USC FBE APPLIED ECONOMICS WORKSHOP
presented by ***Tracy Lewis***
FRIDAY, Mar. 10, 2011
1:30 pm - 3:00 pm, **Room: HOH-706**

Selective Disclosure of Public Information: Who Needs to Know?*

Qi Chen Tracy R. Lewis Yun Zhang

August 2010

Abstract

Credibly communicating information for agents to act upon is a challenge for central planners and managers. This paper studies the optimal disclosure policy when the relative accuracy (reliability) of the public information cannot be verified and when agents' actions are either strategic complements or substitutes for each other. We find that while disclosure of public information improves social surplus when its accuracy is known, access to public information may be restricted to select agents in order to credibly communicate its reliability. Specifically, we show that a separating equilibrium exists where, when agents' actions are strategic complements (substitutes), the planner with more (less) accurate public information discloses the information only to a subset of the agents. The rationale for this pattern of selective disclosure is best understood as a response to agents' misuse or misinterpretation of public announcements in different strategic settings. Further, when agents differ in the accuracy of their private information, we show that it is optimal to selectively disclose to those with relatively less accurate private information. Our analyses have implications for how public information is best disseminated when agents' actions have strategic effects on each other, such as those involving controlling congestion, beauty contests, macroeconomic stabilizations and promoting uniformity and standardization in design and adoption.

*Preliminary draft and comments welcome. Qi Chen and Tracy Lewis acknowledge financial support from the Fuqua School of Business at Duke University. Yun Zhang acknowledges financial support from George Washington University School of Business and Fuqua School of Business at Duke University. We appreciate comments from workshop participants at Duke University, George Washington University, and the Chicago-Minnesota Accounting Theory Conference. We also thank Bjorn Jorgensen for helpful discussions. Address correspondence to Qi Chen at qc2@duke.edu, Tracy Lewis at tracyl@duke.edu and Yun Zhang at yunzhang@gwu.edu.

1 Introduction

It has long been argued that pubic disclosure is an important instrument in the defense against coordination failure. Coordination failure arises when a group of agents working independently outside or within an organization are deciding what action to take to maximize their private surplus. The agents employ private independent information of the state of their environment to choose a best action. Since agents do not fully internalize the impact of their actions on other agents' surplus, a second best outcome results whereby agents rely on inaccurate and incomplete private information to choose actions that are not coordinated with the actions of others. It is typically believed that public disclosure of unbiased and reliable information on the fundamentals of the environment is an effective antidote for coordination failure in most environments.[1] It is argued that individuals who base decisions on public and private information can better predict the behaviors of other agents and thereby increase social surplus.

To what extent does this conclusion apply when individuals are unable to gauge the accuracy of the public information compared to their private information? For many situations, it is unrealistic to suppose individuals know how much more reliable public disclosure is compared to their private information and therefore how they should weigh public and private information in making decisions. For instance, consider a worker deciding whether to return to school for retraining or remain at his current job until the recession ends. Government predictions call for increased unemployment in various sectors. How can the worker assess the accuracy of these predictions without knowing the data on which they are based? What are the government's motives for making such predictions and how should the worker treat this information when deciding what to do?[2] And, as for the government,

[1] See, for instance, the responses of Hellwig (2005), Svenson (2006) and Angeletos and Pavan (2007), to Morris and Shin (2002) who argue that public information may exaccerbate coordination failiure in some circumstances.

[2] The incentives for the government to persuade individuals and the suspicions of the public about the motives of the government in releasing information are present in every day news. For example, a recent report recommending that women younger than 50 should normally not require breast cancer screening was met with much skepticism and suspicion regarding the motives for the release of these findings. The study cited evidence that the incidence of cancer prior to 50 is not sufficiently high nor is it detected sufficiently often to warrant screening. Skeptics cited the real motive for the disclosure was to divert health care resources to other fields and treatment areas. Advocates of the findings who believed that health care leaders have oversimplified and oversold the benefits of cancer screening were however, cautiously optimistic that the disclosure would curtail ineffective screening, claiming that "the biggest problem will be women who don't believe it (the disclosure)."

how much public information should it release, in view of its accuracy and reliability relative to agents' private information? Our goal in this paper is to assess the value and role for public disclosure when the relative accuracy of the information can not be observed.

Our investigation centers on a model due to Angeletos and Pavan (2007), Hellwig (2005) and Morris and Shin (2002) in which a large group of agents, working for themselves or for an organization, decide how to direct their personal effort. The agents each receive a piece of private, conditionally independent and normally distributed information about the state of the world upon which to act. The aggregate surplus the agents generate is quadratic in the agents' actions and the state; and it depends on two factors: (i) the individual output derived from each agent's action and (ii) the joint output resulting from the interaction between the agents' different actions which may complement or substitute for each other. If the agents work in insolation, each individual selects an action to maximize his own surplus ignoring the effects of his decision on other agents. As a result, the agents choose uncoordinated, suboptimal actions based on incomplete private information. On the other hand, a benevolent social planner who provides public information about the environment, can increase social surplus by inducing greater coordination between the agents. However, the way in which the public information is disclosed and the value it has for coordinating agents' behavior depends on whether agents can observe how much more accurate the disclosure is relative to their private information. Specifically, we find that in the setting of our model:

1. When agents can verify the accuracy of public information relative to their own private information, full disclosure of the public information to all agents always increases social welfare.

2. However, if agents can not observe the relative accuracy of the public information, it is not always true that full disclosure to all agents maximizes social welfare. If agents' actions are complementary, then it is optimal to selectively disclose to some but not all agents when public information is relatively accurate. On the other hand, if agents actions are substitutes, then it is optimal to selectively disclose to some agents when public information is relatively less accurate.

The rationale for partial disclosure is that the policy maker may only credibly signal the accuracy of the public information by restricting its access to some agents. The planner's desire to manipulate the agents' attention to public information will, understandably, cause agents to question the planner's motives for providing public information. When agent actions are complements, the planner has an incentive to persuade agents that her disclosure is more accurate. Agents choose similar actions which increases social welfare, if they pay greater attention to public information. The planner signals that her disclosure is more accurate by withholding information from some subset of the agents. To illustrate, assume the planner is either a high type (with more precise information to disclose) or a low type (with less precise information to disclose). Given the value of information for decision making, both types of planners want to provide all agents with more information. However, planners of either type would consider withholding information from some agents, to persuade the remaining agents of the importance of the public disclosure. We demonstrate in what follows, that the cost of persuasion – which is the foregone social surplus from rationing – is lower for the high type than for the low type planner. This leads to a signaling equilibrium where the high type planners ration information to certain agents to signal the reliability of their disclosure, while the low type planners disclose to all agents. The irony that public disclosure is rationed precisely when it is most informative, reflects the planner's cost of credibly communicating the reliability of her information.

When agents' efforts are substitutes, signaling the reliability of public disclosure works in the opposite way. When agents' efforts are substitutes (e.g., conserving energy) so that when they act too much alike (e.g., by turning off their air-conditioning all at the same time), social surplus is diminished. Now the planner wishes to persuade agents that her information is less accurate to induce them to act more independently. We demonstrate the cost of persuasion is lower for the low type than for the high type planner. Hence in equilibrium, information is rationed when it is least accurate and it is widely disseminated when it is most reliable. Recall, that agents' (excess) focus on public disclosure is welfare decreasing in this case, so now the irony is that information is most available precisely, when it is most accurate but least valuable for raising social surplus.

We also analyze which subsets of agents to disclose to when agents differ in the accuracy of their private information. We find that in most cases when the planners ration public information, they would disclose public information only to agents with less accurate private information. The general intuition is that these agents are the subset that minimizes the planner's cost to signal through partial disclosure. Take for example the case of strategic complementarity when more coordination in agents' actions is desired. Separation through information rationing hinges on the fact that the high type planner's signal is more informative about the state of the nature than the low type's. By restricting her disclosure only to agents with less accurate private signals, the high type is able to reinforce such advantage. This is because the actions of those less informed agents will be affected by the disclosed signal the most, and will be closer to the true state when the public signal is of high accuracy. At the same time, if left on their own, agents with more accurate private information will choose actions that are closer to the true state than their less informed counterparts. Consequently, by disclosing to a set of less informed agents (i.e. the inner circle), the high type planner is able to strengthen her ability (relative to the mimicing low type) to avoid large deviations between actions taken by the agents in the inner circle and those by the outer circle (i.e. agents from whom the planner withold information).

Signalling information accuracy through selective disclosure provides some insights about the different ways that information is disseminated to the public. For instance, despite widespread calls for transparency in public disclosure, some news is selectively leaked to a subset of the public while other information is widely disseminated. Our analysis suggests that this may be explained by the disclosers' desire to signal the reliability of the news and how it is interpreted. Our analysis predicts that public disclosure of information will be managed differently across independent issues depending on whether more or less similarity in actions is desired. Hence, more accurate public information disclosure is targeted towards the less informed agents in environments where greater synchronization of actions is desired. In contrast, when greater variations in response to underlying fundamentals of the market is desired, it is the less accurate information that is disclosed to the less informed agents.

Before we move to our model, it is important to place our analysis in the context of the recent literature on disclosure of public information. Our analysis is most closely related to the aforementioned papers of Morris and Shin (2002), Hellwig (2005), and Angeletos and Pavan (2007) on the social value of public information and to the organizational design literature exemplified by Alonso, et al. (2008) and Rantakari (2008) on the value of centralized communication in organizations. This literature demonstrates the value of disclosing public information of known precision. The insights gained from these papers are complementary to our findings. In both cases, public information is shown to reduce coordination failure. However, when the accuracy of the information is unknown, disclosing information is more difficult and therefore of less value in coordinating agents actions. In addition, our analysis shows how agents infer the accuracy of public information by observing how it is disclosed.

Closely related to this is the growing financial literature exemplified by Hirshleifer, et al (2004) and Hirshleifer and Teoh (2003) that study the effects of different presentations of financial information on market prices when investors have limited attention and processing power. These studies rationalize why practitioners care about the choice between equivalent forms of disclosure in a market where investors react differently to equivalent information. Our model offers a complementary explanation for alternative forms of disclosure by showing that different degrees of disclosure can signal to partially informed agents (or investors) how they should interpret public information. Combining the insights from these different approaches to analyzing public disclosure may help us understand why agents apparently under- or over- react to public announcements in different settings.

The "libertarian paternalism" literature also examines ways of providing agents with information to improve their decisions (e.g., Sunstein and Thaler (2003) and Camerer, et al. (2003)). In that case, however, public information is filtered to induce agents to make the "right" decisions, as defined by the policy maker. This approach presumes that agents are unable to process information to act in their own best interest. Here, instead, we presume that agents can process information correctly, if they know how to weight it relative to their private information. One virtue of the model proposed here is that it preserves agents' consumer sovereignty to make their own decisions without coercion or interference from third parties.

Our plan for the rest of the paper is as follows. Section 2 sets up the model, discusses the preferences and information sets for the players (agents and central planner). Section 3 solves the model for both the case when the relative accuracy of public disclosure is known and when it is not unknown. Section 4 extends the model to the case where agents differ in the accuracy of their private information and studies the optimal set of inner circle agents. It also analyzes how the signalling equilibrium affects the planner's incentive to acquire more information in the first place. Section 5 concludes. The appendix contains proofs of formal results that do not appear in the body of the paper.

2 Model Setup

When is public information beneficial and how is it best disclosed? To answer these questions, we propose and analyze the following model that captures the essential features of public disclosure.

2.1 Decision makers

Consider an economy or organization consisting of a continuum of measure of one of agents indexed by i uniformly distributed over $[0,1]$. For simplicity, we adopt a standard reduced form specification due to Angeletos and Pavan (2007) of the agents' preferences in which each agent i is risk neutral with utility,

$$U_i^a = Ae_i - \frac{1}{2}e_i^2. \tag{1}$$

Agent i's surplus, U_i^a, is a function of his action choice (or effort), $e_i \in \mathbb{R}$, and A which represents the marginal return to action. We define $\bar{e} \equiv \int_0^1 e_i di$ which measures the mean action and A as

$$A \equiv (1-\rho)\, v + \rho\bar{e}, \tag{2}$$

where v is a random state (reflecting underlying conditions) with a diffuse prior and $\rho\bar{e}$ measures the impact of other agent's actions on the individual returns to agent i from his action.

It is instructive to rewrite agents' objective function as

$$\begin{aligned} U_i^a &= \left[(1-\rho)\, v + \rho\bar{e}\right] e_i - \frac{1}{2}e_i^2 \\ &= \underbrace{-\frac{1}{2}(1-\rho)(e_i - v)^2}_{\text{"Better Action"}} \underbrace{-\frac{1}{2}\rho\,(e_i - \bar{e})^2}_{\text{"Coordination"}} + \frac{1}{2}\rho\bar{e}^2 + \frac{1}{2}(1-\rho)\, v^2. \end{aligned} \tag{3}$$

Since an individual agent treats the last two terms in (3) as exogenous, he chooses his action to balance two objectives. The first objective is to minimize deviations of his action from the state, captured by the "Better Action" term, $-\frac{1}{2}(1-\rho)(e_i - v)^2$. The second is to minimize the distance between his own action and the average action which is captured by the "Coordination" term, $-\frac{1}{2}\rho\,(e_i - \bar{e})^2$. The sign of the coefficient, ρ, indicates whether agents' actions are complements or substitutes. Specifically, when $\rho > 0$, actions are strategic complements: agents benefit when their actions are closer to each other; in contrast, when $\rho < 0$, actions are strategic substitutes: agents benefit when their actions are further apart.

The random state of nature is unknown prior to each agent's choice of action. For simplicity, agents share a common prior that v is uniformly distributed on $\mathbb{R}$. Prior to choosing e_i, agent i receives a private, noisy signal s_i of v:

$$s_i = v + \varepsilon_i,$$

where $\varepsilon_i \sim N(0, 1/\beta)$ and $cov(\varepsilon_i, v) = cov(\varepsilon_i, \varepsilon_k) = 0$, $\forall i \neq k$. Agents are unable to observe the precision of their signal, β, which is assumed to be the same for all agents for the time being. Agents only know their information is relatively inaccurate compared to the public information that is available (more on this, refer to section 2.2).

The description above could depict a setting where independent agents in the economy decide on how much to invest to maximize their expected private wealth. Alternatively, this could represent separate divisions of a firm that are paid according to the surplus they independently generate. In either case, the actions of different agents may be strategic complements or substitutes for each other depending on the setting.

2.2 Social planner: preferences, information and disclosure policy

A benevolent social planner exists to help manage the agents' behavior to maximize social surplus. Social surplus, U^P, is the simple aggregation of individual agents' surplus as determined by (1) and (2), given by

$$U^P = \int_0^1 U_i di = A\bar{e} - \frac{1}{2}\int_0^1 e_i^2 di = (1-\rho)\, v\bar{e} - (1-2\rho)\,\frac{1}{2}\bar{e}^2 - \frac{1}{2}\int_0^1 (e_i - \bar{e})^2\, di.$$

We assume $\rho < 1/2$ to ensure U^P is concave in $\bar{e}$ and thereby well behaved. Similar to (1), we can rewrite the social surplus as:

$$U^P = \underbrace{-\frac{1}{2}(1-\rho)\int_0^1 (e_i - v)^2\, di}_{\text{"Better Action"}} - \underbrace{\frac{1}{2}\rho\int_0^1 (e_i - \bar{e})^2\, di}_{\text{"Coordination"}} + \frac{1}{2}\rho\bar{e}^2 + \frac{1}{2}(1-\rho)\, v^2.$$

Thus, the social planner, like each agent, cares about taking the right action as well as achieving coordination. The desired coordination is determined by the sign of ρ indicating whether actions are complements or substitutes.

The planner privately observes a public signal, z^τ of v: that may be of two possible types, $\tau \in \{h, l\}$ with

$$z^\tau = v + \eta^\tau,$$

where $\eta^\tau \sim N(0, 1/\alpha^\tau)$ with precision α^τ and $cov\,(v, \eta^\tau) = 0$. Conditional on the true state of nature, v, the planner's signal is independent of the agents' signals, i.e., $cov\,(\varepsilon_i, \eta^\tau) = 0$, $\forall i$.

It is common knowledge that the accuracy of the public signal (z^τ) relative to the private signals s_i is $\lambda^\tau = \alpha^\tau/\beta \in \{\lambda^l, \lambda^h\}$. The type, τ, indicates the relative precision of the signal, with $1 < \lambda^l < \lambda^h$ so that the $h - high$ type of signal is more accurate than the $l - low$ type of signal which in turn is more accurate than the agents' private signals. We assume the ex ante probability that the planner's signal type is h is $q \in [0, 1]$. When q $= 1$ or 0, the precision of the public information is common information, as agents know the type of the signal with certainty. However when $q \in (0, 1)$ we assume agents can not observe the type of signal that results, and therefore are uncertain of the relative accuracy of the public information. This is a noteworthy feature of our model that departs from standard analyses of public disclosure which assume common knowledge of the public and private signal accuracies. Here, this assumption is important in enabling us to analyze how the planner signals the relative accuracy of public information by her disclosure policy.

2.3 Disclosure strategy and disclosure game

The planner serves as a faithful representative for the agents. Whatever public information she discloses is truthful; she can not transmit false or distorted information. The planner may only influence the action each agent selects and thereby the social surplus that is generated by the *transparency* of her public disclosure. The planner discloses to each agent i a signal, z_i^τ, given by

$$z_i^\tau = \delta_i^\tau z^\tau + (1 - \delta_i^\tau)\,\emptyset; \;\; \delta_i^\tau \in \{0, 1\}\,.$$

The agent observes the information public signal z^{τ} when $\delta_i^{\tau} = 1$; otherwise he receives the *null* signal, $\emptyset$, that contains no information.

After observing the type of public signal, τ, but prior to observing the realization of the signal, z^{τ}, the planner selects a disclosure policy D which is observed by all agents. D is a selection of a subset $D \subseteq [0,1]$ of agents who are publicly informed with $\delta_i^{\tau} = 1$. Let $\overline{\delta}^{\tau} = \int_{i \in D} \delta_i^{\tau} di$ be the measure of D. We refer to D ($[0,1] \backslash D$)as the planner's "inner circle" ("outer circle) who are (not) privy to her information. A completely *transparent* D is one where $\overline{\delta}^{\tau} = 1$ and $D = [0,1]$, whereas an *opaque* D is one where the inner circle is a "proper" subset of the agent population, $D \subset [0,1]$. We assume communication between agents is not possible so that an agent $j \notin D$ may not observe or infer anything about the realization of the pubic signal.

Although the planner's information type can not be directly observed, agents may infer the relative precision of public information indirectly by observing the planner's disclosure policy. Recall that a disclosure policy is a selection of agents $D \subseteq [0,1]$ who are publicly informed, independent of the content or realization of the information signal. The planner with a certain *type* of information may signal the informativeness of her signal z by the transparency of her disclosure. For instance, a *high type* planner may wish to persuade agents that her information is very accurate by restricting the disclosure of the public signal to a select set of agents. Since all agents are identical in the sense that their private information precision β is the same, we may without loss of generality, restrict attention to disclosure policies that specify the size of the inner circle, and not the identity of the agents in the circle. Specifically, we consider disclosure policies,

$$DP : \Lambda \rightarrow [0,1]$$

where DP is a mapping from the set of types $\Lambda = \{\lambda^{\tau}\}_{\tau=l,h}$ into a number between 0 and 1 that specifies the fraction $\overline{\delta}^{\tau} \in [0,1]$ of the agents in the economy.

The disclosure of public information is determined by a three stage signaling game. In the *first stage* of the game, the planner observes her accuracy type $\tau = l, h$. In *stage two,* the planner selects, a disclosure policy $\left(\overline{\delta}\right)$ to maximize her expected surplus. Finally, in *stage 3*, the planner discloses the public information according to her already determined disclosure policy and each agent updates his prior beliefs about the signal accuracy and chooses an action to maximize his private surplus. In what follows, we characterize the

Bayesian Perfect Nash equilibrium for this game and describe the role of the disclosure policy as a signal of the value of public information.

3 Model Solution

3.1 Known accuracy equilibrium

As a benchmark for the analysis to follow, we begin with the case in which the absolute and relative accuracy of the public information are known with certainty to be α and λ, respectively (hence the reference to the type "τ" is omitted). To simplify notation, denote $E_i[\cdot] \equiv E[\cdot|\Omega_i]$ as agent i's expectation of a random variable conditional on his information set, Ω_i. Given a disclosure policy DP and realization of private and public signals, each agent updates his beliefs about the state of nature and the behavior of other agents from his information set $\Omega_i \equiv \{s_i, z_i \mid \lambda, D\}$. Agent i's posterior belief about the state v is thus normally distributed with mean

$$E_i[v] = (\delta_i \gamma) z + (1 - \delta_i \gamma) s_i,$$

where $\gamma = \lambda / (1 + \lambda)$. Notice that an agent j who is outside the *inner circle* with $\delta_j = 0$ updates his belief based only on his private information.

Agent i chooses e_i to maximize $E_i [U_i^a]$. The resulting optimal action is given by,

$$e_i = E_i [A] = (1 - \rho) E_i [v] + \rho E_i [\bar{e}]$$

where $E_i [\bar{e}]$ is agent $i's$ expectation of the other agents' average action. As expected, e_i is increasing in the expected state and in the expected average action when $\rho > 0$ and actions are strategic complements. Otherwise, the optimal action is decreasing in the expected average action when $\rho < 0$ and actions are strategic substitutes. Given that the equation for e_i is linear in the updated beliefs about v which are normal, it's reasonable to search for an equilibrium among the set of effort supply functions that are linear in the private and public signals. Indeed, following Angeletos and Pavan (2007), one can easily show there is a unique effort supply equilibrium, and the equilibrium is linear in the private and public signals.

Proposition 1 *A unique effort supply equilibrium exists:*

$$e_i = \begin{cases} wz + (1-w)\, s_i & \text{for} \quad i \in D \\ s_i & \text{for} \quad i \notin D \end{cases}$$

where,

$$\begin{aligned} w &= \frac{\lambda}{\lambda + 1 - \rho\bar{\delta}} = \frac{\gamma}{1 - \rho\bar{\delta}\,(1-\gamma)} \\ \text{and } \gamma &\equiv \frac{\lambda}{\lambda+1}. \end{aligned} \tag{4}$$

The proof is similar to Morris and Shin (2002) and Angeletos and Pavan (2007) and is therefore omitted (and available from the authors upon request). Note that in equilibrium, the weight, w, that agents assign to the public signal in choosing effort is increasing in ρ and λ. The rationale for this is that the public signal is more influential to agents the greater is the complementarity in actions (higher ρ), or the more accurate it is relative to private information (higher λ). Agents wish to imitate one another more in high complementarity environments. By placing greater weight on the public signal, z, agents are able to coordinate their actions to achieve greater uniformity. In equilibrium, each agent expects the other agents to adopt this behavior, so it becomes self-fulfilling.

We are now ready to address the original question that motivated our analysis: *when is public information beneficial and how is it best disclosed?* For a given D with an inner circle of size $\bar{\delta} \in [0,1]$, the social surplus in equilibrium is given by:

$$\begin{aligned} E\left[U^P|v\right] &= \frac{v^2}{2} - \frac{M}{2} \\ \text{where } \ M &\equiv \frac{1}{\beta}\left[\frac{\bar{\delta}\left(1-2\rho\bar{\delta}\right)w^2}{\lambda} + \bar{\delta}\,(1-w)^2 + 1 - \bar{\delta}\right], \end{aligned}$$

and w is the weight agents place on the public signal.

The social planner's goal is to implement an information disclosure policy that minimizes the loss, M. Some loss is inevitable because agents do not fully account for the impact of their private actions on the welfare of other agents. This becomes apparent if we consider a hypothetical situation where the planner directs agents on how to weight public versus private information. In this case, she would select the *first best* weight, w^{FB}, to satisfy,

$$w^{FB} = \arg\min_w M = \frac{\lambda}{\lambda + 1 - 2\rho\bar{\delta}}. \tag{5}$$

Comparing (5) with (4),we have

$$\frac{w^{FB} - w}{w^{FB}} = \frac{\rho\bar{\delta}}{\lambda + 1 - \rho\bar{\delta}}. \tag{6}$$

Therefore, when $\rho > 0$, agents under-weight public information, even though they pay more attention to the public signal when their actions are strategic complements (i.e., the weight is higher than the weight they would place when $\rho = 0$). The opposite is true when $\rho < 0$; relative to the socially optimal level, agents over-weight public signal when their actions are strategic substitutes.

It is clear from (6) that more precise public information (i.e. a higher λ) moves w closer to w^{FB} in a relative sense and therefore could improve social welfare. This is further verified by[3]

$$\begin{aligned}
\frac{\partial E\left[U^P|v\right]}{\partial \lambda} &=_s \left[2\bar{\delta}\,(1-w) - \frac{2w\bar{\delta}\left(1-2\rho\bar{\delta}\right)}{\lambda}\right]\frac{\partial w}{\partial \lambda} + \frac{\bar{\delta}\left(1-2\rho\bar{\delta}\right)w^2}{\lambda^2} \\
&=_s 1 - \rho\delta + \lambda\,(1-2\rho\delta) > 0 \;\forall \rho < 1/2.
\end{aligned}$$

The intuition is that public information always helps agents select actions that match the underlying state of the world. Further, when agents wish to coordinate actions, either to complement others' behavior or to avoid interfering with other agents, the ability to forecast the mean action always improves welfare in our environment.

The arrival of more accurate public information is usually accompanied by increased demand for transparency. A more transparent policy informs a greater number of agents who then harness their knowledge to make better decisions – or so it would seem. But is transparency welfare increasing in all settings? Conditions for full disclosure to be optimal are readily derived and are recorded in the following:

Proposition 2 *Social welfare is maximized at full disclosure DP with* $\bar{\delta} = 1$ *provided that* $\rho \geq \underline{\rho}(\lambda)$ *where* $\underline{\rho}(\lambda) < 0$ *and* $\underline{\rho}'(\lambda) < 0$.

When $\rho > 0$, a higher $\bar{\delta}$ unambiguously increases social welfare for any λ. Increasing disclosure provides more agents with valuable public information to better match their

[3]The notation, $\stackrel{s}{=}$, signifies that two expressions have the same sign, positive or negative.

action to the underlying state. Further, it reduces the inefficiency embedded in agents' equilibrium weight, as

$$\frac{\partial\left(\frac{w^{FB}-w}{w^{FB}}\right)}{\partial\bar{\delta}} = \rho\frac{\lambda+1}{\left(\lambda+1-\rho\bar{\delta}\right)^2} > 0 \text{ iff } \rho > 0. \tag{7}$$

Since agents under-weight public information when $\rho > 0$, (7) implies that a more transparent disclosure policy (i.e. a larger $\bar{\delta}$) reduces inefficient weighting, thus improving welfare. However, in contrast, when $\rho < 0$, agents over-weight public information. Condition (7) then implies that larger $\bar{\delta}$ magnifies the weighting inefficiency more, the more negative ρ is. Hence, full disclosure is desirable only as long as the degree of strategic substitution is not too large. Proposition 2 further indicates that $\underline{\rho}'(\lambda) < 0$ which implies that full disclosure is warranted when the relative precision of public information is sufficiently high. In this case, the benefit of access to better information for personal decision making exceeds the cost of creating greater congestion between different agents.

In summary, Proposition 2 suggests that widespread public disclosure is more attractive the more accurate public information is known to be. In the sections to follow, we enquire if this is also true in settings where the accuracy of public information is not observed by the public.

3.2 Equilibrium with unknown accuracy: signaling through disclosure

Having access to more accurate information is beneficial, at least for individual agents. And, as the previous section illustrates, this conclusion even extends to social situations where individuals' decision may impact positively or negatively on each other. But is access to public information still individually and socially beneficial, when the accuracy of the information itself is not observed by all? The rest of our paper focuses on this question.

3.2.1 Agents' reaction to the disclosure of public information

Once the public information has been disclosed, agents update their beliefs about the planner's type. As a convention, we use $\widehat{}$ to denote any variable computed using agents' perception of the planner's type. Let $\widehat{q}\left(\bar{\delta}\right) \in [0,1]$ be the agent's posterior probability assessment that the planner is of type h after observing her disclosure policy. Upon observing $\bar{\delta}$ and their private and public signals (s_i, z_i^τ), agents i selects his action e_i given his information set $\Omega_i = \{s_i, z_i^\tau, \bar{\delta}\}$ to maximize his expected utility given by,

$$\max_{e_i} \quad \widehat{q}\left(\bar{\delta}\right) E_i \left[Ae_i - \frac{1}{2} e_i^2 | \lambda^h \right] + \left(1 - \widehat{q}\left(\bar{\delta}\right)\right) E_i \left[Ae_i - \frac{1}{2} e_i^2 | \lambda^l \right].$$

A unique solution to this maximization problem is summarized in the following lemma:

LEMMA 1: *A unique action supply equilibrium exists:*

$$e_i = \begin{cases} \hat{w}\left(\hat{q}\left(\bar{\delta}\right)\right) z^\tau + \left(1 - \hat{w}\left(\widehat{q}\left(\bar{\delta}\right)\right)\right) s_i & \text{for} \quad i \in D \\ s_i & \text{for} \quad i \notin D \end{cases},$$

where

$$\hat{w}\left(\hat{q}\left(\bar{\delta}\right)\right) \equiv \frac{\hat{\gamma}}{1 - \rho\bar{\delta}\left(1 - \hat{\gamma}\right)};$$

$$\text{and } \hat{\gamma} \equiv \hat{q}\left(\bar{\delta}\right) \frac{\lambda_h}{\lambda_h + 1} + \left(1 - \hat{q}\left(\bar{\delta}\right)\right) \frac{\lambda_l}{\lambda_l + 1}.$$

PROOF: The proof mirrors the analyses in Morris and Shin (2002) and Angeletos and Pavan (2007) and is therefore omitted.

Note that $\hat{w}\left(\widehat{q}\left(\bar{\delta}\right)\right)$ is similar to the weight obtained in the case of known accuracy except that $\gamma \equiv \frac{\lambda}{\lambda+1}$ in equation (4) is replaced by a weighted average of $\frac{\lambda_h}{\lambda_h+1}$ and $\frac{\lambda_l}{\lambda_l+1}$. In equilibrium, agent $i's$ effort is a weighted sum of his public and private signal. Agents who don't receive a public signal select effort based only on their private information. Otherwise, agents weight their private and public information according to their updated beliefs about the relative accuracy of the public disclosure based on the planner's disclosure policy, $\bar{\delta}$.

3.2.2 Planner's incentive to misrepresent accuracy

Although each type of the planners wishes to maximize the expected social surplus, planners with differing relative precision derive different surplus from the same disclosure policy. We use $EU^P\left[\lambda, \widehat{q}\left(\bar{\delta}\right), \bar{\delta}|v\right]$ to denote a planner's expected social surplus (conditional on v) when her accuracy type is λ and the agents perceive she is of the high type with probability

$\hat{q}\left(\bar{\delta}\right)$ based on her disclosure policy $\bar{\delta}$. Specifically,

$$EU^P\left[\lambda,\widehat{q}\left(\bar{\delta}\right),\bar{\delta}|v\right] \equiv \frac{v^2}{2}-\frac{M\left(\lambda,\hat{q}\left(\bar{\delta}\right),\bar{\delta}\right)}{2}$$

$$\text{where } M\left(\lambda,\hat{q}\left(\bar{\delta}\right),\bar{\delta}\right) = \frac{1}{\beta}\left[\frac{\bar{\delta}\left(1-2\rho\bar{\delta}\right)\hat{w}\left(\widehat{q}\left(\bar{\delta}\right)\right)^2}{\lambda}+\bar{\delta}(1-\hat{w}\left(\widehat{q}\left(\bar{\delta}\right)\right))^2+1-\bar{\delta}\right]$$

It's clear from inspection that the social surplus depends on the agent's perception, $\hat{q}\left(\bar{\delta}\right)$, of the precision of the public information, given $\bar{\delta}$. The planner may try to persuade agents of the precision of the public information by the disclosure policy she selects. Therefore, to predict the planner's choice of disclosure policy, it is useful to know what the planner *would like* the agents to believe about the accuracy of the public disclosure.

LEMMA 2: *When $\rho > 0$, the high type does not benefit from misrepresenting her true accuracy, while the low type strcitly benefits from misrepresenting her true accuracy provided $\frac{\lambda_l}{\lambda_h} \geq \frac{1-2\rho\bar{\delta}}{1-\rho\bar{\delta}}$. When $\rho < 0$, the low type does not benefit from misrepresenting her true accuracy, while the high type strictly benefits from misrespresenting her true accuracy provided $\frac{\lambda_l}{\lambda_h} \geq \frac{1-\rho\bar{\delta}}{1-2\rho\bar{\delta}}$.*

Planners are often *parental* in their disclosure of information and advice, hoping to guide the agents they oversee to make the best decision for their own collective good. Our planner is no exception in this regard. When the agent's individual actions are strategic complements, the planner wishes the agents to better coordinate their actions to maximize the aggregate surplus. If the agents believe that the planner's information is highly accurate, they would pay more attention to the public signal when choosing actions. This results in a higher correlation of independent actions leading to greater social surplus. In contrast, when agents' actions are strategic substitutes, the planner wishes to persuade the agents that her disclosure is of relatively low accuracy. The rationale here is that agents are induced to differentiate their action choices if they believe the public signal is less accurate, and they pay less attention to it in selecting actions. As a result, there is less correlation in individual action choices which leads to greater social surplus.

3.2.3 Planner's preferences for signaling through disclosure

Agents' perception of the relative accuracy of the public disclosure is informed by observing the disclosure policy that the planner selects. This process enables the planners of different types to signal the accuracy of their information through disclosure policy and thereby direct the agents to act in a certain way. For instance, if agents perceive that only low type planners follow an opaque disclosure policy by rationing public information, they will pay less attention to the public disclosure believing that it is relatively inaccurate. However, in order for the agents to predict the disclosure policies that different types of planners employ, it is most useful to determine which *types* of planners benefit the most from signaling their true types. This raises the issue of how a particular type of planner wishes to signal her type via her disclosure policy so that agents can best utilize the public information that the planner discloses and leads to the following lemma.

LEMMA 3: *Consider any two disclosure policies* $\tilde{\delta}_1, \tilde{\delta}_2$ with $0 < \tilde{\delta}_2 < \tilde{\delta}_1 \leq 1$.

When $\rho > 0$, *in any equilibrium where planner* l *weakly prefers* $\tilde{\delta}_2$ to $\tilde{\delta}_1$, *planner* h *strictly prefers* $\tilde{\delta}_2$ to $\tilde{\delta}_1$.

When $\rho < 0$, in any *equilibrium where planner* h *weakly* prefers $\tilde{\delta}_2$ to $\tilde{\delta}_1$, *planner* l *strictly prefers* $\tilde{\delta}_2$ to $\tilde{\delta}_1$.

PROOF: Differentiating the planner's expected surplus, we compute

$$
\begin{aligned}
\frac{d}{d\lambda}\left[\left.\left(\frac{d\widehat{q}}{d\delta}\right)\right|_{dEU^P\left[\lambda,\widehat{q},\overline{\delta}|v\right]=0}\right] &= \frac{d}{d\lambda}\left(-\frac{\frac{\partial EU^P\left[\lambda,\widehat{q},\overline{\delta}|v\right]}{\partial\overline{\delta}}}{\frac{\partial EU^P\left[\lambda,\widehat{q},\overline{\delta}|v\right]}{\partial\widehat{q}}}\right) \\
&= {}^s\frac{d}{d\lambda}\left\{-\frac{\lambda-\lambda\left(1-\widehat{w}\right)^2+\left(4\rho\overline{\delta}-1\right)\widehat{w}^2}{2\left[\left(2\rho\overline{\delta}-1\right)\widehat{w}+\lambda(1-\widehat{w})\right]}\right\} \\
&= {}^s-\left[1-\left(1-\widehat{w}\right)^2\right]\left[\left(2\rho\overline{\delta}-1\right)\widehat{w}+\lambda(1-\widehat{w})\right] \\
&\quad +(1-\widehat{w})\left[\lambda-\lambda\left(1-\widehat{w}\right)^2+\left(4\rho\overline{\delta}-1\right)\widehat{w}^2\right] \\
&= {}^s 1-2\rho\overline{\delta}\widehat{w} > 0. \qquad (8)
\end{aligned}
$$

Suppose $\rho > 0$. Pick two points, $\left(\tilde{\delta}_1, \hat{q}\left(\tilde{\delta}_1\right)\right), \left(\tilde{\delta}_2, \hat{q}\left(\tilde{\delta}_1\right)\right)$ with $\tilde{\delta}_1 > \tilde{\delta}_2$ such that

$$EU^P\left[\lambda^l, \hat{q}\left(\tilde{\delta}_2\right), \tilde{\delta}_2|v\right] - EU^P\left[\lambda^l, \hat{q}\left(\tilde{\delta}_1\right), \tilde{\delta}_1|v\right] = 0$$

Or, equivalently,

$$\int_{\tilde{\delta}_2}^{\tilde{\delta}_1} \left(EU_{\bar{\delta}}^P \left(\lambda^l\right) + EU_{\widehat{q}}^P \left(\lambda^l\right) \left(\frac{d\widehat{q}}{d\bar{\delta}}\right)\Bigg|_{dEU^P\left[\lambda^l,\widehat{q},\bar{\delta}|v\right]=0} \right) d\bar{\delta} = 0,$$

where $EU_{\bar{\delta}}^P\left(\lambda^l\right) \equiv \frac{\partial EU^P\left[\lambda^l,\widehat{q},\tilde{\delta}|v\right]}{\partial \bar{\delta}}$, and $EU_{\widehat{q}}^P\left(\lambda^l\right) \equiv \frac{\partial EU^P\left[\lambda^l,\widehat{q},\tilde{\delta}|v\right]}{\partial \widehat{q}}$. Note (8) implies that

$$\int_{\tilde{\delta}_2}^{\tilde{\delta}_1} \left(EU_{\bar{\delta}}^P \left(\lambda^h\right) + EU_{\widehat{q}}^P \left(\lambda^h\right) \left(\frac{d\widehat{q}}{d\bar{\delta}}\right)\Bigg|_{dEU^P\left[\lambda^l,\widehat{q},\bar{\delta}|v\right]=0} \right) d\bar{\delta} > 0,$$

Then,

$$\begin{aligned} & EU^P\left[\lambda^h, \widehat{q}\left(\tilde{\delta}_2\right), \tilde{\delta}_2|v\right] - EU^P\left[\lambda^h, \widehat{q}\left(\tilde{\delta}_1\right), \tilde{\delta}_1|v\right] \\ = & \int_{\tilde{\delta}_2}^{\tilde{\delta}_1} \left[EU_{\bar{\delta}}^P \left(\lambda^h\right) + EU_{\widehat{q}}^P \left(\lambda^h\right) \left(\frac{d\widehat{q}}{d\bar{\delta}}\right)\Bigg|_{dEU^P\left[\lambda^l,\widehat{q},\bar{\delta}|v\right]=0} \right] d\bar{\delta} \\ > & \ 0. \end{aligned}$$

This completes our proof for the $\rho > 0$ case. A similar argument establishes the ordering of preferences for the $\rho < 0$ case. ■

The implications of Lemma 3 for the complements case of $\rho > 0$ (that is, more coordination in actions is valued) are illustrated in Figure 1. The planners h and l have indifference curves denoting the combinations of $\left(\hat{q}\left(\bar{\delta}\right), \bar{\delta}\right)$ that yield constant surplus, which are shown respectively as I^h and I^l in Figure 1. The I^h curve intersects the I^l curve from below, indicating that h is more willing than l to reduce the size of the inner circle agents, $\bar{\delta}$, on the margin to gain a higher perception of accuracy, $\hat{q}$. The preferred-to-set for type h consists of the shaded area above the I^h curve. This implies whenever l weakly prefers allocation B to A, h strictly prefers B to A. In particular, h is more eager to reduce $\bar{\delta}$ to signal her accuracy than l as h incurs less costs from reducing disclosure than l does. Intuitively, because h's signal is more accurate about the state of nature, she is more confident that the actions taken by the inner circle agents who utilize the disclosed public signal would not be too far apart from the actions taken by the outer circle agents who only employ their private information that is also informative about the state. In other words, h is able to obtain a greater degree of coordination between the inner circle and outer circle for the same degree of information rationing. To see this, take an inner circle agent i and an outer circle agent j. The expected distance between these two agents' actions is given by $E\left[\hat{w}z + (1-\hat{w})\, s_i - s_j\right]^2 = \frac{\hat{w}^2}{\alpha} + \frac{(1-\hat{w})^2}{\beta} + \frac{1}{\beta}$. This distance is larger for the low type planner than for the high type planner, as

$$\frac{E\left[\left[\hat{w}z + (1-\hat{w})\, s_i - s_j\right]^2 |\alpha = \alpha_h\right]}{E\left[\left[\hat{w}z + (1-\hat{w})\, s_i - s_j\right]^2 |\alpha = \alpha_l\right]} = \frac{\frac{\hat{w}^2}{\alpha_h} + \frac{(1-\hat{w})^2}{\beta} + \frac{1}{\beta}}{\frac{\hat{w}^2}{\alpha_l} + \frac{(1-\hat{w})^2}{\beta} + \frac{1}{\beta}} < 1. \tag{9}$$

This means that information rationing creates more coordination (between the two circles) for the high type planner than for the mimicing low type planner, which provides an advantage for the high type to separate by rationing information.

The relative incentives for h and l to disclose are reversed in the strategic substitutes case (i.e. more dispersion in actions is valued), as depicted in Figure 2. Here the indifference curves I^h and I^l slope up with I^h intersecting I^l from below. The preferred-to-set for l lies in the shaded area below the I^l curve. In this instance, it's clear that if h weakly prefers an allocation B to A, then l strictly prefers B to A. This indicates that it is the l type who is more willing to restrict disclosure to persuade agents that her information is inaccurate. The intuition again can be gleaned from (9) which shows that for the the low type planner the actions between the two circles are more dispersed than for the mimicing high type, providing an advantage for the low type to separate by rationing information.

The aforementioned prefence ordering properties implied by Lemma 3 are all that we require to characterize "reasonable equilibria" for our signaling game. By "reasonable equilibria", we mean those that are supported by the Cho and Kreps' Intuitive Criterion that is stated below.

Cho and Kreps' Intuitive Criterion *Consider a type $j \in \{h, l\}$ who makes an out of equilibrium selection of $\tilde{\delta}$* and *her type is correctly perceived. Then if no other type $j' \in \{h, l\}$ is better off mimicing type j, the perception of the agents is "credible".*

The Intuitive Criterion requires that out of equilibrium disclosures are supported by "reasonable beliefs" about the type of planner who would have found it profitable to deviate from the expected equilibrium play.

First, one can easily establish that the condition in Proposition 2 guarantees that $\frac{\partial EU^P\left[\lambda, \widehat{q}, \bar{\delta}|v\right]}{\partial \bar{\delta}} > 0$, $\forall \lambda$. That is, increased disclosure of public information (i.e. a higher $\bar{\delta}$) is beneficial for all types of planner, holding agents' perception $\widehat{q}$ constant. Now, Let's consider the strategic complements case depicted in Figure 1. Lemma 3 implies there is a Pareto dominant separating equilibrium with corresponding strategies and beliefs represented by $A = \left(\bar{\delta}^l = 1, \hat{q}\left(\bar{\delta}^l\right) = 0\right)$ and $B = \left(\bar{\delta}^h = \bar{\delta}^B, \hat{q}\left(\bar{\delta}^B\right) = 1\right)$. In this equilibrium, l selects full disclosure, $\bar{\delta}^l = 1$, and h chooses partial disclosure, $\bar{\delta}^h = \bar{\delta}^B < 1$ in order to signal her type. This separating equilibrium leaves l indifferent between A and B and it therefore minimizes $h's$ signaling cost. By the ordering property of Lemma 3, we know this is the only separating equilibrium that satisfies the Intuitive Criterion. To see why, suppose there is another separating equilibrium in which l selects A and h selects $C = \left(\bar{\delta}^h = \delta^C, \hat{q}\left(\bar{\delta}^C\right) = 1\right)$ as depicted in Figure 1 . This equilibrium must be supported by the belief that any disclosure $\bar{\delta}^h > \bar{\delta}^C$ is not made by an h type for sure, otherwise h would deviate. However, if h were to select another disclosure such as $\bar{\delta}^h = \bar{\delta}^B$ and her type were correctly perceived, there would be no incentive for l to mimic type h, since l could not strictly benefit by doing so. Hence, the belief that the deviating type cannot be h for sure violates the Intuitive Criterion, thus eliminating the proposed separating equilibrium.

The same argument establishes that there are no pooling or semi-pooling equilrium that satisfies the Intuitive Criterion. To illustrate, suppose there is a pooling equilibrium, denoted by D in Figure 1. This equilibrium is supported by the belief that any disclosure $\bar{\delta} > \bar{\delta}^D$ cannot be made by a h type for sure, otherwise h would deviate. However, if h were to select $\bar{\delta}^h = \bar{\delta}^B$ and her type were correctly perceived, there would be no incentive for l to mimic h, since she could not strictly benefit by doing so. Hence again, the belief that the deviating type can not be h for sure, violates the Intuitive Criterion, thus eliminating the proposed pooling equilibrium.

This completes our construction of the unique separating equilibrium for the $\rho > 0$ case. A similar argument can be used to construct the unique separating equilibrium corresponding to the $\rho < 0$ case. Hence we have established Proposition 3 below.

Proposition 3 (i) *Suppose* $\rho > 0$. *There is a unique signaling equilibrium satisfying the Intuitive Criterion where the high type selects* $\bar{\delta}^h < 1$ *and the low type selects* $\bar{\delta}^l = 1$.

(ii) *Suppose* $\rho < 0$. *There is a unique signaling equilibrium satisfying the Intuitive Criterion where the low type selects* $\bar{\delta}^l < 1$ *and* the high type *selects* $\bar{\delta}^h = 1$.

4 Extension and Implication

4.1 Heterogeneous agents

We now relax one of the assumptions in the basic setup and assume that the precision of the agents' private signal varies. Specifically, we index agent i's private precision with $\beta(i) = \exp(\beta_0 + ci)$, where $\beta_0 > 0$, $c \geq 0$, and $i \in [0,1]$. Consistent with the setup for homogenous agents, we assume that while agents do not know β_0, they know parameter c and their own index i and understand that they face a planner with two possible levels of precision: $\alpha^h \equiv \exp\left(\beta_0 + b^h\right)$ and $\alpha^l \equiv \exp\left(\beta_0 + b^l\right)$, with $b^h > b^l > 1$, and $\Pr(\alpha = \alpha^h) = q$ and $\Pr(\alpha = \alpha^l) = 1 - q$. While the agents do not observe the absolute levels of α^h or α^l, they know b^h, b^l and the ratio of $\alpha^h/\alpha^l = \exp\left(b^h - b^l\right)$. The social planner privately observes α. The parameter c, index i and the agents' relative precision $\lambda_i \equiv \exp(b - ci)$ (for both α) are public information, implying that the principal privately knows $\beta(i)$. As with the homogeneous agent case, we assume that $\lambda_i > 1$, $\forall i$. The following lemma characterizes the agents' unique equilibrium action supply in the case of heterogeneous agents.

Lemma 4: *A unique effort supply equilibrium exists:*

$$e_i = \begin{cases} s_i & \text{for} \quad i \notin D; \\ e_i = \hat{w}_i z + (1 - \hat{w}_i)\, s_i & \text{for} \quad i \in D, \end{cases}$$

where,

$$\hat{w}_i = 1 - \frac{1 - \rho\bar{\delta}}{1 - \rho\hat{L}}\hat{L}_i;$$

$$\text{with } \hat{L}_i \equiv \hat{q}\left(\bar{\delta}\right)\frac{1}{\lambda_{h,i} + 1} + \left(1 - \hat{q}\left(\bar{\delta}\right)\right)\frac{1}{\lambda_{l,i} + 1};\ \hat{L} \equiv \int_D \hat{L}_i di.$$

When agents differ in their private precision, their weights on the public signal are not only a function of the measure of the inner circle ($\bar{\delta}$) but also the composition of the inner circle, as captured by the $\hat{L}$ term.

Clearly, when c is zero, we are back to the homogeneous case. Given that the planner and the agents' expected utilities and their corresponding derivatives are continuous in

c, the results established in the previous section carry through to the heterogeneous case provided c is small, among them is the result that information rationing continues to be the only equilibrium that survives the Cho and Krep's Intuitive Criterion. Thus, the remaining question is the choice of optimal inner circle agents to achieve a credible separation among different types of planner. The following Proposition establishes this result.

Proposition 4 *(1) When $\rho < 0$, the optimal separating arrangement is characterized by $D^h = [0, 1]$ and the unique $D^l = \left[0, \bar{\delta}\right] \subset [0, 1]$ such that the high type is indifferent between mimicing and truth telling.*

(2) *When $\rho > 0$*, there exists a $\underline{\rho}$ and $\underline{\lambda}$ such that when $\rho \geq \underline{\rho}$ and $\frac{\alpha_h}{\beta(1)} \geq \underline{\lambda}$, and when α^h/α^l is not too large, *the optimal separating arrangement is characterized by $D^l = [0, 1]$ and the unique $D^h = \left[0, \bar{\delta}\right] \subset [0, 1]$ such that the low type is indifferent between mimicing and truth telling.*

Proposition 4 sheds light on the characteristics of the inner circle and shows that the optimal inner circle is those agents with less precise private information. The intuition is closely linked to the fact that with heterogeneous agents (9) becomes

$$\frac{E\left[\left[\hat{w}_i z + (1 - \hat{w}_i)\, s_i - s_j\right]^2 | \alpha = \alpha_h\right]}{E\left[\left[\hat{w}_i z + (1 - \hat{w}_i)\, s_i - s_j\right]^2 | \alpha = \alpha_l\right]} = \frac{\frac{\hat{w}_i^2}{\alpha_h} + \frac{(1-\hat{w}_i)^2}{\beta(i)} + \frac{1}{\beta(j)}}{\frac{\hat{w}_i^2}{\alpha_l} + \frac{(1-\hat{w}_i)^2}{\beta(i)} + \frac{1}{\beta(j)}} < 1 \tag{10}$$

and is clearly increasing in $\beta(j)$. This implies when $\rho < 0$ (that is, the planner prefers more dispersion in agent actions), the low type can magnify her advantage in separation if the inner (outer) circle consists of agents with less (more) precise private information. The intuition is the following. Two forces are at play when a planner rations information. On the one hand, roughly speaking, an inncer circle that consists of agents with less (more) accurate private information attaches larger (smaller) weights to public disclosure. As a result, if the high type planner mimics the low type, actions by a less privately informed inner circle are influenced to a greater extent by the disclosed signal than actions by a more privately informed inner circle. And by definition, the high type planner expects her signal to be closer to the state of nature than the low type planner. On the other hand, when the outer circle contains more (less) privately informed agents, actions in the outer circle are closer to (further away from) the state of nature. With these two forces reinforcing each other, the mimicing high type planner expects less dispersion (compared to the low type) between the inner circle actions and outer circle ones when the inner circle consists of less privated informed agents. Since divergence in actions between the two circles is preferred when $\rho < 0$, the low type planner's advantage in information rationing is strengthened and separation is better achieved.

The intuition for the case of $\rho > 0$ is similar but works in the opposite direction. Here, the planner values more coordination. Therefore, conditioning on information rationing (which hurts coordination to begin with), it is the high type planner who has the advantage in making the two groups (inner and outer) of agents' actions closer, as expression (10) shows. Further, this advantage is magnified when the inner circle consists of agents with less precise private information, for reasons similar to what is discussed above. Since convergence in actions between the two circles is preferred when $\rho > 0$, the high type planner's advantage in information rationing is strengthened and separation is better achieved.

The additional conditions needed for the case of $\rho > 0$ are due to the fact that after all, high quality public signal is withheld from some agents and the planner does care about "Better Action" to begin with. Thus, information rationing is optimal only when the concern for coordination is reasonably large (i.e., ρ relatively large). In the case of $\rho < 0$, the planner also cares about "Better Action" but it is the low quality information that is being withheld, so the concern is not as overwhelming.

4.2 Incentive for ex-ante investment to acquire information

Having established the use of disclosure policy as a separation device, we now turn to the issue of a planner's ex-ante incentives for information acquisition. For simplicty, let's return to our basic setup of homogenous agents. Suppose now that before knowing her own type, the planner could incur a cost of $g(q)$ to receive high quality information with probability q, where $g' > 0$, $g'' > 0$, $\lim_{q\to 1} g' = \infty$, and $\lim_{q\to 0} g' = 0$. The following proposition compares the planner's information acquisition incentives under separation when type is unknown to agents with when type is public knowledge.

Proposition 5 *Compared to a situation where the social planner's type is publicly known: when $\rho > 0$, the planner under-invests to acquire information; when $\rho < 0$, the social planner over-invests to acquire information.*

We start with when $\lambda' s$ are publicly known (hence $\bar{\delta} = 1$). Recall $EU^p\left[\lambda, \widehat{q}, \bar{\delta}|v\right]$ is the planner's expected utility when her true type is λ, the disclosure set is $\bar{\delta}$, and agents perceive her as the high type with probability $\hat{q}$. The planner chooses q to maximize the expected value of the objective function:

$$\max_q \; qEU^p\left[\lambda_h, 1, 1|v\right] + (1-q)\, EU^p\left[\lambda_l, 0, 1|v\right] - g\left(q\right)$$

The first best level of q^{FB} is given by

$$g'\left(q^{FB}\right) = EU^p\left[\lambda_h, 1, 1|v\right] - EU^p\left[\lambda_l, 0, 1|v\right].$$

When $\lambda' s$ are unknown, the social planner's expected utility depends on which type rations information in equilibrium. With strategic complementarity (i.e., $\rho > 0$), it's the high type that rations the information. Thus, the social planner chooses q to maximize:

$$\max_q \; qEU^p\left[\lambda_h, 1, \bar{\delta} < 1|v\right] + (1-q)\, EU^p\left[\lambda_l, 0, 1|v\right] - g\left(q\right).$$

The optimal $q^{\rho>0}$ is given by

$$g'\left(q^{\rho>0}\right) = EU^p\left[\lambda_h, 1, \bar{\delta} < 1|v\right] - EU^p\left[\lambda_l, 0, 1|v\right].$$

Because, $EU^p\left[\lambda_h, 1, 1|v\right] > EU^p\left[\lambda_h, 1, \bar{\delta} < 1|v\right]$, it is easy to see that $q^{\rho>0} < q^{FB}$, that is, the planner under-invests in acquiring information. This is because in order to separate from the low type, the high type will end up restricting some agents' access to her information, reducing the value of acquiring the information to begin with. As a result, the planner invests less than the first best level.

When $\rho < 0$, it is the low type that rations information. The planner's optimal ex-ante investment in q is given by:

$$g'\left(q^{\rho<0}\right) = EU^p\left[\lambda_h, 1, 1|v\right] - EU^p\left[\lambda_l, 0, \bar{\delta} < 1|v\right].$$

Since $EU^p\left[\lambda_l, 0, 1|v\right] > EU^p\left[\lambda_l, 0, \bar{\delta} < 1|v\right]$, we have $q^{\rho<0} > q^{FB}$, i.e., the planner over-invests in acquiring information.

On a casual glimpse, this result seems counter-intuitive in that while ex post after the type is set, the high type planner would like the agents to believe that she is of the lower type, ex ante the planner actually over-invests to become the high type. This is because it is the low type that is forced to partially disclose despite the fact that full disclosure is the first best. Compared to the earlier case where the incentive to under-invest is due to the dampened reward of becoming a high type, the incentive to over-invest in this case is due to the heightened punishment for ending up as a low type.

4.3 Ex ante restrictions on informed princinpal: designing the constitution

The need to deny some agents' access to public information in an separating equilibrium may provide a justification for a commitment to requiring that all information be disclosed to all agents. Again, for simplicity, let's return to our basic model of homogeneous agents. Suppose before the planner is privately informed of her precision, the agents are able to collectively decide on a constitution that governs the planner's subsequent disclosure behavior. For tractability, let's focus on two "simple" constitutions: one precludes selectively disclosure by the planner (and hence leads to a pooling equilibrium), the other allows (and thus leads to a separating equilibrium).[4] The key question we are interested in here is which of the two constitutions would generate a higher ex ante social surplus.

One the one hand, in the best separating equilibrium that we have been studying so far, the agents are able to perfectly infer the planner's type/precision by looking at her disclosure behavior. But this comes at a cost, as in order to credibly convey her precision, one type of the planners needs to withhold her disclosure from a subset of agents. On the other hand, in a pooling equilibrium with full disclosure, though all agents get to observe the planner's signal, they are not able to perfectly infer the planner's precision and thus can only attach a less efficient "average" weight.

Given that an analytical solution for the expected social surplus under separation is difficult to obtain, we therefore resort to numerical examples to illustrate our observations. Let's start with the strategic complementarity case (i.e. $\rho > 0$). The first issue is how well separation performs relative to pooling as the probability of the high precision planner q varies. Set $\rho = 0.15$, $\lambda_l = 3$, and $\lambda_h = 2$. Straightforward algebra show that separation dominates pooling in terms of ex ante surplus if and only if $q \leq 0.99052$. This is not surprising. Intuitively, when q is sufficiently high, the pooling weight would be quite close

[4] We only consider the best separating equilibrium where one type of the planner fully discloses while the other rations information; and the best pooling equilibrium where disclosure is made available to every agent.

to the separating weight for the high type (that is, the incremental benefit under separation is low), while the expected probability of information rationing under separation is high (that is, the incremental cost of separation is substantial).

Our second result is on the effect of differences between λ_h and λ_l. Figure 3 plots indifference curves (along which pooling and separation generate the same expected surplus) under three sets of parameter combinations in the q-ρ space. Starting with the combination $\lambda_h = 3$ and $\lambda_l = 2$, the area enclosed by the solid line is those ρ-q values under which separation dominates pooling. Now, if we decrease the absolute difference between λ_h and λ_l by holding λ_l constant but decreasing λ_h to 2.8, the indifference curve moves to the dashed line. Clearly, now the area in which separation dominates pooling shrinks. The same can also be observed when we increase the relative difference between λ_h and λ_l (i.e. $\frac{\lambda_h - \lambda_l}{\lambda_l}$) while holding their absolute difference constant. The dotted line denotes the indifference curve with $\lambda_h = 4$ and $\lambda_l = 3$. Again, the area in which separation dominates pooling shrinks. Intuitively, when the difference between the two λ's becomes bigger, achieving separation necessitates withholding disclosure from a larger set of agents, hence making separation less appealing compared to pooling.

The case of strategic substitute ($\rho < 0$) is similar and illustrated in Figure 4. Our results in this section imply that forcing the government to be transparent, i.e. to make disclosures to every member of the society, can be counterproductive, as doing so could deprive a communication channel through which precisions of such disclosures can be conveyed.

5 Conclusion

We model a setting where a social planner who seeks to coordinate actions taken by a continuum of agents is privately informed of an information signal as well as its precision and strategically discloses her signal to the agents. Since coordination is better achieved by inducing agents to overweight or underweight the planner's disclosure, a planner may have incentives to misrepresent her signal precision. To create separation, we show that, when agents' actions are strategic complements (substitutes), the high (low) precision planner optimally restricts disclosure to a subset of the agents, i.e. rationing information by making high (low) quality information only available to a selected inner circle. Furthermore, when agents are heterogeneous in terms of the precision of their own private signals, the optimal separation between different types of planners can be achieved when the inner circle consists of those with more inaccurate private signals and thus are most influenced by the principal's information.

References

[1] Alonso, Richardo, Wouter Dessein and Niko Matouschek, 2008, When does coordination require centralization? *American Economic Review* 98, 145-79.

[2] Angeletos, George-Marios, and Alessandro Pavan, 2004, Transparency of information and coordination in economies with investment complementarities, *American Economic Review* 94, 91-98.

[3] Angeletos, George-Marios, and Alessandro Pavan, 2007, Efficient use of information and social value of information, *Econometrica* 75, 1103-1142.

[4] Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, 2003, Regulation for conservatives: Behavioral economics and the case for "asymmetric paternalism." *University of Pennsylvania Law Review* 1151, 1211-1254.

[5] Hellwig, Christian, 2005, Heterogeneous information and the benefits of transparency," UCLA Mimeo.

[6] Hirshleifer, David and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337-386.

[7] Hirshleifer, David, Sonya Lim, and Siew Hong Teoh, 2004, Disclosure to an audience with limited attention, Dice center working paper series 2004-21, Fisher College of Business, Ohio State University

[8] Morris, Stephen, and Hyun Song Shin, 2002, The social value of public information, *American Economic Review* 92, 1521-1534.

[9] Rantakari, Heikki, 2008, Governing adaptation, *Review of Economic Studies* 75, 1257-1285.

[10] Sunstein, Cass R., and Richard Thaler, 2003, Libertarian paternalism is not an oxymoron, *The University of Chicago Law Review* 70, 1159-1202.

[11] Svensson, Lars, 2006, Social value of public information: Morris and Shin (2002) is actually pro transparency, not con, *American Economic Review* 96, 448-451.

6 Appendix

Proof of Proposition 2:

Express the social welfare function as

$$E\left[U^P|v\right] = \frac{v^2}{2} + \frac{1}{2}\left[\frac{\bar{\delta}-1}{\beta} - \frac{\bar{\delta}\widehat{w}^2}{\alpha} - \frac{\bar{\delta}\left(1-\widehat{w}\right)^2}{\beta}\right] + \frac{\rho\bar{\delta}^2\widehat{w}^2}{\alpha}.$$

Further $\frac{\partial\widehat{w}}{\partial\bar{\delta}} = \frac{\rho\hat{\lambda}}{\left(1+\hat{\lambda}-\rho\bar{\delta}\right)^2} = \frac{\rho\widehat{w}}{1+\hat{\lambda}-\rho\bar{\delta}}$. We use $\widehat{}$ to denote the agent's perception. Hence

$$\begin{aligned}
\frac{\partial E\left[U^P|v\right]}{\partial\bar{\delta}} &= \frac{1}{2}\left[\frac{1-\left(1-\widehat{w}\right)^2}{\beta} - \frac{\widehat{w}^2}{\alpha}\right] + \frac{2\rho\bar{\delta}\widehat{w}^2}{\alpha} + \frac{2\rho\bar{\delta}^2\widehat{w}}{\alpha}\frac{\partial\widehat{w}}{\partial\bar{\delta}} + \left[\frac{\bar{\delta}\left(1-\widehat{w}\right)}{\beta} - \frac{\bar{\delta}\widehat{w}}{\alpha}\right]\frac{\partial\widehat{w}}{\partial\bar{\delta}} \\
&= \frac{1-\left(1-\widehat{w}\right)^2}{2\beta} - \frac{\widehat{w}^2}{2\alpha} + \frac{2\rho\bar{\delta}\widehat{w}^2}{\alpha}\left(1+\frac{\bar{\delta}\rho}{1+\hat{\lambda}-\rho\bar{\delta}}\right) + \frac{\bar{\delta}\rho\widehat{w}}{1+\hat{\lambda}-\rho\bar{\delta}}\left[\frac{1-\widehat{w}}{\beta} - \frac{\widehat{w}}{\alpha}\right] \\
&= \frac{\widehat{w}}{\alpha}\left\{\frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2} + 2\rho\bar{\delta}\widehat{w}\left(\frac{1+\hat{\lambda}}{1+\hat{\lambda}-\rho\bar{\delta}}\right) + \frac{\bar{\delta}\rho}{1+\hat{\lambda}-\rho\bar{\delta}}\left[\left(1-\widehat{w}\right)\lambda-\widehat{w}\right]\right\} \\
&= \frac{\widehat{w}}{\alpha}\left\{\underbrace{\frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2} + \frac{\bar{\delta}\rho}{1+\hat{\lambda}-\rho\bar{\delta}}\left[2\widehat{w}\left(1+\hat{\lambda}\right)+\left(1-\widehat{w}\right)\lambda-\widehat{w}\right]}_{*}\right\}
\end{aligned}$$

Easy to verify that $2\widehat{w}\left(1+\hat{\lambda}\right)+\left(1-\widehat{w}\right)\lambda-\widehat{w}>0$, so the above is positive for sure when $\rho>0$.

When $\rho<0$, the first term is strictly positive while the second term is monotonically increasing in ρ but equals to zero when $\rho=0$. Thus as long as ρ is not too negative, $\frac{\partial E\left[U^P|v\right]}{\partial\bar{\delta}}>0$. More precise bound can be obtained by further simplifying the terms in the curly brackets as

$$\begin{aligned}
* &=_s \frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2}\left(1+\hat{\lambda}-\rho\delta\right) + \bar{\delta}\rho\left[2\widehat{w}\left(1+\hat{\lambda}\right)+\left(1-\widehat{w}\right)\lambda-\widehat{w}\right] \\
&= \frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2}\left(1+\hat{\lambda}\right) + \bar{\delta}\rho\left[2\widehat{w}\left(1+\hat{\lambda}\right)+\left(1-\widehat{w}\right)\lambda-\widehat{w}-\frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2}\right] \\
&=_s \frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2} + \bar{\delta}\rho\widehat{w}\left[2-\frac{1+\lambda}{2\left(1+\hat{\lambda}\right)}\right] \\
&> \frac{\lambda\left(2-\widehat{w}\right)-\widehat{w}}{2} + 2\rho
\end{aligned}$$

where the last inequality follows because $\widehat{w}\left[2-\frac{1+\lambda}{2\left(1+\hat{\lambda}\right)}\right]$ is bounded above by 2. Thus, as long as

$$\rho > -\frac{1}{2}\frac{\left(\lambda\left(2-\widehat{w}\right)-\widehat{w}\right)}{2}$$

we are set. A sufficient condition is

$$\rho > -\frac{2\lambda-(1+\lambda)}{4} = -\frac{\lambda-1}{4}.$$

When λ is known, replace $\hat{\lambda}$ with λ before differentiating $E\left[U^P|v\right]$ with respect to $\bar{\delta}$, we obtain

$$\frac{dE\left[U^P|v\right]}{d\bar{\bar{\delta}}} =_s w^2 + 2w\bar{\bar{\delta}}\frac{dw}{d\bar{\bar{\delta}}} =_s 1 + \frac{2\rho\bar{\delta}}{\lambda+1-\rho\bar{\bar{\delta}}} =_s 1+\lambda+\rho\bar{\bar{\delta}} > 0$$

for all $\rho > -1-\lambda$. ■

Proof of Lemma 2:

First, observe

$$\frac{dEU^P}{d\hat{q}} =_s \left[\lambda\left(1-\hat{w}\right)-\left(1-2\rho\delta\right)\hat{w}\right]\frac{d\widehat{w}}{d\widehat{q}},$$

where $\hat{w}\left(\hat{q}\left(\bar{\delta}\right)\right) = \frac{\hat{\gamma}}{1-\rho\bar{\delta}(1-\bar{\gamma})}$ and $\hat{\gamma} = \hat{q}\left(\bar{\delta}\right)\frac{\lambda_h}{\lambda_h+1} + \left(1-\hat{q}\left(\bar{\delta}\right)\right)\frac{\lambda_l}{\lambda_l+1}$. Since $\frac{d\widehat{w}}{d\widehat{q}} > 0$, $\frac{dEU^P}{d\hat{q}} > 0$ if and only if $\lambda\left(1-\hat{w}\right)-\left(1-2\rho\delta\right)\hat{w} > 0$. Also note that $\forall\hat{q}\in(0,1)$,

$$\begin{aligned}
\frac{\lambda_h}{\lambda_h+1-\rho\bar{\bar{\delta}}} &> \hat{w} > \frac{\lambda_l}{\lambda_l+1-\rho\bar{\bar{\delta}}} \\
\frac{1-\rho\bar{\bar{\delta}}}{\lambda_l+1-\rho\bar{\bar{\delta}}} &> 1-\hat{w} > \frac{1-\rho\bar{\bar{\delta}}}{\lambda_h+1-\rho\bar{\bar{\delta}}}.
\end{aligned}$$

Then for $\rho > 0$, we have:

$$\begin{aligned}
\lambda_j\left(1-\hat{w}\right)-\left(1-2\rho\bar{\bar{\delta}}\right)\hat{w} &> \frac{\lambda_j\left(1-\rho\bar{\bar{\delta}}\right)}{\lambda_h+1-\rho\bar{\bar{\delta}}} - \frac{\lambda_h\left(1-2\rho\bar{\bar{\delta}}\right)}{\lambda_h+1-\rho\bar{\bar{\delta}}} \\
&\Rightarrow \frac{dEU^P}{d\hat{q}} > 0 \text{ for } j=h \\
&\Rightarrow \frac{dEU^P}{d\hat{q}} > 0 \text{ for } j=l \text{ as long as} \frac{\lambda_l}{\lambda_h} \geq \frac{1-2\rho\bar{\bar{\delta}}}{1-\rho\bar{\bar{\delta}}}.
\end{aligned}$$

For $\rho < 0$, we can establish by a similar argument:

$$\begin{aligned}
\lambda_j (1-\hat{w}) - \left(1-2\rho\overline{\delta}\right)\hat{w} &< \frac{\lambda_j\left(1-\rho\overline{\delta}\right)}{\lambda_l + 1 - \rho\overline{\overline{\delta}}} - \frac{\lambda_l\left(1-2\rho\overline{\delta}\right)}{\lambda_l + 1 - \rho\overline{\overline{\delta}}} \\
&\Rightarrow \frac{dEU^P}{d\hat{q}} < 0 \text{ for } j = l \\
&\Rightarrow \frac{dEU^P}{d\hat{q}} < 0 \text{ for } j = h \text{, as long as } \frac{\lambda_l}{\lambda_h} \geq \frac{1-2\rho\overline{\delta}}{1-\rho\overline{\delta}}.
\end{aligned}$$

Q.E.D. ■

Proof of Proposition 4:

As before, symbols with ^ are calculated using the agents' perceptions. Clearly, when c is zero, we are back to the homogeneous case. Given that the planner and agent's expected utilities and their corresponding derivatives is continuous in c, information rationing continues to be the only equilibrium that survives the Cho and Krep's Intuitive Criterion when c is small enough. Accordingly, we focus on the separating equilibrium where $\widehat{q}\left(\overline{\delta}=1\right) = 0$ (1) for the case of $\rho > 0$ ($\rho < 0$).

The principal's expected utility EU^p (conditional on v):

$$\begin{aligned}
EU^p &= -\frac{1}{2}\left[\int_D \left(\frac{\widehat{w}_i^2}{\alpha} + \frac{(1-\widehat{w}_i)^2}{\beta_i}\right) + \int_{ND} \frac{1}{\beta_i}\right] di + \rho \frac{\left(\int_D \widehat{w}_i di\right)^2}{\alpha} \\
&= \frac{1}{\alpha}\left[\rho\left(\int_D \widehat{w}_i di\right)^2 - \frac{1}{2}\int_D \left(\widehat{w}_i^2 + \lambda_i (1-\widehat{w}_i)^2\right) di - \frac{1}{2}\int_{ND} \lambda_i di\right]. \qquad (11)
\end{aligned}$$

When $\rho < 0$:

Let's consider the case $\rho < 0$. Suppose the proposition is not true. Then there must exist a mass of informed agents with positive measure $[x, y]$ with $y > x > 0$, and a continuous mass of uninformed agents just immediately below x. Now, let's slightly decrease x and y such that $EU^P[\alpha_h, \alpha_l, D]$ (i.e. the principal's expected utility when her true type is α_h; agents perceive her as type α_l; and the set of inner circle is D) is constant. Mathematically, holding $EU^P[\alpha_h, \alpha_l, D]$ constant implicitly defines x as a function of y:

$$\frac{dx}{dy} = -\frac{\frac{\partial EU^P[\alpha_h,\alpha_l,D]}{\partial y}}{\frac{\partial EU^P[\alpha_h,\alpha_l,D]}{\partial x}}.$$

The sign of $\frac{dx}{dy}$ must be positive due to the fact that information rationing is costly under the parameter value considered in this paper. The net effect on the low type principal's payoff is

$$\frac{dEU^P}{dy} = \underbrace{\frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial x}}_{<0}\frac{dx}{dy} + \underbrace{\frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial y}}_{>0}.$$

We would like to establish $\frac{dEU^P}{dy} < 0$. Equivalently,

$$\frac{dEU^P}{dy} = \frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial x}\frac{dx}{dy} + \frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial y} < 0$$
$$\Leftrightarrow \frac{dx}{dy} > -\frac{\frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial y}}{\frac{\partial EU^P[\alpha_l,\alpha_l,D]}{\partial x}}$$

At the end of this proof we derive the expression for $\frac{\partial EU^P[\alpha,\widehat{\alpha},D]}{\partial y}$ in general. Applying those expressions, the above inequality is equivalent to

$$\Leftrightarrow \frac{\frac{\widehat{w}_y}{2\alpha_h}\left[\frac{\alpha_h}{\beta(y)} + A(y|\alpha_h,\hat{\alpha}=\alpha_l)\right]}{\frac{\widehat{w}_x}{2\alpha_h}\left[\frac{\alpha_h}{\beta(x)} + A(x|\alpha_h,\hat{\alpha}=\alpha_l)\right]} > \frac{\frac{\widehat{w}_y}{2\alpha_l}\left[\frac{\alpha_l}{\beta(y)} + A(y|\alpha_l,\hat{\alpha}=\alpha_l)\right]}{\frac{\widehat{w}_x}{2\alpha_l}\left[\frac{\alpha_l}{\beta(x)} + A(y|\alpha_l,\hat{\alpha}=\alpha_l)\right]}$$
$$\Leftrightarrow \frac{\alpha_h + \beta(y)A(x|\alpha_h,\hat{\alpha}=\alpha_l)}{\alpha_l + \beta(y)A(x|\alpha_l,\hat{\alpha}=\alpha_l)} > \frac{\alpha_h + \beta(x)A(y|\alpha_h,\hat{\alpha}=\alpha_l)}{\alpha_l + \beta(x)A(y|\alpha_l,\hat{\alpha}=\alpha_l)}$$

The second equivalency is obtained by plugging in the expression for $\frac{dx}{dy}$.

By Lemma A1 proved at the end of the appendix, the last inequality is true if

$$\frac{\alpha_l}{\beta(y)} + A(y|\alpha_l,\hat{\alpha}=\alpha_l) + \frac{B(y|\hat{\alpha}=\alpha_l)}{\widehat{k}\hat{L}_y^2} < 0.$$

Note that

$$
\begin{aligned}
& \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + \frac{B(y|\hat{\alpha} = \alpha_h)}{\widehat{k}\hat{L}_y^2} \\
= \;& \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + \frac{1}{\widehat{k}\hat{L}_y^2}\left[\frac{2\rho}{1-\rho\hat{L}}\int_D \hat{w}_i\left(2-\hat{L}_i\right)di - \underbrace{\left(1-\widehat{k}\hat{L}_y\right)}_{=w_y}\right] \\
= \;& \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + \frac{2\rho}{\widehat{k}\hat{L}_y^2\left(1-\rho\hat{L}\right)}\int_D \hat{w}_i\left(2-\hat{L}_i\right)di - \frac{1}{\widehat{k}\hat{L}_y^2} + \frac{1}{\hat{L}_y} \\
= \;& \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + \underbrace{1+\lambda_y}_{=1/\hat{L}_y} - \frac{1}{\hat{L}_y^2}\underbrace{\left(\frac{1}{\widehat{k}} - \frac{2\rho}{1-\rho\delta}\int_D \hat{w}_i\left(2-\hat{L}_i\right)di\right)}_{\equiv\hat{\Sigma}} \\
= \;& 2\left(\frac{\alpha_l}{\beta(y)}+1\right) - 1 + A(y|\alpha_l, \hat{\alpha} = \alpha_l) - \left(\frac{\alpha_l}{\beta(y)}+1\right)^2\hat{\Sigma} \qquad (**)
\end{aligned}
$$

Note that

$$
\begin{aligned}
\hat{\Sigma} &= \frac{1}{\widehat{k}} - \frac{2\rho}{1-\rho\delta}\int_D \hat{w}_i\left(2-\hat{L}_i\right)di \\
&= \frac{1}{\widehat{k}} - \frac{2\rho}{1-\rho\delta}\left[2\int_D \hat{w}_i di - \int_D\left(1-\widehat{k}\hat{L}_i\right)\hat{L}_i d\right] \\
&= \frac{1}{\widehat{k}} - \frac{2\rho}{1-\rho\delta}\left[2\left(\bar{\delta}-\widehat{k}\hat{L}\right) - \hat{L} + \hat{k}\hat{L}_2\right] \;\; (\text{where } \widehat{L}_2 \equiv \int_D \widehat{L}_i^2 di) \\
&= \frac{1}{\widehat{k}} - \frac{2\rho}{1-\rho\delta}\left[\frac{2\left(\bar{\delta}-\hat{L}\right)}{1-\rho\hat{L}} - \hat{L} + \hat{k}\hat{L}_2\right] \\
&= \frac{1-\rho\hat{L}}{1-\rho\delta} - \frac{2\rho}{1-\rho\delta}\left[\frac{2\left(\bar{\delta}-\hat{L}\right)}{1-\rho\hat{L}} - \hat{L} + \hat{k}\hat{L}_2\right] \\
&= \frac{1-\rho\hat{L}}{1-\rho\delta} - 4\underbrace{\frac{\rho\left(\bar{\delta}-\hat{L}\right)}{(1-\rho\delta)\left(1-\rho\hat{L}\right)}}_{=\frac{1}{(1-\rho\delta)}-\frac{1}{(1-\rho\hat{L})}} + \frac{2\rho\hat{L}}{1-\rho\delta} - \frac{2\rho\hat{L}_2}{1-\rho\delta}\hat{k} \\
&= \frac{1+\rho\hat{L}}{1-\rho\delta} - \frac{4}{1-\rho\delta} + \frac{4}{1-\rho\hat{L}} - \frac{2\rho\hat{L}_2}{1-\rho\hat{L}} \\
&= \frac{\rho\hat{L}-3}{1-\rho\delta} + \frac{4-2\rho\hat{L}_2}{1-\rho\hat{L}}.
\end{aligned}
$$

Clearly, $\hat{\Sigma} = 1$ when $\rho = 0$. Also,

$$\begin{aligned}
\frac{\partial \Sigma}{\partial \rho} &= \frac{(1-\rho\delta)\hat{L} + \delta\left(\rho\hat{L} - 3\right)}{(1-\rho\delta)^2} + \frac{-2\hat{L}_2\left(1-\rho\hat{L}\right) + \hat{L}\left(4 - 2\rho\hat{L}_2\right)}{\left(1-\rho\hat{L}\right)^2} \\
&= \frac{\hat{L} - 3\delta}{(1-\rho\delta)^2} + \frac{4\hat{L} - 2\hat{L}_2}{\left(1-\rho\hat{L}\right)^2} \\
&< \frac{\hat{L} - 3\delta + 4\hat{L} - 2\hat{L}_2}{\left(1-\rho\hat{L}\right)^2} \text{ b/c } \hat{L} < 3\delta \text{ and } \frac{1}{(1-\rho\delta)^2} > \frac{1}{\left(1-\rho\hat{L}\right)^2} \\
&= \frac{5\hat{L} - 3\hat{L} - 3\widehat{P} - 2\hat{L}_2}{\left(1-\rho\hat{L}\right)^2} = \frac{2\left(\hat{L} - \hat{L}_2\right) - 3\left(\bar{\delta} - \hat{L}\right)}{\left(1-\rho\hat{L}\right)^2} < 0
\end{aligned}$$

The last inequality is obtained because

$$\begin{aligned}
\hat{L} - \hat{L}_2 &= \int_D \left[\frac{1}{1+\hat{\lambda}_i} - \frac{1}{\left(1+\hat{\lambda}_i\right)^2}\right] di \\
&= \int_D \left[\frac{\hat{\lambda}_i}{\left(1+\hat{\lambda}_i\right)^2}\right] di < \int_D \left[\frac{\hat{\lambda}_i}{1+\hat{\lambda}_i}\right] di \\
&= \int_D \left[1 - \frac{1}{1+\hat{\lambda}_i}\right] di = \bar{\delta} - \hat{L}.
\end{aligned}$$

Now, let's add and subtract $(\lambda_y + 1)^2$ from (**) and then combine terms:

$$\begin{aligned}
(**) &= -\left[\left(\frac{\alpha_h}{\beta(y)} + 1\right)^2 - 2\left(\frac{\alpha_h}{\beta(y)} + 1\right) + 1\right] + \underbrace{\left(\frac{\alpha_h}{\beta(y)} + 1\right)^2 \left(1 - \hat{\Sigma}\right)}_{=0 \text{ when } \rho=0} + A\left(y|\alpha_h, \hat{\alpha} = \alpha_h\right) \\
&= -\left[\frac{\alpha_h}{\beta(y)}\right]^2 + \left(\frac{\alpha_h}{\beta(y)} + 1\right)^2 \left(1 - \hat{\Sigma}\right) + A\left(y|\alpha_h, \hat{\alpha} = \alpha_h\right).
\end{aligned}$$

Since $A\left(y|\alpha_l, \hat{\alpha} = \alpha_l\right) = 0$ and $1 - \hat{\Sigma} = 0$ when $\rho = 0$ and both are strictly increasing in ρ,

$$-\left[\frac{\alpha_l}{\beta(y)}\right]^2 + \left(\frac{\alpha_l}{\beta(y)} + 1\right)^2 \left(1 - \hat{\Sigma}\right) + A\left(y|\alpha_l, \hat{\alpha} = \alpha_l\right) < 0, \forall \rho < 0.$$

When $\rho > 0$:

Let's consider the case $\rho > 0$. Suppose not. There must exist a mass of informed agents with positive measure $[x, y]$ with $y > x > 0$, and a continuous mass of uninformed agents just immediately below x. Now, let's slightly decrease x and y such that $EU^P[\alpha_l, \alpha_h, D]$ (i.e. the principal's expected utility when her true type is α_l; agents perceive her as type α_h; and the set of inner circle is D) is constant. Mathematically, holding $EU^P[\alpha_l, \alpha_h, D]$ constant implicitly defines x as a function of y:

$$\frac{dx}{dy} = -\frac{\frac{\partial EU^P[\alpha_l,\alpha_h,D]}{\partial y}}{\frac{\partial EU^P[\alpha_l,\alpha_h,D]}{\partial x}} > 0 \text{ (by partial disclosure hurts result).}$$

We need to show that move y down a bit closer to 0 (taking into its effect on the necessary change in x to maintain $EU^P(\alpha_l, \alpha_h, D)$ benefits the high type planner, i.e.,

$$\frac{dEU^P}{dy} = \underbrace{\frac{\partial EU^P[\alpha_h, \alpha_h, D]}{\partial x}}_{<0}\frac{dx}{dy} + \underbrace{\frac{\partial EU^P[\alpha_h, \alpha_h, D]}{\partial y}}_{>0} < 0$$

$$\Leftrightarrow -\frac{\frac{\partial EU^P[\alpha_l,\alpha_h,D]}{\partial y}}{\frac{\partial EU^P[\alpha_l,\alpha_h,D]}{\partial x}} = \frac{dx}{dy} > -\frac{\frac{\partial EU^P[\alpha_h,\alpha_h,D]}{\partial y}}{\frac{\partial EU^P[\alpha_h,\alpha_h,D]}{\partial x}}$$

$$\Leftrightarrow \frac{\frac{\widehat{w}_y}{2\alpha_l}\left[\frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_h)\right]}{\frac{\widehat{w}_x}{2\alpha_l}\left[\frac{\alpha_l}{\beta(x)} + A(x|\alpha_l, \hat{\alpha} = \alpha_h)\right]} > \frac{\frac{\widehat{w}_y}{2\alpha_h}\left[\frac{\alpha_h}{\beta(y)} + A(y|\alpha_h, \hat{\alpha} = \alpha_h)\right]}{\frac{\widehat{w}_x}{2\alpha_h}\left[\frac{\alpha_h}{\beta(x)} + A(y|\alpha_h, \hat{\alpha} = \alpha_h)\right]}$$

$$\Leftrightarrow \frac{\alpha_l + \beta(y) A(y|\alpha_l, \hat{\alpha} = \alpha_h)}{\alpha_h + \beta(y) A(y|\alpha_h, \hat{\alpha} = \alpha_h)} > \frac{\alpha_l + \beta(x) A(x|\alpha_l, \hat{\alpha} = \alpha_h)}{\alpha_h + \beta(x) A(x|\alpha_h, \hat{\alpha} = \alpha_h)}.$$

By Lemma A1 (proved at the end of the appendix), the last inequality is true if

$$\frac{\alpha_h}{\beta(y)} + A(y|\alpha_h, \hat{\alpha} = \alpha_h) + \frac{B(y|\hat{\alpha} = \alpha_h)}{\widehat{k}\hat{L}_y^2} > 0.$$

Using a similar approach in obtaining (**) above, we get:

$$\begin{aligned} & \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + \frac{B(y|\hat{\alpha} = \alpha_l)}{\widehat{k}\hat{L}_y^2} \\ = \; & \frac{\alpha_l}{\beta(y)} + A(y|\alpha_l, \hat{\alpha} = \alpha_l) + 1 + \lambda_y - \frac{1}{\widehat{k}\hat{L}_y^2}\left[1 - \frac{2\rho}{1-\rho\widehat{L}}\int_D \hat{w}_i\left(2 - \hat{L}_i\right) di\right]. \end{aligned}$$

This expression is positive for sure when $1 - \frac{2\rho}{1-\rho\widehat{L}}\int_D \hat{w}_i\left(2 - \hat{L}_i\right) di < 0$. Note that at

$\rho = 1/2$ and $\alpha_l \to \infty$, $\hat{L}_i \to 0$ and $\hat{w}_i \to 1$. Hence,

$$1 - \frac{2\rho}{1-\rho\hat{L}} \int_D \hat{w}_i \left(2 - \hat{L}_i\right) di \to 1 - 2\bar{\delta}.$$

Note that when $\alpha_h = \alpha_l$, $\bar{\delta} = 1$. As such, as long as the two α's are not too far apart, $\lim_{\alpha_l \to \infty, \rho \to 1/2} \left[1 - \frac{2\rho}{1-\rho\hat{L}} \int_D \hat{w}_i \left(2 - \hat{L}_i\right) di\right] < 0.$ Q.E.D. ■

Derive the expression for $\frac{\partial EU^P}{\partial y}$:

Define $\hat{k} \equiv \frac{1-\rho\bar{\delta}}{1-\rho\hat{L}}$, $\hat{\lambda}_y \equiv \frac{\hat{\alpha}}{\beta_y}$, $\hat{L}_y \equiv \frac{1}{\hat{\lambda}_y+1}$and $\widehat{w}_y = 1 - \frac{1-\rho\bar{\delta}}{1-\rho\hat{L}}\hat{L}_y$. Thus, from (11), we can compute $\frac{\partial EU^P}{\partial y}$ as:

$$\frac{1}{\alpha}\left[2\rho\left(\int_D \widehat{w} di\right)\left(\widehat{w}_y + \int_D \frac{\partial \widehat{w}_i}{\partial y} di\right) - \frac{1}{2}\left[\widehat{w}_y^2 + \lambda_y\left(1-\widehat{w}_y\right)^2 - \lambda_y\right] - \frac{1}{2}\int_D \frac{\partial\left[\widehat{w}_i^2 + \lambda_i\left(1-\widehat{w}_i\right)^2\right]}{\partial y}\right]$$

$$= \frac{1}{\alpha}\left[2\widehat{w}_y\rho\left(\int_D \widehat{w}_i di\right)\left(1 + \frac{\rho\hat{L}}{1-\rho\hat{L}}\right) - \frac{1}{2}\left[\widehat{w}_y^2 + \lambda_y\left(1-\widehat{w}_y\right)^2 - \lambda_y\right] - \int_D \left[\widehat{w}_i - \hat{k}\frac{\alpha}{\hat{\alpha}}\left(1-\hat{L}_i\right)\right]\frac{\partial \widehat{w}_i}{\partial y} di\right]$$

$$= \frac{\widehat{w}_y}{\alpha}\left[\frac{2\rho\int_D \widehat{w}_i di}{1-\rho\hat{L}} - \frac{1}{2}\left[\widehat{w}_y - \lambda_y\left(2-\widehat{w}_y\right)\right] - \frac{\rho}{1-\rho\hat{L}}\int_D \left[\widehat{w}_i - \hat{k}\frac{\alpha}{\hat{\alpha}}\left(1-\hat{L}_i\right)\right]\hat{L}_i di\right]$$

$$= \frac{\widehat{w}_y}{2\alpha}\left[\lambda_y + A\left(y|\alpha,\hat{\alpha}\right)\right] \ (*),$$

where

$$A\left(y|\alpha,\hat{\alpha}\right) \equiv \frac{2\rho}{1-\rho\hat{L}}\int_D \left\{2\widehat{w}_i - \left[\widehat{w}_i - \hat{k}\frac{\alpha}{\hat{\alpha}}\left(1-\hat{L}_i\right)\right]\hat{L}_i\right\} di - \left[\widehat{w}_y - \hat{k}\frac{\alpha}{\hat{\alpha}}\left(1-\hat{L}_y\right)\right].$$

Since

$$\frac{\partial \widehat{w}_i}{\partial y} = -\hat{L}_i\frac{\partial \hat{k}}{\partial y} = \rho\hat{L}_i\frac{1-\rho\hat{L}-\left(1-\rho\delta\right)\hat{L}_y}{\left(1-\rho\hat{L}\right)^2} = \frac{\rho\hat{L}_i}{1-\rho\hat{L}}t_y,$$

we can express $A\left(y|\alpha,\hat{\alpha}\right)$ as

$$A\left(y|\alpha,\hat{\alpha}\right) = B\left(y|\hat{\alpha}\right) + \frac{\alpha}{\hat{\alpha}}C\left(y|\hat{\alpha}\right) \tag{12}$$

$$\text{where } B\left(y|\hat{\alpha}\right) \equiv \frac{2\rho}{1-\rho\hat{L}}\int_D \widehat{w}_i\left(2-\hat{L}_i\right) di - \widehat{w}_y$$

$$C\left(y|\hat{\alpha}\right) \equiv \hat{Z} + \hat{k}\left(1-\widehat{L}_y\right) \text{ with } \hat{Z} \equiv \hat{k}\frac{2\rho}{1-\rho\hat{L}}\int_D \left(1-\widehat{L}_i\right)\hat{L}_i di.$$

Note when $\hat{\alpha} = \alpha$, $\widehat{w}_i - \hat{k}\frac{\alpha}{\hat{\alpha}}\left(1-\hat{L}_i\right) = 1 - \widehat{k}\hat{L}_i - \widehat{k}\left(1-\hat{L}_i\right) = 1 - \hat{k}$ for all i. Substitute this into the (*) part above, we have

$$\begin{aligned}
A\left(y|\alpha,\hat{\alpha}=\alpha\right) &= \frac{2\rho}{1-\rho\hat{L}}\int_{D}\left[2\widehat{w}_{i}-\left(1-\hat{k}\right)\hat{L}_{i}\right]di-\left(1-\hat{k}\right) \\
&= \frac{2\rho\int_{D}\widehat{w}_{i}di}{1-\rho\hat{L}}+\frac{2\rho\left(\delta-\hat{L}\right)}{1-\rho\hat{L}}-\left(1-\hat{k}\right) \\
&= \frac{2\left(1-\hat{k}\right)}{1-\rho\hat{L}}+\left(1-\hat{k}\right),\ \text{as } \rho\left(\int_{D}\widehat{w}_{i}di\right)=\frac{\rho\left(\delta-\hat{L}\right)}{1-\rho\hat{L}}=1-\hat{k} \\
&= \left(1-\hat{k}\right)\left(\frac{2}{1-\rho\widehat{L}}+1\right) \qquad (13) \\
&= {}_{s}1-\hat{k}=_{s}\rho.
\end{aligned}$$

Note that in fact $A\left(y|\alpha,\hat{\alpha}=\alpha\right)$ is independent of y, strictly increases in ρ and equals 0 when $\rho=0$.

Lemma A1: *Define*

$$R\left(y|\alpha_{1},\alpha_{2}\right)=\frac{\alpha_{1}+\beta\left(y\right)A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)}{\alpha_{2}+\beta\left(y\right)A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)},$$

where $\alpha_1,\alpha_2>0$ *and* $A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)$ *and* $A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)$ *are defined in* (12) *and* (13) *above. Then,*

$$\begin{aligned}
sign\left[\frac{\partial R\left(y|\alpha_{1},\alpha_{2}\right)}{\partial y}\right] &= sign\left(1-\frac{\alpha_{1}}{\alpha_{2}}\right)\ \textit{if}\ \frac{\alpha_{2}}{\beta\left(y\right)}+A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)+\frac{B\left(y|\hat{\alpha}=\alpha_{2}\right)}{\widehat{k}\hat{L}_{y}^{2}}>0\ ;\ \textit{and} \\
sign\left[\frac{\partial R\left(y|\alpha_{1},\alpha_{2}\right)}{\partial y}\right] &= sign\left(\frac{\alpha_{1}}{\alpha_{2}}-1\right)\ \textit{if}\ \frac{\alpha_{2}}{\beta\left(y\right)}+A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)+\frac{B\left(y|\hat{\alpha}=\alpha_{2}\right)}{\widehat{k}\hat{L}_{y}^{2}}<0.
\end{aligned}$$

Proof of Lemma A1:

$$\begin{aligned}
\frac{\partial R\left(y|\alpha_{1},\alpha_{2}\right)}{\partial y} &= {}_{s}\left[\alpha_{2}+\beta\left(y\right)A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)\right]\left[A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)\frac{\partial\beta\left(y\right)}{\partial y}+\beta\left(y\right)\frac{\partial A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)}{\partial y}\right] \\
&\quad -\left[\alpha_{1}+\beta\left(y\right)A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)\right]A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)\frac{\partial\beta\left(y\right)}{\partial y} \\
&= \alpha_{2}A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)\frac{\partial\beta\left(y\right)}{\partial y}+A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)\beta\left(y\right)\frac{\partial\beta\left(y\right)}{\partial y} \\
&\quad +\left(\alpha_{2}+\beta\left(y\right)A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)\right)\beta\left(y\right)\frac{\partial A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)}{\partial y} \\
&\quad -\alpha_{1}A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)\frac{\partial\beta\left(y\right)}{\partial y}-A\left(y|\alpha_{2},\hat{\alpha}=\alpha_{2}\right)A\left(y|\alpha_{1},\hat{\alpha}=\alpha_{2}\right)\beta\left(y\right)\frac{\partial\beta\left(y\right)}{\partial y}
\end{aligned}$$

Then factoring out α_2, the above has the same sign as

$$\begin{aligned} &= {}_sA\left(y|\alpha_1,\hat{\alpha}=\alpha_2\right)\frac{\partial\beta\left(y\right)}{\partial y}-\frac{\alpha_1}{\alpha_2}A\left(y|\alpha_2,\hat{\alpha}=\alpha_2\right)\frac{\partial\beta\left(y\right)}{\partial y} \\ &\quad+\left[\alpha_2+\beta\left(y\right)A\left(y|\alpha_2,\hat{\alpha}=\alpha_2\right)\right]\frac{\beta\left(y\right)}{\alpha_2}\frac{\partial A\left(y|\alpha_1,\hat{\alpha}=\alpha_2\right)}{\partial y}\ . \end{aligned}$$

Substituting in $A\left(y|\alpha_1,\hat{\alpha}=\alpha_2\right)=B\left(y|\hat{\alpha}=\alpha_2\right)+\frac{\alpha_1}{\alpha_2}C\left(y|\hat{\alpha}=\alpha_2\right)$, $A\left(y|\alpha_2,\hat{\alpha}=\alpha_2\right)=B\left(y|\hat{\alpha}=\alpha_2\right)+C\left(y|\hat{\alpha}=\alpha_2\right)$ and

$$\begin{aligned} \frac{\partial A\left(y|\alpha_1,\hat{\alpha}=\alpha_2\right)}{\partial y} &= -\frac{\partial\widehat{w}_y}{\partial y}+\hat{k}\frac{\alpha_1}{\alpha_2}\frac{\partial\left(1-\hat{L}_y\right)}{\partial y}=-\hat{k}\frac{\partial\left(1-\hat{L}_y\right)}{\partial y}+\hat{k}\frac{\alpha_1}{\alpha_2}\frac{\partial\left(1-\hat{L}_y\right)}{\partial y} \\ &= \frac{\alpha_2}{\beta\left(y\right)^2}\widehat{k}\hat{L}_y^2\frac{\partial\beta\left(y\right)}{\partial y}\left(1-\frac{\alpha_1}{\alpha_2}\right),\text{ with }\frac{\partial\left(1-\hat{L}_y\right)}{\partial y}=-\frac{\alpha_2}{\beta\left(y\right)^2}\hat{L}_y^2\frac{\partial\beta\left(y\right)}{\partial y} \end{aligned}$$

and because $\frac{\partial\beta(y)}{\partial y}>0$, we have

$$\begin{aligned} \frac{\partial R\left(y|\alpha_1,\alpha_2\right)}{\partial y} &= {}_sB\left(y|\hat{\alpha}=\alpha_2\right)\left(1-\frac{\alpha_1}{\alpha_2}\right)+\left(\frac{\alpha_1}{\alpha_2}-\frac{\alpha_1}{\alpha_2}\right)C\left(y|\hat{\alpha}=\alpha_2\right) \\ &\quad+\left[\frac{\alpha_2}{\beta\left(y\right)}+A\left(y|\alpha_2,\hat{\alpha}=\alpha_2\right)\right]\widehat{k}\hat{L}_y^2\left(1-\frac{\alpha_1}{\alpha_2}\right) \\ &= \widehat{k}\hat{L}_y^2\left[\frac{\alpha_2}{\beta\left(y\right)}+A\left(y|\alpha_2,\hat{\alpha}=\alpha_2\right)+\frac{B\left(y|\hat{\alpha}=\alpha_2\right)}{\widehat{k}\hat{L}_y^2}\right]\left(1-\frac{\alpha_1}{\alpha_2}\right) \end{aligned}$$

The lemma is proved by noting that $\hat{k}L_y^2>0$ and $\frac{\partial\beta(y)}{\partial y}>0$. Q.E.D. ■