R \to \gamma_S^+} \begin{pmatrix} F_S(\widetilde{\sigma}+\eta^*)\mu_S - (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma}+\eta^*) \\ -F_S(\widetilde{\sigma})\mu_S + (\gamma_S^2 + \gamma_\varepsilon^2) f_S(\widetilde{\sigma}) \end{pmatrix} \\
= & \ 0
\end{aligned} \tag{A.148}$$

where the third equality follows by Lemma A.1, and the last equality holds since $\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} + \eta^* = \infty$ and $\lim_{\gamma_R \to \gamma_S^+} \widetilde{\sigma} = \infty$ by Claim 2.1 from the proof of Claim 2 (given that $\mu_S > \mu_R$ by assumption).

(A.145)-(A.148) together imply that for $\mu_S > \mu_R = 0$ we have

$$\lim_{\gamma_R \to \gamma_S^+} \frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = -\varphi \frac{2\gamma_R \gamma_\varepsilon^2}{(\gamma_R^2 + \gamma_\varepsilon^2)^2} \mu_S < 0,$$

which completes the proof of Claim 3.

Claim 4. *The expected perceived disagreement from S's ex ante perspective $E_S[\widetilde{\Delta}]$ strictly increases with γ_R if $\mu_S = \mu_R$ and $\gamma_R > \gamma_S$.*

Proof. Recall from (A.145) that for $\gamma_R > \gamma_S$ we have

$$\begin{aligned}
\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = & \ P_S^{\varnothing} \frac{(f_h + f_l)\gamma_R \gamma_\varepsilon^2 (\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h\varphi + F_l\varphi)} \\
& + \varphi \frac{2\gamma_R \gamma_\varepsilon^2}{(\gamma_R^2 + \gamma_\varepsilon^2)^2} \left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right).
\end{aligned} \tag{A.149}$$

Note that $\widetilde{\sigma} = \frac{\mu_S(\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = 0$ as far as $\mu_S = \mu_R$ (given that μ_R is normalized to 0). Besides,

$\widetilde{\sigma} - \eta^* < \mu_R = 0 < \widetilde{\sigma} + \eta^*$ by Corollary A.1. It follows

$$\begin{aligned} & \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \\ = & \int_{0}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{0} \sigma f_S(\sigma) d\sigma > 0. \end{aligned}$$

This together with (A.149) implies $\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} > 0$ for any $\gamma_R > \gamma_S$ if $\mu_S = \mu_R = 0$.

Claim 5. *Let $\mu_S > \mu_R$. Then, there exists γ'_R, such that the expected perceived disagreement from S's ex ante perspective $E_S[\widetilde{\Delta}]$ strictly increases in γ_R if $\gamma_R > \gamma'_R$.*

Proof. Recall from (A.145) that for $\gamma_R > \gamma_S$ we have

$$\begin{aligned} \frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} = & P_S^{\varnothing} \frac{(f_h + f_l)\gamma_R \gamma_\varepsilon^2 (\gamma_S^2 + \gamma_R^2 + 2\gamma_\varepsilon^2)\varphi}{(\gamma_S^2 + \gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)(1 - F_h\varphi + F_l\varphi)} \\ & + \varphi \frac{2\gamma_R \gamma_\varepsilon^2}{(\gamma_R^2 + \gamma_\varepsilon^2)^2} \left(\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right). \end{aligned} \tag{A.150}$$

The first term is strictly positive. Let us show that the second term is also positive for sufficiently large γ_R. Recall that by Proposition 1(b), it holds $\widetilde{\sigma} - \eta^* < \underline{\sigma} \leq \overline{\sigma} < \widetilde{\sigma} + \eta^*$, where $\underline{\sigma}$ and $\overline{\sigma}$ are the lower and upper status quo signals, respectively. For the case $\gamma_R > \gamma_S$, by (A.134) and (A.136) these signals are equal to:

$$\begin{aligned} \overline{\sigma} &= \frac{(\gamma_S^2 + 2\gamma_\varepsilon^2)(\gamma_R^2 + \gamma_\varepsilon^2)\mu_S}{(\gamma_R^2 - \gamma_S^2)\gamma_\varepsilon^2} > 0, \\ \underline{\sigma} &= \frac{\gamma_S^2 (\gamma_R^2 + \gamma_\varepsilon^2)\mu_S}{(\gamma_S^2 - \gamma_R^2)\gamma_\varepsilon^2} < 0. \end{aligned}$$

Then, we have

$$\begin{aligned} & \int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \\ = & \int_{\widetilde{\sigma}}^{\overline{\sigma}} \sigma f_S(\sigma) d\sigma - \int_{\underline{\sigma}}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma + \int_{\overline{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\underline{\sigma}} \sigma f_S(\sigma) d\sigma \\ > & \int_{\widetilde{\sigma}}^{\overline{\sigma}} \sigma f_S(\sigma) d\sigma - \int_{\underline{\sigma}}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma, \end{aligned} \tag{A.151}$$

where the inequality follows from $\overline{\sigma} > 0$ and $\underline{\sigma} < 0$. Note that

$$\lim_{\gamma_R \to \infty} \widetilde{\sigma} = \lim_{\gamma_R \to \infty} \frac{\mu_S (\gamma_R^2 + \gamma_\varepsilon^2)}{\gamma_R^2 - \gamma_S^2} = \mu_S, \tag{A.152}$$

$$\lim_{\gamma_R \to \infty} \overline{\sigma} = \frac{(\gamma_S^2 + 2\gamma_\varepsilon^2)\mu_S}{\gamma_\varepsilon^2}, \tag{A.153}$$

$$\lim_{\gamma_R \to \infty} \underline{\sigma} = -\frac{\gamma_S^2 \mu_S}{\gamma_\varepsilon^2} = 2\mu_S - \lim_{\gamma_R \to \infty} \overline{\sigma}. \tag{A.154}$$

Denote $l = \lim_{\gamma_R \to \infty} \overline{\sigma}$. Then, taking the limit of the right-hand side of (A.151) we obtain

$$\begin{aligned}
& \lim_{\gamma_R \to \infty} \left(\int_{\widetilde{\sigma}}^{\overline{\sigma}} \sigma f_S(\sigma) d\sigma - \int_{\underline{\sigma}}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right) \\
= & \int_{\mu_S}^{l} \sigma f_S(\sigma) d\sigma - \int_{2\mu_S - l}^{\mu_S} \sigma f_S(\sigma) d\sigma \\
= & \int_{\mu_S}^{l} \sigma f_S(\sigma) d\sigma - \int_{\mu_S}^{l} (2\mu_S - \sigma) f_S(2\mu_S - \sigma) d\sigma \\
= & \int_{\mu_S}^{l} \sigma f_S(\sigma) d\sigma - \int_{\mu_S}^{l} (2\mu_S - \sigma) f_S(\sigma) d\sigma \\
= & \ 2 \int_{\mu_S}^{l} (\sigma - \mu_S) f_S(\sigma) d\sigma > 0,
\end{aligned}$$

where the first equality is by (A.152)-(A.154), the second equality is obtained by integration by substitution, and the inequality is due to $l = \frac{(\gamma_S^2 + 2\gamma_\varepsilon^2)\mu_S}{\gamma_\varepsilon^2} > \mu_S$. Thus,

$$\lim_{\gamma_R \to \infty} \left(\int_{\widetilde{\sigma}}^{\overline{\sigma}} \sigma f_S(\sigma) d\sigma - \int_{\underline{\sigma}}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma \right) > 0,$$

which together with (A.151) implies that there exists sufficiently large γ'_R such that $\int_{\widetilde{\sigma}}^{\widetilde{\sigma}+\eta^*} \sigma f_S(\sigma) d\sigma - \int_{\widetilde{\sigma}-\eta^*}^{\widetilde{\sigma}} \sigma f_S(\sigma) d\sigma > 0$ for any $\gamma_R > \gamma'_R$. Finally, by (A.150) this yields that there exists sufficiently large γ'_R such that $\frac{dE_S[\widetilde{\Delta}]}{d\gamma_R} > 0$ for any $\gamma_R > \gamma'_R$.

To sum up, by the preceding claims we have shown the following properties of the expected perceived disagreement from S's ex ante perspective $E_S[E_R[\Delta(\sigma)|d]] = E_S[\widetilde{\Delta}]$ (given that assumption $\mu_S \geq \mu_R$ is without loss of generality):

- $E_S[\widetilde{\Delta}]$ strictly decreases in γ_R if $\gamma_R < \gamma_S$ (Claim 1);
- $E_S[\widetilde{\Delta}]$ is continuous in γ_R at any $\gamma_R \in (0, \infty)$ (Claim 2);
- $E_S[\widetilde{\Delta}]$ strictly decreases in γ_R if $\mu_S \neq \mu_R$, $\gamma_R > \gamma_S$ and γ_R is sufficiently close to γ_S (Claim 3);
- $E_S[\widetilde{\Delta}]$ strictly increases in γ_R if $\mu_S = \mu_R$ and $\gamma_R > \gamma_S$ (Claim 4);
- $E_S[\widetilde{\Delta}]$ strictly increases in γ_R if $\mu_S \neq \mu_R$ and γ_R is sufficiently large (Claim 5).

Claims 1, 2 and 4 lead to the claim of Proposition 7(a). Claims 1, 2, 3 and 5 lead to the claim of Proposition 7(b). ■

A.5 Application to matching

A.5.1 Proof of Proposition 8

The proof is provided in the main text. ■

A.6 Robustness

A.6.1 Proof of Proposition 9

In what follows, we denote posterior mean and posterior variance of player $i \in \{S, R\}$ given signal σ as $\widetilde{\mu}_i(\sigma)$ and $(\widetilde{\gamma}_R(\sigma))^2$, respectively. We also denote by Δ_0 the prior KL divergence from $N(\mu_S, \gamma_S^2)$ to $N(\mu_R, \gamma_R^2)$:

$$\Delta_0 = \log\left(\frac{\gamma_S}{\gamma_R}\right) + \frac{\gamma_R^2}{2\gamma_S^2} - \frac{1}{2} + \frac{(\mu_S - \mu_R)^2}{2\gamma_S^2}.$$

Claim 1: *For* $\sigma \in \mathbb{R}$, $\Delta(\sigma) = D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ *is symmetric around the agreement signal* $\widetilde{\sigma}$*, strictly decreasing in* σ *for* $\sigma < \widetilde{\sigma}$*, and convex.*

Proof. Generally, for the case of two Gaussian distributions $N(\mu_i, \gamma_i^2)$ and $N(\mu_j, \gamma_j^2)$, $D_{KL}(p, q)$ given in (9) simplifies to the following algebraic expression:

$$D_{KL}(N(\mu_i, \gamma_i^2), N(\mu_j, \gamma_j^2)) = \log\left(\frac{\gamma_j}{\gamma_i}\right) + \frac{\gamma_i^2 + (\mu_i - \mu_j)^2}{2\gamma_j^2} - \frac{1}{2}. \tag{A.155}$$

Hence, $D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ rewrites as

$$\begin{aligned} D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma)) &= \log\left(\frac{\widetilde{\gamma}_S(\sigma)}{\widetilde{\gamma}_R(\sigma)}\right) + \frac{(\widetilde{\gamma}_R(\sigma))^2}{2\left(\widetilde{\gamma}_S(\sigma)\right)^2} - \frac{1}{2} + \frac{(\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma))^2}{2\left(\widetilde{\gamma}_S(\sigma)\right)^2} \\ &= c_1 + c_2\left(\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma)\right)^2, \end{aligned} \tag{A.156}$$

where c_1 and c_2 are some constants (given that $\widetilde{\gamma}_i(\sigma) = \frac{1}{\sqrt{\frac{1}{\gamma_i^2} + \frac{1}{\gamma_\varepsilon^2}}}$ is independent of σ). Recall that $|\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma)|$ is a linear function of $\sigma \in \mathbb{R}$ that is symmetric around the agreement signal $\widetilde{\sigma}$ and strictly decreasing in σ for $\sigma < \widetilde{\sigma}$ if $\gamma_S \neq \gamma_R$ (by (A.20)). It follows that $\left(\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma)\right)^2 = |\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma)|^2$ is symmetric around the agreement signal $\widetilde{\sigma}$, strictly decreasing in σ for $\sigma < \widetilde{\sigma}$ and convex in σ. Consequently, by (A.156) these properties carry over to $D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ (given that $c_2 > 0$).

Claim 2: $\Delta_0 \geq \Delta(\widetilde{\sigma})$.

Proof. We have (noting that $\widetilde{\mu}_S(\widetilde{\sigma}) - \widetilde{\mu}_R(\widetilde{\sigma}) = 0$ by definition):

$$\begin{aligned} &\Delta_0 - \Delta(\widetilde{\sigma}) \\ =\ &\log\left(\frac{\gamma_S}{\gamma_R}\right) + \frac{\gamma_R^2}{2\gamma_S^2} - \frac{1}{2} + \frac{(\mu_S - \mu_R)^2}{2\gamma_S^2} \\ &-\left(\log\left(\frac{\widetilde{\gamma}_S(\widetilde{\sigma})}{\widetilde{\gamma}_R(\widetilde{\sigma})}\right) + \frac{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2}{2\left(\widetilde{\gamma}_S(\widetilde{\sigma})\right)^2} - \frac{1}{2}\right) \geq M(\gamma_S, \gamma_R), \end{aligned} \tag{A.157}$$

where

$$M(\gamma_S, \gamma_R) \equiv \log\left(\frac{\gamma_S}{\gamma_R}\right) + \frac{\gamma_R^2}{2\gamma_S^2} - \left(\log\left(\frac{\widetilde{\gamma}_S(\widetilde{\sigma})}{\widetilde{\gamma}_R(\widetilde{\sigma})}\right) + \frac{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2}{2\left(\widetilde{\gamma}_S(\widetilde{\sigma})\right)^2}\right). \tag{A.158}$$

Substituting $\widetilde{\gamma}_i(\sigma) = \frac{1}{\sqrt{\frac{1}{\gamma_i^2}+\frac{1}{\gamma_\varepsilon^2}}}$ and taking the derivative of $M(\gamma_S, \gamma_R)$ with respect to γ_S we obtain:

$$\frac{dM(\gamma_S, \gamma_R)}{d\gamma_S} = \frac{\gamma_S^4(\gamma_R^2 + \gamma_\varepsilon^2) - \gamma_R^4(\gamma_S^2 + \gamma_\varepsilon^2)}{\gamma_S^3(\gamma_R^2 + \gamma_\varepsilon^2)(\gamma_S^2 + \gamma_\varepsilon^2)} \geq 0 \text{ iff } \gamma_S \geq \gamma_R.$$

Thus, $M(\gamma_S, \gamma_R)$ is U-shaped in γ_S reaching its minimum at $\gamma_S = \gamma_R$, which is equal to 0 by (A.158). Hence, $M(\gamma_S, \gamma_R) \geq 0$ for any γ_S and γ_R that together with (A.157) implies $\Delta_0 \geq \Delta(\widetilde{\sigma})$.

Claim 3: *In equilibrium,* $\widetilde{\Delta}(\varnothing)$ *must be finite while* $\widetilde{\Delta}(\varnothing) > \Delta(\widetilde{\sigma})$.

Proof. First, note that full disclosure of all signals (FD) is never an equilibrium under $\gamma_S \neq \gamma_R$. Indeed, in this case $\widetilde{\Delta}(\varnothing) = \Delta_0$, as non-disclosure happens if and only if S is uninformed. Consequently, S would have an incentive to deviate by concealing any signal σ such that $\Delta(\sigma) > \Delta_0$ (that must be true for sufficiently high and low signals by the convexity of $\Delta(\sigma)$ shown in Claim 1).

The fact that $\widetilde{\Delta}(\varnothing)$ should be finite can then be shown by contradiction. Suppose indeed that $\widetilde{\Delta}(\varnothing)$ is not finite. Then there would exist the FD-equilibrium, as S would always favour disclosing over not disclosing. But the FD-equilibrium cannot exist, as shown above.

The fact that $\widetilde{\Delta}(\varnothing) > \Delta(\widetilde{\sigma})$ is proved also by contradiction. Suppose that $\widetilde{\Delta}(\varnothing) \leq \Delta(\widetilde{\sigma})$. Then, since $\Delta(\widetilde{\sigma}) < \Delta(\sigma)$ for any non-empty signal $\sigma \neq \widetilde{\sigma}$ by Claim 1, it must hold that also $\widetilde{\Delta}(\varnothing) < \Delta(\sigma)$ for any non-empty signal $\sigma \neq \widetilde{\sigma}$. Consequently, in equilibrium all signals (except for possibly $\widetilde{\sigma}$) must be concealed. But this implies that after non-disclosure, R acknowledges that with a positive probability S holds a non-empty signal different from $\widetilde{\sigma}$. But for any such signal σ, it holds $\Delta(\sigma) > \Delta(\widetilde{\sigma})$ by Claim 1. Since $\widetilde{\Delta}(\varnothing)$ is a probability weighting over $\Delta(\sigma)$ at all concealed signals and Δ_0, while $\Delta_0 \geq \Delta(\widetilde{\sigma})$ by Claim 2, it follows that $\widetilde{\Delta}(\varnothing) > \Delta(\widetilde{\sigma})$, which yields a contradiction.

Claim 4: Any equilibrium under $\gamma_S \neq \gamma_R$ *features finite* $\eta^* > 0$ *such that* S *discloses* σ *if and only if* $\sigma \in [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$.

Proof. As shown in Claim 3, in any equilibrium $\widetilde{\Delta}(\varnothing)$ must be finite and $\widetilde{\Delta}(\varnothing) > \Delta(\widetilde{\sigma})$. Consequently, by Claim 1 there exists a unique $\eta^* > 0$ such that $\Delta(\widetilde{\sigma} - \eta^*) = \Delta(\widetilde{\sigma} + \eta^*) = \widetilde{\Delta}(\varnothing)$ while for any non-empty signal, $\Delta(\sigma) \leq \widetilde{\Delta}(\varnothing)$ iff $\sigma \in [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$. It follows that the equilibrium disclosure set is given by $[\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$. ■

Proposition A.1 *Define the symmetric KL divergence as*

$$D^\sharp_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma)) = D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma)) + D_{KL}(\widetilde{g}_S(\cdot\,|\sigma), \widetilde{g}_R(\cdot\,|\sigma)),$$

and Bhattacharyya distance between two Gaussian distributions $N(\mu_i, \gamma_i^2)$ *and* $N(\mu_j, \gamma_j^2)$ *as*

$$D_B(N(\mu_i, \gamma_i^2), N(\mu_j, \gamma_j^2)) = \frac{1}{4}\log\left(\frac{1}{4}\left(\frac{\gamma_i^2}{\gamma_j^2} + \frac{\gamma_j^2}{\gamma_i^2} + 2\right)\right) + \frac{1}{4}\left(\frac{(\mu_i - \mu_j)^2}{\gamma_i^2 + \gamma_j^2}\right). \quad \text{(A.159)}$$

Then, if $\Delta(\sigma)$ *is given either by* $D_{KL}(\widetilde{g}_S(\cdot\,|\sigma), \widetilde{g}_R(\cdot\,|\sigma))$, $D^\sharp_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ *or* $D_B(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$, *any equilibrium features a finite non-degenerate disclosure interval*

$I = [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$ such that S discloses σ if and only if $\sigma \in I$, where $\widetilde{\sigma}$ is the unique signal for which the posterior means of S and R are equal.

Proof. As in the proof of Proposition 9, denote posterior mean and posterior variance of player $i \in \{S, R\}$ given signal σ as $\widetilde{\mu}_i(\sigma)$ and $(\widetilde{\gamma}_R(\sigma))^2$, respectively. We also denote by Δ_0 the prior disagreement according to the respective disagreement measure.

Claim 1: *For $\sigma \in \mathbb{R}$, $\Delta(\sigma)$ is symmetric around the agreement signal $\widetilde{\sigma}$, strictly decreasing in σ for $\sigma < \widetilde{\sigma}$, and convex.*

Proof. The claim for $\Delta(\sigma) = D_{KL}(\widetilde{g}_S(\cdot\,|\sigma), \widetilde{g}_R(\cdot\,|\sigma))$ follows by the same arguments as for $\Delta(\sigma) = D_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ (see Claim 1 of Proposition 9). Thus, by definition $D^{\sharp}_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ is the sum of two functions that are symmetric in $\sigma \in \mathbb{R}$ around the agreement signal $\widetilde{\sigma}$, strictly decreasing in σ for $\sigma < \widetilde{\sigma}$ and convex in $\sigma \in \mathbb{R}$. It follows immediately that these properties carry over to $D^{\sharp}_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$.

Finally, note that by (A.159) $D_B(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$ also rewrites as $c_1 + c_2\left(\widetilde{\mu}_S(\sigma) - \widetilde{\mu}_R(\sigma)\right)^2$, for some constants c_1 and $c_2 > 0$ (given that $\widetilde{\gamma}_i(\sigma) = \frac{1}{\sqrt{\frac{1}{\gamma_i^2} + \frac{1}{\gamma_\varepsilon^2}}}$ is independent of σ). The conclusion then follows by the same argument as the proof of Claim 1 of Proposition 9.

Claim 2: $\Delta_0 \geq \Delta(\widetilde{\sigma})$.

Proof. If $\Delta(\sigma) = D_{KL}(\widetilde{g}_S(\cdot\,|\sigma), \widetilde{g}_R(\cdot\,|\sigma))$ or $\Delta(\sigma) = D^{\sharp}_{KL}(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$, the claim follows by the same argument as the proof of Claim 2 of Proposition 9.

Consider the remaining case $\Delta(\sigma) = D_B(\widetilde{g}_R(\cdot\,|\sigma), \widetilde{g}_S(\cdot\,|\sigma))$. Then,

$$\begin{aligned} & \Delta_0 - \Delta(\widetilde{\sigma}) \\ = \; & \frac{1}{4}\log\left(\frac{1}{4}\left(\frac{\gamma_S^2}{\gamma_R^2} + \frac{\gamma_R^2}{\gamma_S^2} + 2\right)\right) + \frac{1}{4}\frac{(\mu_S - \mu_R)^2}{\gamma_S^2 + \gamma_R^2} \\ & -\frac{1}{4}\log\left(\frac{1}{4}\left(\frac{(\widetilde{\gamma}_S(\widetilde{\sigma}))^2}{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2} + \frac{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2}{(\widetilde{\gamma}_S(\widetilde{\sigma}))^2} + 2\right)\right) \geq M_2(\gamma_S, \gamma_R), \end{aligned} \tag{A.160}$$

where

$$\begin{aligned} M_2(\gamma_S, \gamma_R) \;\equiv\; & \frac{1}{4}\log\left(\frac{1}{4}\left(\frac{\gamma_S^2}{\gamma_R^2} + \frac{\gamma_R^2}{\gamma_S^2} + 2\right)\right) \\ & -\frac{1}{4}\log\left(\frac{1}{4}\left(\frac{(\widetilde{\gamma}_S(\widetilde{\sigma}))^2}{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2} + \frac{(\widetilde{\gamma}_R(\widetilde{\sigma}))^2}{(\widetilde{\gamma}_S(\widetilde{\sigma}))^2} + 2\right)\right). \end{aligned} \tag{A.161}$$

Substituting $\widetilde{\gamma}_i(\sigma) = \frac{1}{\sqrt{\frac{1}{\gamma_i^2} + \frac{1}{\gamma_\varepsilon^2}}}$ and taking the derivative of $M_2(\gamma_S, \gamma_R)$ with respect to γ_R we obtain:

$$\frac{dM_2(\gamma_S, \gamma_R)}{d\gamma_R} = \frac{\gamma_R\left(\gamma_R^2 - \gamma_S^2\right)\left(\gamma_R^2\gamma_\varepsilon^2 + \gamma_S^2\left(2\gamma_R^2 + 3\gamma_\varepsilon^2\right)\right)}{2\left(\gamma_S^2 + \gamma_R^2\right)\left(\gamma_R^2 + \gamma_\varepsilon^2\right)\left(\gamma_R^2\gamma_\varepsilon^2 + \gamma_S^2\left(2\gamma_R^2 + \gamma_\varepsilon^2\right)\right)} \geq 0 \text{ iff } \gamma_R \geq \gamma_S.$$

Thus, $M_2(\gamma_R, \gamma_S)$ is U-shaped in γ_R reaching its minimum at $\gamma_R = \gamma_S$, which is equal to 0 by (A.161). Hence, $M_2(\gamma_S, \gamma_R) \geq 0$ for any γ_S and γ_R that together with (A.160) implies $\Delta_0 \geq \Delta(\widetilde{\sigma})$.

Claim 3: *In equilibrium,* $\widetilde{\Delta}(\varnothing)$ *must be finite while* $\widetilde{\Delta}(\varnothing) > \Delta(\widetilde{\sigma})$.

The claim follows by the same argument as the proof of Claim 3 of Proposition 9 (given current Claim 1 and Claim 2).

Claim 4: Any equilibrium under $\gamma_S \neq \gamma_R$ *features finite* $\eta^* > 0$ *such that* S *discloses* σ *if and only if* $\sigma \in [\widetilde{\sigma} - \eta^*, \widetilde{\sigma} + \eta^*]$.

The claim follows by the same argument as the proof of Claim 4 of Proposition 9 (given current Claims 1-3). ■

A.7 The value of public information

A.7.1 Proof of Proposition 10

a) Consider the case of equal prior means $\mu_1 = \mu_2 = \mu$. Since $\gamma_1 \neq \gamma_2$ by assumption, we have $\Delta(\sigma) = |E_1[\omega|\sigma] - E_2[\omega|\sigma]| > 0$ for any $\sigma \neq \widetilde{\sigma}$. Hence, for $i = 1, 2$, it holds $E_i[\Delta(\sigma)] > 0$ and thus

$$V_i(\mu_1, \mu_2, \gamma_1, \gamma_2) = |\mu - \mu| - E_i[\Delta(\sigma)] = -E_i[\Delta(\sigma)] < 0.$$

It follows, given the continuity of V_i in μ_2, that $V_i(\mu_1, \mu_2, \gamma_1, \gamma_2) < 0$ if μ_2 slightly deviates from μ_1, i.e., if $|\mu_1 - \mu_2|$ is sufficiently small.

b) Throughout the proof, without loss of generality we assume $\mu_2 \geq \mu_1$ (the proof for the case $\mu_2 \leq \mu_1$ is fully symmetric). In Steps 1 and 2 below we show, respectively, that V_1 and V_2 strictly increase with μ_2 while becoming strictly positive for sufficiently large μ_2 (which proves point b) of the proposition given that $\mu_2 \geq \mu_1$ by assumption).

Step 1. Let us show that $\frac{\partial V_1}{\partial \mu_2} > 0$ while V_1 becomes strictly positive for sufficiently large μ_2. We have:

$$\begin{aligned} V_1 &= |\mu_1 - \mu_2| - \int_{-\infty}^{\infty} \Delta(\sigma) f_1(\sigma) d\sigma \\ &= (\mu_2 - \mu_1) \\ &\quad - \left(\int_{-\infty}^{\widetilde{\sigma}} -(k_1\sigma + k_2) f_1(\sigma) d\sigma + \int_{\widetilde{\sigma}}^{\infty} (k_1\sigma + k_2) f_1(\sigma) d\sigma \right), \end{aligned} \tag{A.162}$$

where the last equality is by (A.20) once we replace S and R with 1 and 2, respectively, so that

$$k_1 = \operatorname{sgn}[\gamma_1 - \gamma_2] \left(\frac{\gamma_1^2}{\gamma_1^2 + \gamma_\varepsilon^2} - \frac{\gamma_2^2}{\gamma_2^2 + \gamma_\varepsilon^2} \right), \tag{A.163}$$

$$k_2 = \operatorname{sgn}[\gamma_1 - \gamma_2] \left(\frac{\gamma_\varepsilon^2 \mu_1}{\gamma_1^2 + \gamma_\varepsilon^2} - \frac{\gamma_\varepsilon^2 \mu_2}{\gamma_2^2 + \gamma_\varepsilon^2} \right). \tag{A.164}$$

Taking the derivative with respect to μ_2, we obtain

$$\begin{aligned}
\frac{dV_1}{d\mu_2} &= 1 - \frac{d\widetilde{\sigma}}{d\mu_2}\Delta(\widetilde{\sigma})f_1(\widetilde{\sigma}) + \frac{d\widetilde{\sigma}}{d\mu_2}\Delta(\widetilde{\sigma})f_1(\widetilde{\sigma}) \\
&\quad -\frac{dk_2}{d\mu_2}\left(\int_{\widetilde{\sigma}}^{\infty} f_1(\sigma)d\sigma - \int_{-\infty}^{\widetilde{\sigma}} f_1(\sigma)d\sigma\right) \\
&= 1 + \text{sgn}\,[\gamma_1 - \gamma_2]\frac{\gamma_\varepsilon^2}{\gamma_2^2 + \gamma_\varepsilon^2}\left(1 - 2F_1(\widetilde{\sigma})\right) \\
&> 1 - \frac{\gamma_\varepsilon^2}{\gamma_2^2 + \gamma_\varepsilon^2} \\
&> 0. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (A.165)
\end{aligned}$$

Further, since $\lim_{\mu_2\to\infty}\widetilde{\sigma} = \lim_{\mu_2\to\infty}\frac{\mu_2(\gamma_1^2+\gamma_\varepsilon^2)-\mu_1(\gamma_2^2+\gamma_\varepsilon^2)}{\gamma_1^2-\gamma_2^2}$ $=\text{sgn}[\gamma_1 - \gamma_2]\infty$ so that $\lim_{\mu_2\to\infty} F_1(\widetilde{\sigma}) = 1(0)$ if $\gamma_1 > (<)\gamma_2$, (A.165) yields

$$\lim_{\mu_2\to\infty}\frac{dV_1}{d\mu_2} = 1 - \frac{\gamma_\varepsilon^2}{\gamma_2^2 + \gamma_\varepsilon^2} > 0,$$

i.e., V_1 becomes linearly increasing in the limit as $\mu_2 \to \infty$, and hence it should become positive for sufficiently large μ_2.

Thus, we have shown the claim of the proposition for V_1.

Step 2. Let us show that V_2 also strictly increases in μ_2 at any $\mu_2 \geq \mu_1$ and becomes strictly positive for sufficiently large μ_2. The proof proceeds via three claims.

Claim 1: *For any $\mu_2 \geq \mu_1$, it holds $\frac{dV_2}{d\mu_1} < 0$. Besides, $V_2 > 0$ for sufficiently small μ_1.*

Proof: The claim follows by the same arguments as in Step 1.

Claim 2: V_2 strictly increases in μ_2 at any $\mu_2 \geq \mu_1$.

Proof: Consider some prior means $\mu_2' \geq \mu_1'$. Note that, generally, both components of V_2, the prior disagreement $|\mu_1 - \mu_2|$ and the expected posterior disagreement $E_2[\Delta(\sigma)]$, do not change if both prior means μ_1 and μ_2 are shifted by the same amount. Then, for any $x > 0$, we have

$$V_2(\mu_1', \mu_2' + x) = V_2(\mu_1' - x, \mu_2'). \qquad (A.166)$$

In turn, by Claim 1

$$V_2(\mu_1' - x, \mu_2') > V_2(\mu_1', \mu_2').$$

This together with (A.166) implies that for any $x > 0$,

$$V_2(\mu_1', \mu_2' + x) > V_2(\mu_1', \mu_2').$$

Since this holds for any given $\mu_2' \geq \mu_1'$ and $x > 0$, V_2 strictly increases in μ_2 at any $\mu_2 \geq \mu_1$.

Claim 3: *$V_2 > 0$ for sufficiently large μ_2.*

Proof: Consider again arbitrary prior means $\mu_2' \geq \mu_1'$. For sufficiently large x we have

$$V_2(\mu_1', \mu_2' + x) = V_2(\mu_1' - x, \mu_2') > 0,$$

where the equality is by (A.166) and the inequality is by Claim 1. It follows that for any given μ_1, $V_2(\mu_1, \mu_2) > 0$ for sufficiently large μ_2. ■