

# Divide and Inform: Rationing Information to Facilitate Persuasion\*

Beatrice Michaeli

University of California, Los Angeles

December 16, 2016

\*I would like to thank two anonymous referees and Thomas Hemmer (the Editor), as well as Philip Berger, Alexander Bleck, Judson Caskey, Hans Christensen, Patricia Dechow, Wouter Dessein, Ron Dye, Paul Fischer, Henry Friedman, Pingyang Gao, Jonathan Glover, Tim Gray, Ilan Guttman, Jeffrey Hales, Barney Hartman-Glaser, Mirko Heinle, Moritz Hiemann, Christian Hoffmann, John Hughes, Emir Kamenica, Andrei Kovrijnykh, Christian Leuz, Xiaojing Meng, Mark Maffett, Bugra Ozel, James Ryans, Haresh Sapra, Catherine Schrand, Philip Stocken, Brett Truman and seminar participants at Booth School of Business, Carnegie Mellon University, Columbia Business School, Kellogg School of Management, Tel Aviv University, Tuck School of Business, UC Berkeley, UCLA, Wharton and Yale School of Management for their helpful comments and suggestions. I am very grateful to Navin Kartik, Nahum Melumad and Amir Ziv for their valuable discussions that benefited this project. My deepest gratitude goes to Tim Baldenius for the continuous support of my PhD study and the extraordinary ongoing guidance in research.

## Abstract

### Divide and Inform: Rationing Information to Facilitate Persuasion

This article develops a Bayesian persuasion model examining a manager's incentives to gather information when the manager can disseminate this information selectively to interested parties ("users") and when the objectives of the manager and the users are not perfectly aligned. The model predicts that, if the manager can choose the subset of users to receive the information, then the manager may gather more precise information. The article identifies conditions under which a regime that allows managers to grant access to information selectively maximizes aggregate information. Strikingly, this happens when the objectives of managers and users are sufficiently *misaligned*. This finding is robust to variations of the model such as information acquisition cost, unobservable precision, sequential noisy actions taken by the users and delayed choice of the subset of users in "the know." These results call into doubt the common belief that forcing managers to provide unrestricted access to information to all potential users is always beneficial.

**Key words:** Bayesian persuasion, ex ante commitment to information design, endogenous quality of information, verifiable messages, selective dissemination, persuasion with multiple receivers

# 1 Introduction

Fairness considerations, common practices and regulations require that managers communicate information to all potential users of this information. The standard view is that fairness and greater availability of information at an aggregate level go hand in hand. The justification for this view is that, provided managers report truthfully, they disseminate the same quality of information regardless of who receives it. However, this argument ignores the fact that the quality of information may be endogenous. Managers choose not only to whom to disclose the information but also whether and what quality of information to gather in the first place, and this choice may be affected by the size of the audience and the managers' objectives. This article studies the effects of managers' discretion to limit access to information to a subset of users on the managers' incentives to gather information and thereby on the aggregate information available to the public. Within the confines of the model, this allows inferences about ex ante efficiency and the trade-off between the fairness and the information objectives.

I develop a model of Bayesian persuasion with information control that builds on the model of Kamenica and Gentzkow (2011). The players—a firm manager (“she”) and a group of identical users (“they” in plural, “he” in singular)—have misaligned preferences in a sense that the users prefer to take actions aligned with the state of nature, while the manager prefers that the users' aggregate action be partially aligned with the state and partially aligned with a target that may differ from that state.<sup>1</sup> This setting creates an incentive for the manager, if not constrained by regulation, to choose a quality of information and subset of users resulting in posterior expectations that improve the odds of satisfying the manager's preferences. Perhaps more surprising, allowing managers to do so may make all parties at least weakly better off (some strictly better off).

While stylized, the model captures essential features of several examples of misaligned preferences that could lead to limited dissemination in equilibrium. A manager may seek to induce consensus earnings forecasts from financial analysts (users) that the firm is more likely to meet or beat.<sup>2</sup> A CEO (manager) may increase the likelihood of a majority vote

---

<sup>1</sup>All results hold when the users are also biased as long as they prefer actions that are more sensitive to the state than the aggregate action the manager prefers.

<sup>2</sup>The manager may also seek to induce forecast that is high (to increase the stock price and the managers' stock-based compensation) or low (to trigger a low exercise price before an option grant date).

on the part of shareholders (users) for continuation of her employment or for approving an acquisition that contributes to her empire building. Investors (users) may choose whether to participate in supplying capital to an entrepreneur (manager) based on an expected future payoff, while the entrepreneur may suffer a private loss not borne by investors if a desired level of funding is not achieved. Similar conditions may arise when a CFO (manager) who seeks to obtain a loan from a syndicate of lenders (users), or when a CEO (manager) provides access to internally generated data to divisional and sales managers (users) in hope of achieving a prescribed performance threshold or implementing a product strategy. A government official (manager) may likewise want to achieve a certain target through dissemination of data and influencing the decisions of market participants (users). In these examples, the manager may find some combination of information quality and limited dissemination to be optimal. In many cases, such limited dissemination is common. For example, Chinese walls are frequently used for restricting the dissemination of information within firms.<sup>3</sup> In other cases, selective dissemination is prohibited. For example, Regulation FD prohibits selective disclosure by managers, but its effect on the informational efficiency is controversial.<sup>4</sup> This article provides a theoretical framework that links the managers' ability to disseminate information selectively with factors that influence the quality of information and, by doing so, illuminates the tension between fairness and informational efficiency.

The first part of the article considers a setting in which the manager has to choose the fraction of informed users *ex ante* and the implementation of an information system is cost-free. The analysis shows that, if the manager does not have discretion over the access to information, she implements a perfectly revealing information system *only* if her preferences are sufficiently aligned with those of the users.<sup>5</sup> If the players' objectives are very misaligned, the users react too sensitively to the signal from the manager's point of view. Hence the manager is better off not providing information and leaving the users to act

---

<sup>3</sup>I thank an anonymous referee for pointing this out.

<sup>4</sup>Herrmann, Hope and Thomas (2008), for example, show an increase in the average analyst forecast accuracy after Reg FD, while Agrawal, Chadha and Chen (2006) document a decrease. Francis, Nanda and Wang (2006) and Heflin, Subramanyam and Zhang (2003) find no evidence of significant change. At a post-adoption roundtable discussion, analysts and other panelists expressed concerns that Reg FD had diminished the quantity and quality of the information disseminated by firms. See Regulation Fair Disclosure Revisited, Report by Commissioner Laura S. Unger, <http://www.sec.gov/news/studies/regfdstudy.htm>.

<sup>5</sup>This is a multiple-receivers variation of the findings in Kamenica and Gentzkow (2011).

on their prior.<sup>6,7</sup> In contrast, if the manager can restrict access to information, she *always* implements a perfectly revealing system. This happens despite the fact that the users are *ex ante* identical. By keeping some of the users “in the dark,” the manager can regulate the sensitivity of the users’ *aggregate* action to changes in the state of nature. As a result of this additional degree of freedom, the manager now finds it in her interest to implement the most informative system, the output of which is only selectively disseminated. As one would expect, the optimal fraction of privileged users who get to observe the signal is increasing in the degree of players’ objective alignment.

A corollary of the preceding discussion is that, within the confines of the base model, if the players’ preferences are sufficiently *misaligned*, a regime that allows for information discrimination (“unregulated dissemination”) not only maximizes aggregate information, but it also *Pareto dominates* a regime that requires information dissemination to all users (“mandated dissemination”). It makes the manager better off (by revealed preference) as well as some users (because they receive information that would not be available otherwise) without making the other users worse off (because they do not observe the information anyway).<sup>8</sup> Paradoxically, when the players’ preferences over actions are *misaligned*, their preferences over regimes are *aligned*. The opposite is also true: when the players’ preferences over actions are sufficiently *aligned*, their preferences over regimes are *misaligned*. The manager then prefers unregulated dissemination, while all users at least weakly prefer mandated dissemination. This result calls into doubt the conventional wisdom that regulating information dissemination and requiring equal access is especially needed when the incentive conflict is severe. Ironically, under this scenario, regulations forcing equal access will promote fairness but at the expense of reduced overall information.

With costless information acquisition, the optimal precision is a bang-bang solution: the manager implements either a perfectly revealing information system or one that does not convey any information at all. However, interior precision levels and information acquisition costs are frequently observed in practice. With costly implementation, I find that, if the

---

<sup>6</sup>As anecdotal evidence, in July 2013, the SEC required Urban Outfitters to publicly disclose the effect of direct-to-customer sales on the net retail segment sales. In response, Urban Outfitters declared that, effective the first quarter of 2014, the company will no longer gather this information even for internal purposes.

<sup>7</sup>In 1970s, the car manufacturers did not perform certain safety tests to avoid the disclosure of the results.

<sup>8</sup>In subsection 5.3.1, I consider an extension in which the users are rewarded for relative performance. In this case, no Pareto ranking can be made.

fraction of informed users were exogenous, then the interior optimal precision level would be lower, the farther this fraction is from the one the manager would choose if she could.

When the preferences between the players are misaligned, the Pareto ranking of mandated and unregulated dissemination regimes remains the same. However, when preferences over actions are aligned, the introduction of precision costs creates disagreement between the users regarding the preferred regime. The assessment of the aggregate users' welfare depends on two countervailing effects: (i) a *precision effect*—the information collected under unregulated dissemination is more precise than under mandated dissemination—and (ii) an *omission effect*—the fraction of users who observe the information under unregulated dissemination is lower than under mandated dissemination. This article identifies sufficient conditions under which the users are better off, on an aggregate level, under unregulated dissemination. Put differently, even the users themselves, at some prior state, would collectively agree to unregulated dissemination, as long as they are behind the veil of ignorance (Harsanyi, 1955), that is, before each learns whether he will be included in the group of informed users.

In the last part of the article, I consider variations of the model. First, I relax the assumption that the manager has to choose the fraction of informed users ex ante and examine the manager's dissemination strategy and incentives to implement an information system. Next, I consider a setting in which the information precision is unobservable and show that the same equilibrium persists, but it is no longer unique. Lastly, I extend the results by allowing the users to choose the timing of their actions. I identify conditions under which, in equilibrium, (i) the users act simultaneously and (iii) the users act sequentially but no unraveling occurs.

The article belongs to the persuasion literature. Its primary theoretical antecedent is the Bayesian persuasion model of Kamenica and Gentzkow (2011). As in Alonso and Camara (2014) and Wang (2013), this article focuses on information control with multiple receivers. Similar to Gentzkow and Kamenica (2013), the model allows persuasion to be costly. The article relates to the literature that studies ex ante commitment to information system design (Baiman 1975; Arya, Glover and Sivaramakrishnan 1997; Göx and Wagenhoffer 2009; Shavel 1991) and literature that links ex ante information acquisition with ex post choice of communication strategy (Che and Kartik 2009; Fischer and Stocken 2010; Hughes and Pae 2004; Pae 1999). One of the key applications of the study is financial reporting to

external users, and so the article relates to the literature on costs and benefits of information dissemination, which is reviewed by Verrecchia (2001), Dye (2001), Beyer, Cohen, Lys, and Walther (2010) and Stocken (2012). The model has features in common with the mandatory and the voluntary disclosure literature. The key difference is the timing: the information dissemination is only voluntary *ex ante* when the manager decides whether to implement an information system. However, once an implemented system has generated a signal, the manager must truthfully share it with the predetermined group of users.

Prior literature on selective disclosure considers models that focus on the *ex post* strategic communication between players with asymmetric information (e.g., cheap-talk and disclosure models). I assume that the players have symmetric information and focus on the *ex ante* strategic control of the information. This allows for identification of unintended consequences of prohibiting selective dissemination that differ from the already studied effects driven by herding (Arya, Glover, Mittendorf and Narayanamoorthy 2005), externalities (Chen, Lewis, Schipper and Zhang 2016), users' incentives to gather information (Jorgensen, Li and Melumad 2013; Demski and Feltham 1994; Kim and Verrecchia, 1991; McNichols and Trueman 1994), price efficiency (Dutta 1996) and private information sale (Bushman 1991; Sabino 1993). Other related studies consider *ex ante* commitment to dissemination policy (Edmans, Heinle and Huang 2013; Gao and Liang 2013) and choice of report precision that is publicly disclosed (Penno 1996; Titman and Trueman 1986).

The article proceeds as follows. Section 2 describes the model. Section 3 analyzes the access to information and considers policy implications. Section 4 extends the results to costly implementation of information systems. Section 5 discusses the robustness of the results to variations of the model. Section 6 concludes. All proofs are in the appendix.

## 2 Model Setup

I consider a manager (e.g., CEO, CFO, government official) and a group of individuals labeled “users” (e.g., analysts, shareholders, divisional managers, investors) who are interested in learning about a state of nature to make their decisions. There is a continuum of identical users uniformly distributed on the interval  $[0, 1]$ .<sup>9</sup> In the main part of the article, the users

---

<sup>9</sup>This is a convenience assumption. The only place where this assumption affects the analysis is subsection 5.3.1.

act simultaneously. In subsection 5.3, I extend the analysis by assuming the users can choose the date at which they act. The payoff of user  $i$  depends on his own action  $a_i \in \mathbb{R}$  (e.g., analyst's forecast, divisional manager's production decision, investment level) and the state of nature  $\omega \in \mathbb{R}$  (e.g., economic earnings, inventory levels)

$$u(a_i, \omega) = -(a_i - \omega)^2.$$

For any realization of the state of nature, the interior solution that maximizes the payoff of user  $i$  is

$$a^*(\omega) \in \arg \max_{a_i} u(a_i, \omega) = \omega, \tag{1}$$

i.e., a representative user prefers an action that is fully aligned with the state of nature.<sup>10,11</sup>

In the main part of the article, the payoff of user  $i$  does not depend on the actions of other users. In Section 5, I relax this assumption. The payoff of the manager depends on the aggregate action of all users, denoted by

$$A \equiv \int_0^1 a_i di,$$

(e.g., consensus forecast, overall firm production) and on the state of nature  $\omega$  and is denoted by:

$$v(A, \omega) = -(A - k\omega - (1 - k)\bar{\omega})^2,$$

where  $k \in (0, 1)$  and  $\bar{\omega} \in \mathbb{R}$  are commonly known parameters. For any realization of the state of nature, the interior solution that maximizes the manager's payoff is

$$A^*(\omega) \in \arg \max_A v(A, \omega) = k\omega + (1 - k)\bar{\omega}. \tag{2}$$

The manager prefers an aggregate action that is partially aligned with the state of nature  $\omega$  and partially biased toward some exogenous value  $\bar{\omega}$ .<sup>12</sup> For example, managers may prefer

---

<sup>10</sup>For example, analysts care about their reputation for accuracy; divisional managers want to accelerate production if there is not enough inventory.

<sup>11</sup>Analysts might be upward biased in good states and downward biased in bad states to amplify the trading volume. Allowing for different sensitivity to the state of nature and bias in the users' bliss point will not change the results qualitatively as long as the users' preferred actions are more sensitive to the state of nature than the manager's preferred aggregate action.

<sup>12</sup>A similar preference was introduced in a cheap talk setting by Melumad and Shibano (1991) and by Kamenica and Gentzkow (2011) in their lobbying example of a Bayesian persuasion game. In Melumad and Shibano (1991), the communication game is affected by preference reversal, i.e., the bliss point of the manager can be lower or higher than the bliss point of the user in different environments. In my model, although preference reversal is possible depending on the relative magnitude of  $\bar{\omega}$  vis-a-vis  $\omega$ , it does not affect the persuasion game because the manager can persuade the users to take an aggregate action that is closer to  $k\omega$  but cannot persuade them to take an action closer to  $(1 - k)\bar{\omega}$ .

consensus forecasts that are less sensitive to the economic earnings and are biased toward a specific forecast that maximizes the manager’s compensation;<sup>13</sup> CEOs prefer acquisitions that satisfy their empire building preferences or a rate of production that fits their product strategy;<sup>14</sup> directors of reserve banks may have a target inflation rate;<sup>15</sup> and entrepreneurs want to receive funding as long as their probability of success is not very low. I refer to  $k$  as the measure of “preference alignment” between the manager and a representative user. As  $k \rightarrow 1$ , the players’ preferences are perfectly aligned because then the manager, just like the users, prefers an action that is fully aligned with the state of nature.

None of the players observes the state of nature  $\omega$ , and all players share the same prior beliefs. The manager can implement an information system that will provide a noisy signal  $s$  of the state of nature.<sup>16</sup> In the basic setting, the information system implementation is cost-free. This assumption is relaxed in Section 4. Each signal realization leads to a posterior belief. Accordingly, an information system creates a distribution over posterior beliefs. This distribution is chosen by the manager and has to be Bayes-plausible, i.e., the expected posterior distribution equals the prior. To keep the analysis of the setting in Section 4 tractable, it will be useful to assume normally distributed signals. To facilitate comparisons across the settings, I impose normality on the distributions of  $\omega$  and  $s$  throughout the article:<sup>17</sup>

$$\begin{aligned}\omega &= \mu_0 + \varepsilon, & \varepsilon &\sim \mathcal{N}\left(0, \frac{1}{\alpha}\right), & \alpha &\geq 0; \\ s &= \omega + \delta, & \delta &\sim \mathcal{N}\left(0, \frac{1}{\beta}\right), & \beta &\geq 0.\end{aligned}$$

The error terms  $\varepsilon$  and  $\delta$  are independent:  $Cov(\varepsilon, \delta) = 0$ . Upon observing the signal realiza-

---

<sup>13</sup>The preferred forecast can be high (to increase the stock price and the managers’ stock-based compensation), low (to trigger low exercise price before option grant date) or mean (that is easy to meet or beat).

<sup>14</sup>High rates if they plan to aggressively penetrate the market or low rates if they plan to discontinue the product.

<sup>15</sup>On the one hand, high inflation rates are more volatile and can generate distortions in the economy. On the other hand, low inflation rates correspond to low interest rates and the central bank might need to reduce its benchmark interest rate to zero. This may make the bank helpless in the face of recession. Currently the Fed’s target is 2%. See September 13, 2015, *The Economist*, “Why the Fed targets 2 % inflation?”

<sup>16</sup>For example, pay for appraisals to assess the value of the firm’s long-lived assets, purchase inventory management software, hire economists to provide forecasts or collect statistical data.

<sup>17</sup>The results in Sections 3 and 5 can be shown for more general distributions of the state of nature and the posterior beliefs.

tion, the players form a posterior belief regarding the state of nature:

$$\omega|s \sim \mathcal{N}\left(\frac{\alpha\mu_0 + \beta s}{\alpha + \beta}, \frac{1}{\alpha + \beta}\right).$$

Let  $\mu(s|\beta) = \frac{\alpha\mu_0 + \beta s}{\alpha + \beta}$  denote the posterior expectation. Under the maintained assumptions, the manager’s choice of a Bayes-plausible distribution over posterior beliefs simplifies to a choice of  $\beta$ , which represents the precision of the signal. A choice of  $\beta \rightarrow \infty$  indicates a signal that fully reveals the realization of  $\omega$ , while a choice of  $\beta \rightarrow 0$  means the signal does not convey any information. The manager’s choice of  $\beta$  is observed by the users but is not contractible. The observability of the precision assumption reflects the fact that corporate governance mechanisms such as internal controls and due diligence procedures that affect the informativeness of disclosure are observable to external and internal users of information.<sup>18</sup> The prior expectation of the state of nature  $\mu_0$  and the variance  $\frac{1}{\alpha}$  are common knowledge.

To analyze the effects of the ability (or lack thereof) of the manager to limit access to information, I compare two alternative regimes: “mandatory dissemination” (hereafter, “MD”) and “unregulated dissemination” (hereafter, “UD”). Under MD the manager is required to disclose the same signal to all users, while under UD the manager can choose the fraction  $x \in [0, 1]$  of users who will observe the signal.<sup>19,20</sup> I refer to those users as “informed” and to the rest of the users as “uninformed.” In other words, ex ante identical users become heterogeneous endogenously by virtue of selective information dissemination.

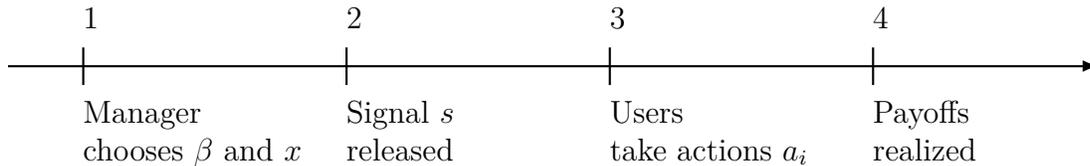
In the main part of the article, the manager chooses the fraction  $x$  of informed users at date 1 and cannot revise her choice afterwards, i.e. has commitment power. This assumption comports with the ex ante commitment assumption of Bayesian persuasion models and reflects the fact that often managers have to provide (or commit to) the access to information early on. For example, managers may need to (i) set up the access of the informed users to a database in advance; (ii) send invitations to conference calls, meetings, or speeches ahead of time to ensure participation; (iii) reserve (and pay for) a conference venue; (iv) contact a syndicate of institutional investors in an IPO. In all these examples, ex post deviation for the chosen subset of informed users is either impractical or very costly (requires additional

---

<sup>18</sup>I thank an anonymous referee for pointing this out.

<sup>19</sup>Alternatively,  $x$  can be interpreted as the probability that user  $i$  observes the signal.

<sup>20</sup>For example, top executives often meet in private with handpicked investors. See Serena NG and Anton Troianovski, “How Some Investors Get Special Access to Companies,” *The Wall Street Journal*, September 27, 2015, <http://www.wsj.com/articles/how-some-investors-get-special-access-to-companies-1443407097>



**Figure 1:** Timeline of events

set-up time, cancellation fees/penalties or results in displeasure of the users). In subsection 5.1, I extend the analysis by relaxing the ex ante commitment assumption and considering cases in which the manager can delay her choice to date 2 (or revise her prior choice) after observing the signal. For most of the article, the observability of  $x$  is irrelevant. It only becomes relevant in subsection 5.1.

I restrict attention to cases in which the signal is truthfully communicated to the informed users. This assumption reflects the litigation threat in case of concealing or distorting information. In my model, the users do not gather information on their own. This assumption is motivated by the fact that, in many cases, firm managers have access to sources of information unavailable to the users.<sup>21</sup>

Figure 1 shows the timeline of the events. At date 1, the manager chooses whether to implement the information system, its precision  $\beta$  and the fraction  $x$  of informed users (under UD). At date 2, the information system reveals signal  $s$  to the fraction  $x$  of informed users. At date 3, the users take actions, and, at date 4, the payoffs are realized.

### 3 Access to Information

I solve the model by backward induction. To avoid confusion I use a subscript  $t = 1$  for the expectation operator to denote the expectation at date 1 over the random variables  $\omega$  and  $s$  and a subscript  $t = 3$  to denote the posterior expectation of  $\omega$  at date 3 (after observing the realization of the signal  $s$ ). At date 3, after observing the choice of information

---

<sup>21</sup>An additional motivation to focus solely on the information gathered by managers is that prior literature has analyzed the effects of selective disclosure on the incentives of the users to acquire information on their own (Jorgensen, Li and Melumad 2011).

system precision and the signal realization, the informed users form Bayesian rational beliefs regarding the state of nature and, given the quadratic loss nature of their payoff functions, take actions that equal the posterior expectation:

$$\hat{a}(s, \beta) \equiv \arg \max_{a_i} E_{t=3}[u(a_i, \omega)|s, \beta] = E_{t=3}[\omega|s, \beta] = \mu(s|\beta). \quad (3)$$

Lacking information, the uninformed users take actions that equal the prior expectation:<sup>22</sup>

$$\hat{a}(\mu_0) \equiv \arg \max_{a_i} E_{t=3}[u(a_i, \omega)|\emptyset] = E_{t=3}[\omega] = E_{t=1}[\omega] = \mu_0. \quad (4)$$

As a result, the aggregate action of the users at date 3 is a weighted average of the posterior (for the informed users) and the prior (for the uninformed users):

$$\hat{A}(x, s, \beta) \equiv \int_0^x \hat{a}(s, \beta) di + \int_x^1 \hat{a}(\mu_0) di = x\mu(s|\beta) + (1-x)\mu_0. \quad (5)$$

The expected payoff of the manager at date 1 is given by

$$V(x, \beta) \equiv E_{t=1}[v(\hat{A}(x, s, \beta), \omega)].$$

Firm managers choose the precision of the information system that they implement, but they cannot always choose the fraction of users who observe the information. For example, even when regulators gravitate toward rules that ensure equal access to information, not all users may observe the available information for various exogenous reasons. If the fraction of informed users is exogenous, then, at date 1, the manager chooses the precision of the information system to maximize her expected utility:

$$\hat{\beta}(x) \in \arg \max_{\beta} V(x, \beta).$$

By implementing an information system with optimally chosen precision the manager “persuades” the users, i.e., she convinces the users to take actions that are closer to her preferred actions and differ from the actions the users would have taken without the information. Kamenica and Gentzkow (2011) find that a sender—in my model, the manager—benefits from persuading a single user if her expected payoff is convex in the single user’s beliefs regarding the state of nature. Lemma 1 translates this result to my setting and extends it to an exogenously given fraction  $x \in [0, 1]$  of informed users.

---

<sup>22</sup>The informed users have no incentive to communicate their information to the uninformed users, because the payoff of each user is affected only by his own action (I assume the users cannot make side payments and collude on the information). The uninformed users cannot infer the signal realization from the actions of the informed users before choosing their actions, because all actions are taken simultaneously (to be relaxed in subsection 5.3).

**Lemma 1** *Suppose  $x \in [0, 1]$  is exogenously given. The manager then implements a perfectly revealing information system if and only if  $k > \frac{x}{2}$ . Otherwise, she does not implement an information system.*

The formal proof is omitted as it follows from the discussion below. Note that at date 2, upon observing the signal, the manager and the users share the same beliefs. Then, by the Law of Iterated Expectations, the manager's expected payoff at date 1 can be conveniently presented as:

$$\begin{aligned}
V(x, \beta) &= E_{t=1} \left[ \underbrace{E_{t=3}[v(\hat{A}(x, s, \beta), \omega)|s, \beta]}_{\text{posterior expectation at date 3}} \right] \\
&\quad \underbrace{\hspace{10em}}_{\text{expectation over signal realizations at date 1}} \\
&= E_{t=1}[E_{t=3}[-(x\mu(s|\beta) + (1-x)\mu_0 - k\omega - (1-k)\bar{\omega})^2|s, \beta]] \\
&= E_{t=1} \left[ \underbrace{2x \left(k - \frac{x}{2}\right) (\mu(s|\beta))^2}_{\text{term 1}} - \underbrace{k^2 (\mu_0^2 + Var(\omega))}_{\text{term 2}} \right. \\
&\quad \left. - \underbrace{[(x - 2k + 1)\mu_0 - (1 - k)\bar{\omega}][(1 - x)\mu_0 - (1 - k)\bar{\omega}]}_{\text{term 3}} \right]. \tag{6}
\end{aligned}$$

The second and third terms in (6) are constant across signal realizations,  $s$ . However, term 1 depends on the posterior expectation that a signal induces, and the manager can control its magnitude by choosing the precision of the information system. Given that the posterior expectation is increasing in  $s$ , then  $(\mu(s|\beta))^2$  is convex in  $s$ . Noting that term 1 is an expectation of a quadratic function of the random variable  $\mu(s|\beta)$  whose expectation equals the prior  $\mu_0$  and, using Jensen's inequality,

$$E_{t=1} \left[ 2x \left(k - \frac{x}{2}\right) (\mu(s|\beta))^2 \right] \begin{cases} > 2x \left(k - \frac{x}{2}\right) \mu_0^2 \geq 0 & \text{if } k > \frac{x}{2}, \\ < 2x \left(k - \frac{x}{2}\right) \mu_0^2 \leq 0 & \text{if } k < \frac{x}{2}. \end{cases}$$

If the preferences of the manager are sufficiently aligned with those of the users ( $k > \frac{x}{2}$ ), then the manager's payoff is convex in  $\mu(s|\beta)$ , and the manager is better off inducing variance in the users' actions by making a signal available than leaving the users with their prior expectation. The opposite logic holds when the preferences of the manager are sufficiently misaligned with those of the users ( $k < \frac{x}{2}$ ). To gain further insight note that using (6),

$$\begin{aligned}
V(x, \beta) &= E_{t=1} \left[ 2x \left( k - \frac{x}{2} \right) (\mu(s|\beta))^2 \right] + \text{const} \\
&= 2x \left( k - \frac{x}{2} \right) (\text{Var}(\mu(s|\beta)) + \mu_0^2) + \text{const},
\end{aligned} \tag{7}$$

where  $\text{Var}(\mu(s|\beta)) = \frac{\beta}{\alpha(\alpha+\beta)}$  is the variance of the posterior expectation. The variance is increasing in the precision of the signal and, by the Law of Total Variances, is bounded from above by the prior variance  $\text{Var}(\omega) = \frac{1}{\alpha}$ . Put differently, when the signal perfectly reveals the state of nature, the variance of the posterior expectation equals the variance of the state of nature. When the preferences of the players are aligned ( $k > \frac{x}{2}$ ), the manager wants to induce as much variance in the posterior expectation as possible, which she achieves by setting  $\beta \rightarrow \infty$ . The opposite is true when the preferences of the players are misaligned ( $k \leq \frac{x}{2}$ ). The manager then minimizes the variance by setting  $\beta \rightarrow 0$ . In the knife-edge case, when  $k = \frac{x}{2}$ , the manager's expected payoff is the same regardless of the variance in posterior expectation that she induces. For the remainder of the article, I assume that, whenever the manager is indifferent, she does not implement an information system.

The preceding discussion is stated in terms of normal distributions of the state of nature and the posterior beliefs. However, as seen from (7), all that matters is the sign of the coefficient  $k - \frac{x}{2}$  attached to the variance in posterior expectation and the ability of the manager to increase or reduce this variance by optimally choosing the distribution of the posterior beliefs. Therefore, the result of Lemma 1 can be shown for more general distributions with well-defined first and second moments.<sup>23</sup>

As a next step, I consider the manager's ability to control the aggregate information flow by limiting access to information of a subset of users. To do so, I relax the assumption that  $x$  is exogenously given and let the manager choose not only the information system precision but also the fraction of informed users. At date 1, the full-fledged optimization problem of the manager is

### Program $\mathcal{P}$

$$\max_{\beta \geq 0, x \in [0,1]} V(x, \beta).$$

---

<sup>23</sup>The expected payoff of the sender depends only on the first and the second moments of the distribution. Hence all results in Section 3 hold for general distributions with well-defined first and second moments. Detailed analysis available upon request.

Let  $(\hat{x}, \hat{\beta})$  denote the solution to this program.

**Proposition 1** *Under unregulated dissemination (UD), for any  $k$ , it is optimal for the manager to implement an information system that perfectly reveals the state of nature to a fraction  $k$  of users, i.e.,  $(\hat{x}, \hat{\beta}) = (k, \infty)$ .*

The manager finds it optimal to provide maximally precise information but to restrict access to it.<sup>24</sup> Given that the manager has two instruments at her disposal,  $x$  and  $\beta$ , it may seem counterintuitive at first that she chooses to restrict access rather than provide opaque information. The expected utility of the manager is:

$$\begin{aligned} V(x, \beta) &= -E_{t=1}[(x\mu(s|\beta) + (1-x)\mu_0 - k\omega - (1-k)\bar{\omega})^2] \\ &= -Var(x\mu(s|\beta) + (1-x)\mu_0 - k\omega - (1-k)\bar{\omega}) \\ &\quad - (E_{t=1}[x\mu(s|\beta) + (1-x)\mu_0 - k\omega - (1-k)\bar{\omega}])^2 \\ &= -Var(x\mu(s|\beta) - k\omega) - (1-k)^2(\mu_0 - \bar{\omega})^2. \end{aligned}$$

Given that  $-(1-k)^2(\mu_0 - \bar{\omega})^2$  is a constant, maximizing  $V(x, \beta)$  boils down to minimizing  $Var(x\mu(s|\beta) - k\omega) \geq 0$ . The only way to minimize this term to zero is to choose  $\beta \rightarrow \infty$  and  $x = k$  because, otherwise, if  $\beta$  is finite, some variance remains for any  $x$ .

To gain additional intuition, recall that the aggregate action of the users at date 3 is as stated in equation (5):

$$\hat{A}(x, s, \beta) = x\mu(s|\beta) + (1-x)\mu_0,$$

while the manager would want it to be as close as possible to her bliss point as stated in equation (2):

$$A^*(\omega) = k\omega + (1-k)\bar{\omega},$$

for any realization of  $\omega$ . If the state of nature were observable, the manager would want the users to react to the realization of  $\omega$  with a response coefficient of  $k$  (as, by (2),  $\frac{dA^*(\omega)}{d\omega} = k$ ). In the model, the users only see the signal,  $s$ , and not the state,  $\omega$ , directly. To regulate the

---

<sup>24</sup>Allowing for different sensitivity of a representative user's bliss point to the state of nature does not change this result qualitatively as long as the manager prefers smoother aggregate action across states. Under this scenario, the optimal fraction of informed users equals the ratio of the sensitivities of the manager's and a representative user's bliss points.

sensitivity with which the users react to the signal, which, by (5), is  $\frac{d\hat{A}(x,s,\beta)}{ds} = x\frac{d\hat{a}(s,\beta)}{ds} + \frac{d\hat{a}(\mu_0)}{ds}$ , the manager has two instruments at her disposal:  $\beta$ , which, by (3), determines the response coefficient  $\frac{d\hat{a}(s,\beta)}{ds} = \frac{\beta}{\alpha+\beta}$  for an informed user, and  $x$ , which determines the fraction of informed users.<sup>25</sup> It is straightforward to show that one way to implement an expected response coefficient of  $k$  is to set  $x = 1$  and a precision level (calculated as a plug) of  $\beta = \alpha\frac{k}{1-k}$ . This, however, introduces noise into the signal, which is costly to the users and to the manager. Alternatively, the manager can set  $x = k$  and  $\beta \rightarrow \infty$ . This would yield the same response coefficient in expectation but avoids the noise in the signal, which makes this the optimal solution.

Note that, by Proposition 1, if the manager is allowed to choose the group of informed users, she always finds it optimal to implement an information system, regardless of the preference misalignment between the players. This result is in stark contrast to the result for exogenously given fraction of informed users in Lemma 1. The ability to limit access to information mitigates the reluctance to implement an information system due to preference misalignment between the players. Technically speaking, this difference arises because the expected payoff of the manager at the optimal fraction  $\hat{x} = k$  is always convex in the users' beliefs. As a result, the manager always finds it optimal to implement a system that releases a perfectly informative signal to the optimally chosen fraction of users.<sup>26</sup>

Regulators often gravitate toward rules that ensure equal access to information for all agents in an economy. To consider the efficiency of such rules, I compare two alternative regimes: mandated dissemination (MD), under which the manager is required to disclose information to all users, i.e.,  $x = 1$  is exogenously set,<sup>27</sup> and unregulated dissemination (UD), under which the manager can optimally choose the fraction of informed users in her

---

<sup>25</sup>The uninformed users, by default, have response coefficient  $\frac{d\hat{a}(\mu_0)}{ds} = 0$ .

<sup>26</sup>In my model, the manager implements an information system solely for information dissemination. The results will qualitatively hold if the manager were to use the output of the information system for other purposes, as long as the manager sufficiently cares about the dissemination. To illustrate, consider for example an ex post payoff  $v = -\rho(A - k\omega - (1 - k)\bar{\omega})^2 - (1 - \rho)(d - \omega)^2$ , where  $d$  is an operating decision that the manager takes after observing the signal and  $\rho \in (0, 1)$  represents the importance of dissemination. Then, under mandated dissemination (MD), the preference alignment threshold above which the manager implements a perfectly revealing information system decreases to  $\frac{2\rho-1}{2\rho} < \frac{1}{2}$  but remains strictly positive as long as  $\rho > \frac{1}{2}$ . The results pertaining to unregulated dissemination (UD) remain unaffected.

<sup>27</sup>In many cases, even if the regulator sets  $x = 1$ , the actual fraction of informed users is strictly lower as some users will not observe the information (for example, unsophisticated investors). Considering this possibility does not change qualitatively the results.

own best interest.

## Proposition 2

(i) If  $k \leq \frac{1}{2}$ , UD Pareto dominates MD.

(ii) If  $k > \frac{1}{2}$ , the manager is better off under UD, but all users are at least weakly better off under MD.

Under UD, the manager can choose any fraction  $x \in [0, 1]$ , but she finds it optimal to limit access to information to some users (i.e., sets  $\hat{x} = k < 1$ ). Hence, by revealed preference, she always prefers UD over MD. The users' preference is less obvious. At the heart of the comparison lie the observations that the manager sets an infinite precision whenever she chooses to implement an information system and that, given their quadratic loss payoff, the users benefit from fully revealing information.

By Lemma 1 and Proposition 1, if  $k \leq \frac{1}{2}$ , then the manager implements an information system only under UD. The informed users are better off under UD (because it is the only regime under which the manager implements an information system). The uninformed users are indifferent (because they do not observe information under either regime). As a result, UD Pareto dominates MD. However, if  $k > \frac{1}{2}$ , the manager implements a perfectly revealing information system under both regimes. The informed users are indifferent (because they perfectly observe  $\omega$  under both regimes), while the users who are uninformed under UD are better off under MD (because it allows them to observe information).

Paradoxically, Proposition 2 shows that, when the players preferences over the actions are *misaligned*, their preferences over regimes are *aligned*, and vice versa. If regulators believe that firm managers and users have very different objectives, then by part (i), they should gravitate to UD, because it Pareto dominates MD. If  $k > \frac{1}{2}$ , Pareto ranking of the regimes is not possible, but MD ensures all users are at least weakly better off. Hence, if regulators care exclusively about the welfare of the users and believe the objectives of the managers and the users are sufficiently aligned, they should gravitate to MD.

The result of Proposition 2 may be surprising at first, because conventional wisdom would say that, if the preferences of the players are misaligned, then there is need for the regulator to intervene as players will not arrive at a socially efficient result on their own, and vice

versa. However, my model predicts exactly the opposite—the regulator’s intervention when the players objectives are misaligned may suppress socially beneficial information acquisition. A natural question that arises is what is the socially optimal fraction of informed users, which I examine in subsection 4.4 after I generalize the model to costly information system implementation.

## 4 Costly Persuasion

In Section 3, the optimal precision under MD has a bang-bang character and jumps to infinity at  $k = \frac{1}{2}$ . This result does not seem particularly descriptive. Therefore, in this subsection, I extend the results by introducing a cost borne by the manager associated with the implementation of the information system.<sup>28</sup> I show that when implementation is costly the manager sacrifices some amount of precision, and, under standard regularity conditions, her choice is an interior solution. This allows (i) for analyzing how the ability of the manager to limit access to information affects the information precision and (ii) for providing new predictions.

I assume that the manager bears an implementation cost  $C(c, \beta)$  with

$$C_c(\cdot) \geq 0, \quad C_\beta(\cdot) \geq 0, \quad C_{cc}(\cdot) \geq 0, \quad C_{\beta\beta}(\cdot) \geq 0, \quad C_{c\beta}(\cdot) > 0.$$

The parameter  $c > 0$  represents the cost of information system technology in the economy and might be a function of the level of competition in that market. I assume  $c$  is exogenously given and refer to it as the “information cost” or just “cost.” To ensure interior solutions I assume that  $\lim_{\beta \rightarrow \infty} C_\beta(c, \beta) \rightarrow \infty$  and  $\lim_{\beta \rightarrow 0} C_\beta(c, \beta) = 0$  for any cost.<sup>29</sup> The total implementation cost is zero whenever information is cost-free or information system is completely imprecise, i.e.,  $\lim_{c \rightarrow 0} C(c, \beta) = \lim_{\beta \rightarrow 0} C(c, \beta) = 0$ . Lastly, I assume that the implementation cost is independent of the fraction of users observing the signal.<sup>30</sup>

---

<sup>28</sup>Information system implementation costs are frequently observed in practice. For example, providing a guidance requires hiring an economist, inventory management requires purchasing software, evaluating an asset’s fair value requires paying for an appraisal, etc.

<sup>29</sup>This is the case for many commonly used cost functions, including the quadratic cost function  $\frac{c\beta^2}{2}$  as a special case. All results hold qualitatively with few minor adjustments for cost functions with  $\lim_{\beta \rightarrow 0} C_\beta(c, \beta) > 0$  (linear cost function  $c\beta$  as a special case), i.e., for which the marginal cost from implementing a system with even very small precision is positive. Whenever applicable, I will outline in a footnote the minor adjustments needed under the assumption that  $\lim_{\beta \rightarrow 0} C_\beta(c, \beta) > 0$ .

<sup>30</sup>It seems reasonable to think that, once the information system is implemented, it does not matter

## 4.1 The Optimal Precision

Similar to the analysis with cost-free implementation, I start by considering the manager's problem when  $x \in [0, 1]$  is exogenously given. Let

$$\Delta x \equiv |x - k|$$

denote the distance between the fraction of informed users and the preference misalignment. At date 1, the manager chooses

$$\beta^c(x) \in \arg \max_{\beta} V(x, \beta) - C(c, \beta), \quad (8)$$

where the superscript “c” denotes costly information system implementation.<sup>31</sup>

**Lemma 2** *Suppose  $x \in [0, 1]$  is exogenously given and the manager bears an implementation cost  $C(c, \beta)$ . The manager then implements an information system with precision  $\beta^c(x) \in (0, \infty)$  if and only if  $k > \frac{x}{2}$ .  $\beta^c(x)$  is decreasing in  $c$  and in  $\Delta x$  and increasing in  $\frac{1}{\alpha}$  and in  $k$ .*

Similar to the results in Lemma 1 for cost-free implementation, the manager implements an information system if and only if her preferences are sufficiently aligned with those of the users.<sup>32</sup> The cutoff for implementation ( $k > \frac{x}{2}$ ) does not depend on the cost. If  $c \rightarrow 0$ , then  $\beta^c(x) \rightarrow \infty$  and therefore the results from the preceding section are recouped as a special case. However, as long as  $c > 0$ , the manager finds it optimal to sacrifice some precision. The higher the cost, the less precise the information system that the manager implements. In the limit, as  $c \rightarrow \infty$ ,  $\beta^c(x) \rightarrow 0$ , i.e., if the implementation is extremely costly, the signal will be uninformative.

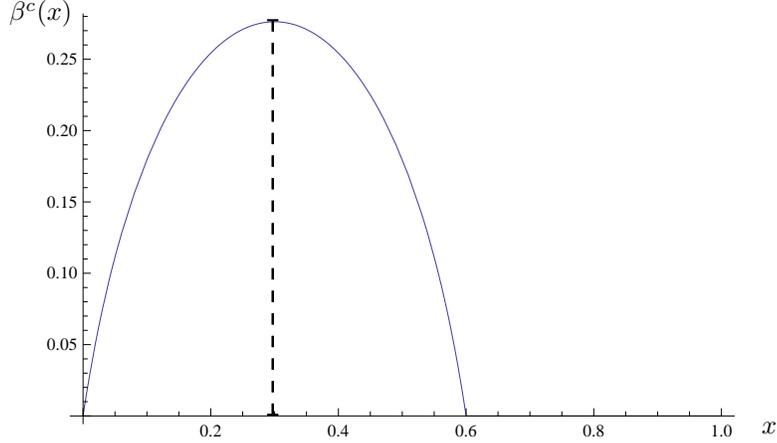
The closer the exogenous fraction  $x$  to the preference alignment  $k$ , the higher the precision of the information system that the manager implements. This is graphically shown in Figure

---

to how many users its output is communicated. Assuming otherwise, i.e., that the implementation cost is increasing in the number of the users to which the signal is conveyed (which is the reasonable alternative assumption), would only mechanically facilitate finding that it is optimal for the manager to restrict access to information without qualitatively changing the results.

<sup>31</sup>The optimal precision  $\beta^c(x)$  depends, in addition to  $x$ , on other exogenous parameters  $c$ ,  $\alpha$  and  $k$ . I depress them to avoid clutter.

<sup>32</sup>If  $\lim_{\beta \rightarrow 0} C_{\beta}(c, \beta) > 0$ , i.e., if the implementation of an information system with even very small precision is costly, then the manager will implement an information system only if  $c$  is below a certain threshold. It can be shown that the threshold is decreasing in  $\Delta x$ . Analysis available upon request.



**Figure 2:** Optimal precision as a function of  $x$   
Numerical example with  $C(c, \beta) = \frac{c\beta^2}{2}$ ,  $\alpha = 1$ ,  $c = 0.2$ ,  $k = 0.3$

2. The intuition for this result is that, when  $x = k$ , the players' preference misalignment is minimized, and this provides stronger incentives for the manager to gather information. The more aligned the manager's preferences with those of the users, the more precise the information system she implements, so that the users' actions will be more in line with the realization of  $\omega$ . As shown in Lemma 2, the signal precision is increasing in  $\frac{1}{\alpha}$ , the prior variance of  $\omega$ . Further, the optimal precision does not depend on  $\bar{\omega}$ , i.e., by disseminating information about  $\omega$ , the manager cannot persuade the users to take an action close to  $\bar{\omega}$ .

As a next step, I consider the ability of the manager to limit access to information to a subset of users by optimally choosing  $x$ . At date 1, the manager's full-fledged optimization problem is:

**Program  $\mathcal{P}^c$ :**

$$\max_{\beta \geq 0, x \in [0,1]} V(x, \beta) - C(c, \beta). \quad (9)$$

Let  $(x_{UD}^c, \beta_{UD}^c)$  denote the solution to this program. The next result extends Proposition 1 to the costly implementation setting.

**Proposition 3** *Suppose the manager bears an implementation cost  $C(c, \beta)$ . Under UD, for any  $k$ , it is optimal for the manager to implement an information system with precision  $\beta_{UD}^c \equiv \beta^c(x = k)$  and to disseminate the signal to a fraction  $x_{UD}^c = k$  of informed users.*

The key finding here is that the optimal subset of informed users is unaffected by the cost. The choice of  $x$  is a device that induces the right aggregate signal-response coefficient of  $k$ , and that logic is unaffected by any implementation costs. However, when implementation is costly, the manager implements an information system that provides a noisy signal. The comparative statics of the optimal precision with respect to  $\frac{1}{\alpha}$ ,  $c$ ,  $k$  and  $\bar{\omega}$  are similar in nature to those of  $\beta^c(x)$ .

## 4.2 Regime Preferences

Straightforward application of Lemma 2 reveals that under MD the manager implements an information system with precision

$$\beta_{MD}^c \equiv \beta^c(x = 1) > 0$$

if  $k > \frac{1}{2}$  and zero otherwise. By Lemma 2, the information system implemented under UD is more precise than the one implemented under MD:

$$\Delta\beta \equiv \beta_{UD}^c - \beta_{MD}^c > 0. \tag{10}$$

Similar to the cost-free setting, the manager always prefers (by revealed preference) UD, and the uninformed users are always weakly better off under MD.<sup>33</sup> However, the comparison with regards to the informed users is more complicated and requires taking into account that the equilibrium precisions under both regimes are different as shown in (10) and that the users want as precise information as possible. The analysis shows that the informed users are always strictly better off under UD, because it ensures they observe a more precise signal. The result below extends Proposition 2 to the costly implementation setting.

**Proposition 4** *Suppose the manager bears an implementation cost  $C(c, \beta)$ .*

(i) *If  $k \leq \frac{1}{2}$ , UD Pareto dominates MD.*

(ii) *If  $k > \frac{1}{2}$ , the uninformed users are strictly better off under MD, while the informed users and the manager are strictly better off under UD.*

---

<sup>33</sup>If  $k > \frac{1}{2}$ , it is the only regime that enables them to observe a signal, while if  $k \leq \frac{1}{2}$ , then they are indifferent because they do not observe a signal under either regime.

As before, if  $k \leq \frac{1}{2}$ , UD Pareto dominates MD. Hence, if regulators believe the preferences of firm managers and users in an economy are sufficiently misaligned, they should not enforce equal access to information of all potential users. If  $k > \frac{1}{2}$ , there is an additional dimension to the disagreement issue discussed in the cost-free setting, because now even the different types of users, endogenously divided into informed and uninformed ones, prefer different regimes.

### 4.3 Welfare Analysis with Sufficiently Aligned Preferences

Proposition 4 shows that, when the preferences of the players are sufficiently aligned, neither of the regimes ensures all users are at least weakly better off simultaneously. In this subsection, I conduct welfare analysis to evaluate under which regime the users are better off at an *aggregate* level when  $k > \frac{1}{2}$ . Let

$$\phi^c(x) \equiv E_{t=1} [u(\hat{\alpha}(s|\beta^c(x)), \omega)] - E_{t=1} [u(\hat{\alpha}(\mu_0), \omega)]$$

denote a representative user's expected gain of becoming informed (hereafter, "information gain"). Then,

$$\begin{aligned} W^c(x) &= \int_0^x E_{t=1} [u(\hat{\alpha}(s|\beta^c(x)), \omega)] di + \int_x^1 E_{t=1} [u(\hat{\alpha}(\mu_0), \omega)] di \\ &= \int_0^1 E_{t=1} [u(\hat{\alpha}(\mu_0), \omega)] di + \int_0^x \phi^c(x) di \\ &= \underbrace{E_{t=1} [u(\hat{\alpha}(\mu_0), \omega)]}_{\text{base welfare}} + \underbrace{x\phi^c(x)}_{\text{aggregate gain}} \end{aligned} \quad (11)$$

is the aggregate users' welfare.<sup>34</sup> The base welfare represents the aggregate payoff of all users when the manager does not implement an information system. The aggregate gain is the information gain of the informed users, on an aggregate level. While the base welfare is independent of  $x$ , the comparative statics of the aggregate gain with respect to  $x$  is ambiguous. As seen from (15), there are two effects: (i) a direct effect—as  $x$  increases, more users benefit from information, and (ii) an indirect effect—the fraction  $x$  affects the information gain indirectly through the optimal precision.

**Corollary 1**  $\phi^c(x)$  is decreasing in  $\Delta x \equiv |x - k|$ .

<sup>34</sup>The gain  $\phi^c(x)$  and the aggregate welfare  $W^c(x)$  depend, in addition to  $x$ , on other exogenous parameters:  $c$ ,  $\alpha$  and  $k$ . I depress them to avoid clutter.

The intuition behind this result is that the information gain is increasing in the precision, which is single-peaked at  $x = k$  by Lemma 2. As a result, the informed users are better off when the manager discloses information only to a fraction  $k$  of users. Equation (15) and Corollary 1 imply that broader dissemination of information is not always better for the users on an aggregate level. To compare the users' welfare under both regimes, let

$$\phi_{MD}^c \equiv \phi^c(x = 1) \quad \text{and} \quad \phi_{UD}^c \equiv \phi^c(x = k)$$

denote the information gain of a representative informed user and

$$W_{MD}^c \equiv W^c(x = 1) \quad \text{and} \quad W_{UD}^c \equiv W^c(x = k)$$

the aggregate welfare of the users under MD and UD, respectively. Comparing UD and MD on an aggregate level requires signing the welfare differential

$$\Delta W \equiv W_{UD}^c - W_{MD}^c = \underbrace{k\phi_{UD}^c}_{\text{aggregate gain UD}} - \underbrace{\phi_{MD}^c}_{\text{aggregate gain MD}}$$

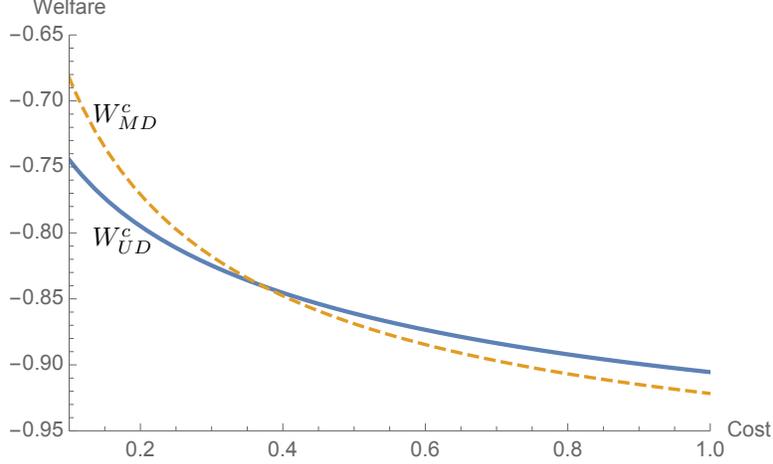
The comparison of the aggregate gains is affected by two countervailing effects: (i) *an omission effect*—under UD, the proportion of informed users is lower than the proportion under MD ( $x_{UD}^c = k < 1$ )—and (ii) *a precision effect*—by Lemma 2, the signal precision under UD is higher than the one under MD. As a result, by Corollary 1, the information gain under UD is larger than the one under MD:

$$\Delta\phi \equiv \phi_{UD}^c - \phi_{MD}^c > 0. \tag{12}$$

The next result presents sufficient conditions for the aggregate welfare under UD to exceed the one under MD.

**Proposition 5** *Suppose  $k > \frac{1}{2}$  and the manager bears an implementation cost  $C(c, \beta) = \frac{c\beta^2}{2}$ . Then there exist  $\tilde{k} \in (\frac{1}{2}, 1)$ , such that  $W_{UD}^c \geq W_{MD}^c$  if  $k \leq \tilde{k}$ , and  $c$  is sufficiently high.*

This result is graphically shown in Figure 3. As  $c \rightarrow 0$ , the manager sets the same (infinite) precision under UD and MD. Hence the gain from information of a representative informed user under both regimes is the same, and  $\Delta W \propto k - 1 < 0$  so that the users prefer MD, in aggregate. As  $c$  increases, the optimal precision and, as a result, the respective aggregate gain decrease. As  $c \rightarrow \infty$ , the aggregate gain under both regimes reaches zero. However, as I



**Figure 3:** Users' welfare comparison  
 Numerical example with  $C(c, \beta) = \frac{c\beta^2}{2}$ ,  $\alpha = 1$ ,  $k = 0.55$ .  
 $W_{UD}^c$  ( $W_{MD}^c$ ) is the users' welfare under unregulated (mandated) dissemination

show in the proof, the gain under MD does so faster than the one under UD if the preference alignment is sufficiently bounded away from one.<sup>35</sup> As a result, for some sufficiently high cost, the aggregate gain under MD is larger than the aggregate gain under UD.

#### 4.4 Regulated Information Dissemination

A natural question that arises is what is the socially optimal fraction of informed users. To answer this question, I assume that a benevolent regulator (“he”) chooses the subset of users who get to observe the signal prior to the manager’s choice of information system and call this regime regulated dissemination (“RD”). The regulator chooses  $x$  to maximize the aggregate expected payoffs of all players, subject to the constraint that the information system precision is chosen by the manager in her own interest:

$$\begin{aligned} & \max_{x \in [0,1]} \lambda[V(x, \beta^c(x)) - C(c, \beta^c(x))] + W^c(x) \\ \text{subject to} & \quad \beta^c(x) \equiv \arg \max_{\beta} V(x, \beta) - C(c, \beta). \end{aligned}$$

The parameter  $\lambda \geq 0$  represents the weight that the regulator puts on the manager’s welfare. It is straightforward that, if the regulator cares only about the manager’s welfare (i.e.,  $\lambda \rightarrow \infty$ ), he will choose the same fraction of informed users that the manager would have set

<sup>35</sup>When the preference alignment is sufficiently close to one, the aggregate gains under both regimes are of similar magnitude and decrease in  $c$  at the same rate.

in her own interest ( $x_{RD}^c = k$ ). However, if the regulator cares sufficiently about the users' welfare, the socially beneficial level of  $x$  may go beyond  $k$  and, under certain conditions, may even reach one.

**Corollary 2** *Suppose  $\lambda \leq 1$  and the manager bears an implementation cost  $C(c, \beta) = \frac{c\beta^2}{2}$  with  $c \geq 0$ . Then the socially beneficial fraction of informed users  $x_{RD}^c \in [k, 1]$  is decreasing in  $c$ .*

When the players' preferences are sufficiently misaligned, the regulator wants to restrict access to information for some users even when he does not care at all about the manager's welfare (i.e.,  $\lambda = 0$ ), because he wants to ensure the manager has incentives to implement an information system and disseminate socially valuable information.<sup>36</sup> However, when the players' preferences are sufficiently aligned and information acquisition is cost-free, the regulator chooses a corner solution for  $x$  and enforces equal access to information. The rationale behind this observation is that when  $c \rightarrow 0$  the gain in collective users' welfare is larger than the loss in manager's payoff caused by increasing  $x$  beyond  $k$ .<sup>37</sup> Introduction of costs changes the corner solution character of the socially beneficial fraction of informed users. Put differently, unlike the fraction of informed users set by the manager in her own interest, the fraction  $x$  that the regulator enforces is decreasing in the implementation cost. The reason for this finding is that, as  $c$  increases, the manager chooses a lower precision. Then, to provide incentives for the manager to increase the precision of the socially valuable information, even the users themselves, at some prior state, would collectively agree to some  $x < 1$ , as long as they are behind the veil of ignorance (Harsanyi, 1955), i.e., before each learns whether he will be included in the group of informed users.

---

<sup>36</sup>If  $k \leq \frac{1}{2}$ , then  $2k \leq 1$ . To ensure that the manager implements an information system, the planner needs to make sure that  $x < 2k \leq 1$ . This observation confirms the result in Proposition 4 (i).

<sup>37</sup>When  $k > \frac{1}{2}$ , the implementation constraint is satisfied for any  $x$ , and with  $c = 0$  the manager implements a perfectly revealing information system. Hence the derivative of the regulator's objective function with respect to  $x$  reads  $\frac{2\lambda(k-x)}{\alpha} + \frac{1}{\alpha}$ . If  $x < k$ , both terms are positive, i.e., all players benefit from increasing  $x$ . If  $x > k$ , then the first term is negative and represents the decrease in manager's payoff from increasing the fraction of informed users beyond  $k$ . The second term represents the gain in aggregate users' welfare from an increase in  $x$ . If  $\lambda \leq 1$  and  $k > \frac{1}{2}$ , then  $2\lambda(k-x) + 1 > 2k - 1 > 0$ , i.e., the gain in users' welfare is larger than the loss in manager's payoff.

## 5 Discussion of Model Variations

In this section, I discuss the robustness of the model with respect to different variations. I focus on pure strategies and return to the maintained assumption in Section 3 that the implementation of an information system is cost-free.

### 5.1 Delayed or Revised Choice of Informed Users at Date 2

In the main part of the article, the manager chooses the fraction of informed users at date 1 and cannot revise her choice later on, i.e., she has commitment power. This comports with the commitment assumption of Bayesian persuasion models and reflects the fact that managers often must choose the subset of “privileged” users early on (e.g., to provide access to a database in advance; send invitations to conference calls, meetings, or speeches ahead of time to ensure participation; reserve a conference venue) and deviation at a later date is impractical or costly. Now I relax this assumption and allow the manager to delay or revise (increase or decrease) the choice of  $x$  after observing the signal  $s$ . I assume that the manager’s choice of  $x$  is not observed by the uninformed users.<sup>38</sup> At date 2, the manager’s delayed or revised choice of subset of users “in the know,” for given  $\beta$  and  $s$ , is:

$$x_{t=2}(s|\beta) \in \arg \max_{x \in [0,1]} E_{t=2}[v(\hat{A}(x, s, \beta), \omega)|s, \beta].$$

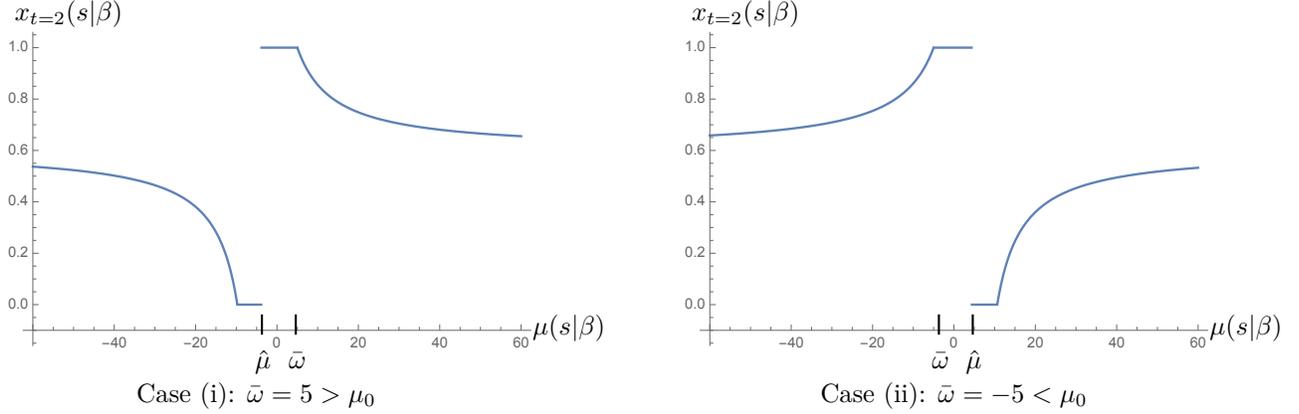
**Observation 1** *Suppose that the manager can delay (or revise) her choice of unobservable  $x$  to date 2 and that  $\bar{\omega} = \mu_0$ . Then, for any signal realization,  $x_{t=2}(s|\beta) = k$ .*

When the uninformed users act based on their prior beliefs, their actions equal the value toward which the manager is biased (in the knife-edge case where  $\bar{\omega} = \mu_0$ ). Hence the manager chooses  $x = k$  because this is the fraction that minimizes to zero the difference between the actual aggregate action of the users, as stated in equation (5), and the expectation of the manager’s most preferred action, as stated in equation (2), for any signal realization. Put differently, if the manager could delay or revise her choice at date 2, she would choose exactly the same fraction as the one she would choose if she had to commit at date 1. Then, the analysis from the main part of the article remains the same for this special case.

For the remainder of this subsection, I assume that  $\bar{\omega} \neq \mu_0$ . Furthermore, I assume that the users (i) are not ordered on  $[0, 1]$  and (ii) realize they are included “in the know” only

---

<sup>38</sup>At the end of this subsection I briefly discuss the case of observable  $x$ .



**Figure 4:** Illustration of Lemma 3 (unobservable delayed or revised choice of  $x$ )  
 Numerical example with  $k = 0.6$ ,  $\mu_0 = 0$ ,  $\alpha = 1$ ,  $\beta \rightarrow \infty$

after receiving the signal. Hence, if a user ends up being uninformed, he only infers that the signal was such that the manager did not choose  $x = 1$ .

**Lemma 3** *Suppose that the manager can delay (or revise) her choice of unobservable  $x$  to date 2 (after observing  $s$ ) and that  $\bar{\omega} \neq \mu_0$ . Then, at date 2, the manager disseminates the signal to a fraction  $x_{t=2}(s|\beta) \neq k$  of users. At date 3, the informed users take an action that equals  $\mu(s|\beta)$ , and the uninformed users take an action that equals  $\hat{\mu} \equiv E[\omega | x_{t=2}(s|\beta) \neq 1]$ .*

(i) *If  $\bar{\omega} > \mu_0$ , then  $\hat{\mu} \in (-\infty, \mu_0)$ . For any signal such that  $\mu(s|\beta) < \hat{\mu}$ , the revised fraction is  $x_{t=2}(s|\beta) = \max \left\{ 0, k - (1 - k) \frac{(\bar{\omega} - \hat{\mu})}{(\hat{\mu} - \mu(s|\beta))} \right\} \in [0, k)$ . Otherwise,  $x_{t=2}(s|\beta) = \min \left\{ 1, k - (1 - k) \frac{(\bar{\omega} - \hat{\mu})}{(\hat{\mu} - \mu(s|\beta))} \right\} \in (k, 1]$ .*

(ii) *If  $\bar{\omega} < \mu_0$ , then  $\hat{\mu} \in (\mu_0, \infty)$ . For any signal such that  $\mu(s|\beta) < \hat{\mu}$ , the revised fraction is  $x_{t=2}(s|\beta) = \min \left\{ 1, k - (1 - k) \frac{(\bar{\omega} - \hat{\mu})}{(\hat{\mu} - \mu(s|\beta))} \right\} \in (k, 1]$ . Otherwise,  $x_{t=2}(s|\beta) = \max \left\{ 0, k - (1 - k) \frac{(\bar{\omega} - \hat{\mu})}{(\hat{\mu} - \mu(s|\beta))} \right\} \in [0, k)$ .*

The results of Lemma 3 are illustrated in Figure 4. The manager deviates from  $x = k$  but restricts the access to information for a large subset of signal realizations. To understand the intuition, first consider the case where the manager is biased toward high values ( $\bar{\omega} > \mu_0$ ). Note that, when the informed users take actions that equal  $\mu(s|\beta)$  and the uninformed users

take actions that equal their belief  $\hat{\mu}$ , the manager wants to minimize the gap between the conditional expectation of her preferred action,

$$A^*(s|\beta) \equiv E_{t=2}[A^*(\omega)|s, \beta] = k\mu(s|\beta) + (1 - k)\bar{\omega}, \quad (13)$$

and the actual aggregate action,

$$\hat{A}(x, s, \beta) = x\mu(s|\beta) + (1 - x)\hat{\mu}. \quad (14)$$

Suppose that the uninformed users' expectation of the state is below the action toward which the manager is biased,  $\hat{\mu} < \bar{\omega}$ . (As I show later, this is the case in equilibrium.) Then, if the observed signal is such that  $\mu(s|\beta) < \hat{\mu}$ , disclosing the signal to a subset  $x = k$  of users will result in lower actual aggregate action than the one the manager prefers. To minimize this gap, the manager disseminates the signal to fewer users ( $x < k$ ). This is because the actions that the uninformed users take are closer to the value toward which the manager is biased, i.e.,  $\mu(s|\beta) < \hat{\mu} < \bar{\omega}$ . The higher the observed signal (but still lower than the one leading to homogenous beliefs across users, i.e.,  $\mu(s|\beta) = \hat{\mu}$ ), the higher the actions that the informed users take and hence the higher  $\hat{A}(x, s, \beta)$ . However, at the same time,  $A^*(s|\beta)$  also increases, and it does so at a higher rate. Hence, to minimize the gap between the actual and the preferred action, the manager needs to further restrict the access to information (as this results in more uninformed users taking a higher action that is closer to the value towards which the manager is biased, i.e.,  $\mu(s|\beta) < \hat{\mu} < \bar{\omega}$ ). However, given that  $x \geq 0$ , eliminating the gap may not be feasible. In such a case, the best the manager can do is to withhold the signal from everyone.

Similar arguments hold when the observed signal is such that  $\mu(s|\beta) > \hat{\mu}$ . Then, given that  $\bar{\omega} > \hat{\mu}$ , the manager wants to disseminate the signal more broadly ( $x > k$ ) because the actions of the informed users are closer to the value towards which she is biased. The closer the beliefs of the informed users to the value towards the manager is biased,  $\bar{\omega}$ , the more broadly the manager disseminates the signal. When the realized signal is such that  $\mu(s|\beta) = \bar{\omega}$ , the manager's preferred action collapses to  $A^*(s|\beta) = \bar{\omega}$ , and the manager disseminates the signal to everyone. For any signal that leads to  $\mu(s|\beta) \in (\hat{\mu}, \bar{\omega})$ , minimizing the gap between actual and preferred action requires setting  $x > 1$ . This, however, is not feasible, and so the best the manager can do is to publicly disseminate the signal. Lastly,

in the knife-edge case where the beliefs of the informed and uninformed users coincide, i.e.,  $\mu(s|\beta) = \hat{\mu}$ , the manager is indifferent because  $\hat{A}(x, s, \beta) = \hat{\mu}$  for any  $x$ .<sup>39</sup>

I next specify the beliefs of the uninformed users. In equilibrium, they satisfy  $\hat{\mu} = E[\omega|x_{t=2}(s|\beta) \neq 1]$ . As shown in the proof, there exists a unique and interior  $\hat{\mu} < \mu_0$  that satisfies this condition (and, given the assumption  $\bar{\omega} > \mu_0$ , the conjecture  $\hat{\mu} < \bar{\omega}$  is correct). When the manager is biased towards high values, the uninformed users revise their beliefs downward upon realizing they are uninformed.

Turning to the case where the manager is biased toward low values ( $\bar{\omega} < \mu_0$ ) the arguments from the preceding discussion are reversed. In equilibrium, the uninformed users, upon realizing the signal realization is such that  $x < 1$ , revise their prior upward, i.e.,  $\hat{\mu} > \mu_0$ . If the observed signal is such that  $\mu(s|\beta) > \hat{\mu}$ , disclosing the signal to a subset of  $x = k$  results in a lower actual aggregate action than the one the manager prefers, and so the manager disseminates the signal to fewer users ( $x < k$ ) because the actions of the uninformed users are closer to the action that the manager prefers:  $\mu(s|\beta) > \hat{\mu} > \bar{\omega}$ . In contrast, when the signal is such that  $\mu(s|\beta) < \hat{\mu}$ , the manager disseminates the signal to more users ( $x > k$ ), as the actions of uninformed users are further away from the action that the manager prefers:  $\hat{\mu} > \mu(s|\beta) > \bar{\omega}$ .

At date 1, the manager takes into account that the fraction of informed users is  $x_{t=2}(s|\beta)$  and chooses  $\beta$  to maximize her expected payoff:

$$\max_{\beta \geq 0} E_{t=1}[v(\hat{A}(x_{t=2}(s|\beta), s, \beta), \omega)].$$

**Proposition 6** *Suppose that the manager can delay (or revise) her choice of unobservable  $x$  to date 2. Then, at date 1, the manager implements a perfectly revealing information system, i.e.,  $\beta \rightarrow \infty$ .*

If the manager can delay (or revise) the choice of  $x$  to date 2, she induces an actual aggregate action that is as close as possible to her expected bliss point. The more precise the signal, the more precise the expectation of the manager about her bliss point, and hence the more efficient the alignment with the aggregate action of the users. Therefore it is optimal for the manager to implement a perfectly revealing information system.<sup>40</sup>

<sup>39</sup>With a continuous distribution, the probability that  $\mu(s|\beta) = \hat{\mu}$  is zero.

<sup>40</sup>Technically speaking, this result occurs because the manager's payoff conditional on the signal is always lower than its concave closure.

In the preceding discussion, I focus on the case where the delayed (or revised) choice of  $x$  is unobservable. In the observable case, the manager's choice of  $x$  not only determines the subset of privileged users who are informed of the signal directly, but may also indirectly convey information about the signal to the rest of the users. Therefore, the communication game between the manager and the users becomes complex and is beyond the scope of this article. What can be shown is that, despite the informational spillover, there does not exist a fully revealing equilibrium in which all users either directly observe, or infer the signal.<sup>41,42</sup>

## 5.2 Unobservable Signal Precision

In the main part of the article, I assume that the manager's choice of precision is observable. Now I relax this assumption and reexamine the prior results. In this subsection, I return to the baseline model in which the manager commits to the fraction  $x$  at date 1. If the manager's payoff is commonly known and the fraction  $x$  is observable, the users can conjecture the manager's choice. Then, the same equilibrium described in Section 3 persists, but it is no longer unique.

**Observation 2** *Suppose the manager's choice of precision  $\beta$  is not observable but the fraction  $x$  is observable. Then, there exist two equilibria:*

- (i) *an informative equilibrium in which the manager sets  $x = k$  and  $\beta \rightarrow \infty$ ;*
- (ii) *an uninformative equilibrium in which the manager does not implement an information system.*

The formal proof is omitted because it follows from the discussion below. Consider a conjecture of the precision level  $\beta^o$  made by the users. If  $\beta^o > 0$ , the manager's objective

---

<sup>41</sup>I thank an anonymous referee for pointing this out.

<sup>42</sup>The proof is by contradiction. Briefly, suppose that the manager's strategy of choosing  $x_{t=2}(s|\beta)$  is invertible in  $s$  everywhere. This is a necessary and sufficient condition for a fully revealing equilibrium, so the uninformed users infer the signal realization and the actual aggregate action collapses to  $\mu(s|\beta)$ . The manager's preferred action,  $A^*(s|\beta)$ , as defined in (13), is a weighted average of  $\mu(s|\beta)$  and the value  $\bar{\omega}$  towards which the manager is biased. Given the uninformed users' strategy, a manager who has observed  $s$  such that  $\mu(s|\beta) \neq \bar{\omega}$ , is better off choosing a fraction from which the uninformed users will infer that the signal is  $s' \neq s$ , where  $\mu(s'|\beta)$  lies between  $\mu(s|\beta)$  and the manager's preferred action  $A^*(s|\beta)$ . In such a case, the actual aggregate action will be a weighted average of  $\mu(s|\beta)$  and  $\mu(s'|\beta)$  and therefore closer to the manager's preferred action. This violates the postulated fully revealing equilibrium. Formal proof is provided in the Appendix.

is increasing in  $\beta$ , and she implements infinite precision.<sup>43</sup> In other words, as long as the users react to the signal ( $\beta^o > 0$ ), the manager has incentives to communicate a perfectly informative signal ( $\beta \rightarrow \infty$ ) because this is the only way she can minimize the noise in the users' actions. In equilibrium, the users' conjecture has to be true. Hence there exists an informative equilibrium in which the manager sets  $x = k$  and  $\beta \rightarrow \infty$ .

To see why an additional uninformative equilibrium arises, consider a conjecture of precision level  $\beta^o = 0$  made by the users. Regardless of the precision of the signal that the manager sets, the users will ignore the signal. In the absence of implementation costs, the manager is indifferent between setting any precision and so (as assumed in the case of indifference) chooses  $\beta = 0$ .

### 5.3 Sequential Users' Actions

The maintained assumption in the main part of the article is that the users act simultaneously. While this may be true in the case of a shareholder vote, it need not always be the case. Analysts, divisional managers and market participants may act sequentially, releasing forecasts or making decisions at different times.<sup>44</sup> In this subsection, I extend the analysis by allowing the users to act at two consecutive dates: 3.1 and 3.2. I return to the assumptions that the manager commits to the fraction  $x$  and the precision  $\beta$  is observable. Within the confines of the baseline model, the uninformed users would choose to act at date 3.2, hoping to observe the actions of the informed users and update their beliefs about  $\omega$ . With any user's payoff independent of the other users' actions and the timing of his own action, the informed users will be indifferent between acting at date 3.1 or at date 3.2. The equilibrium depends on the tie-breaking rule.

**Observation 3** *Suppose that, when indifferent, the informed users act late. Then, in equilibrium, all users act simultaneously at date 3.2.*

In this case, all results described in Section 3 hold. However, if, when indifferent, the informed users act early, i.e., at date 3.1, then (in the absence of additional frictions) the uninformed

---

<sup>43</sup>Specifically, the derivative with respect to  $\beta$  of the manager's objective is  $\frac{\partial V(x,\beta)}{\partial \beta} = \frac{(\beta^o)^2 x^2}{\beta^2 (\alpha + \beta^o)^2} \geq 0$  and so the manager sets  $\beta \rightarrow \infty$ .

<sup>44</sup>For example, O'Brien, McNichols and Lin (2005) empirically show that financial analysts differ in the speed with which they release forecasts.

users will learn the observed signal, and the equilibrium under UD will be identical to the one under MD. For the remainder of this subsection, I focus on the behavior of the uninformed users and assume that the informed users act early for exogenous reasons. I refer to the updating of beliefs by uninformed users after observing the actions of informed users as “information spillover.” In many settings, the assumptions that (i) the payoff of the users is independent of the actions of the other users, (ii) there is no cost for delaying the action and (iii) the uninformed users perfectly observe the actions of their peers do not hold. In the discussion below, I relax these assumptions and reexamine the results.

### 5.3.1 Reward for Relative Performance and Cost of Action Delay

In many cases, the payoff of the users depends not only on their own actions but also on the actions of the other users.<sup>45</sup> Analysts, divisional managers and government officials may care not only about the accuracy of their forecasts and decisions but also about their performance relative to that of their peers. For example, Hong, Kubik and Solomon (2000) and Wu and Zang (2009) empirically document that financial analysts who issue the least accurate forecasts are more likely to be fired and less likely to be promoted. In addition, the users of information may incur costs for delaying their actions such as lost clientele or missed opportunities. To analyze such settings, I assume in this subsection that the ex-post payoff of user  $i$  is

$$\begin{aligned} u_i(a_i, A, \omega, t) &= -(1 - \pi) \underbrace{(a_i - \omega)^2}_{\text{accuracy}} - \pi \underbrace{((a_i - \omega)^2 - (A - \omega)^2)}_{\text{relative performance}} - \mathbb{1}_{t_i=3.2} \underbrace{D}_{\text{delay cost}}, \\ &= -(a_i - \omega)^2 + \pi(A - \omega)^2 - \mathbb{1}_{t=3.2}D, \end{aligned}$$

where  $D > 0$  is an exogenous cost for delaying the action,  $\pi \in [0, 1]$  represents the importance of the users’ relative performance and  $\mathbb{1}_{t_i=3.2}$  is an indicator variable that equals 1 if user  $i$  acts late and 0 otherwise. It is apparent that user  $i$  gains if his action is accurate but loses if the other users’ actions are also accurate, on average.

**Lemma 4** *There exists a unique threshold  $D(\beta) \in [0, \frac{1}{\alpha}]$  such that, for given  $\beta$ , all uninformed users act early if  $D \geq D(\beta)$ . Otherwise, they act late. The threshold  $D(\beta)$  is increasing in  $\beta$  and is independent of  $x$ .*

---

<sup>45</sup>I thank two anonymous referees for suggesting this extension.

The users are atomistic, and so the action of a single user does not affect the accuracy of the aggregate action. Hence, when choosing the timing of his action, a representative uninformed user considers only the expected increase in the accuracy of his action due to learning the signal if he acts late (which depends on the precision of the signal  $\beta$  and is independent of  $x$ ) net of the delay cost. Given that all users are identical, a symmetric equilibrium arises, in which all uninformed users act early if the delay cost is sufficiently high and vice versa. The higher the precision of the signal, the more the uninformed users gain by acting late, and so the larger the set of costs  $D$  for which the uninformed users delay.

At date 1, the manager chooses the precision of the information system  $\beta$  and commits to the fraction of informed users  $x$ . I assume that if the manager is indifferent between disseminating the signal only to a subset of users or to everyone, she publicly disseminates the signal.

**Proposition 7**

- (i) *Suppose  $D \geq \frac{1}{\alpha}$ . Then, the manager implements a perfectly revealing information system and disseminates the signal to a fraction  $k$  of users. All users act early so there is no information spillover.*
- (ii) *Suppose  $D \in [\max\{0, \frac{2k-1}{\alpha k^2}\}, \frac{1}{\alpha})$ . Then, the manager implements an information system with precision  $\beta(D) \in (0, \infty)$  and disseminates the signal to a fraction  $k$  of users. All users act early so there is no information spillover. The precision  $\beta(D)$  is increasing in the delay cost  $D$ .*
- (iii) *Suppose  $D \in [0, \max\{0, \frac{2k-1}{\alpha k^2}\})$ . Then, the manager implements a perfectly revealing information system and disseminates the signal to all users.*

If the delay cost is sufficiently high ( $D \geq \frac{1}{\alpha}$ ), then for any feasible  $\beta$ , all users act early. There is no information spillover, and therefore the manager can implement the optimal solution of the baseline model (when the users are required to act simultaneously and there is no spillover threat), i.e.,  $(\infty, k)$ . For lower delay cost ( $D < \frac{1}{\alpha}$ ), if the manager implements a perfectly revealing information system ( $\beta \rightarrow \infty$ ), the uninformed users act late and learn the disseminated signal. Recall that, by Lemma 4, the threshold  $D(\beta)$  is increasing in  $\beta$  and independent of  $x$ . Hence, to avoid the information spillover, the only thing the manager

can do is to sacrifice some amount of precision and choose  $\beta$  just low enough to satisfy  $D = D(\beta)$ , which induces the uninformed users to act early. The manager avoids the spillover, and so her incentives to disseminate the signal to a fraction  $k$  of informed users remain unaffected. The lower the delay cost, the lower the precision level needed to avoid spillover. For sufficiently low delay cost ( $D < \max\{0, \frac{2k-1}{\alpha k^2}\}$ ), it becomes too costly for the manager to keep a subset of users “in the dark” (by further reduction in the precision of the signal), and she prefers to allow the spillover and provide maximally precise information. (This scenario arises only if the preferences of the manager and the users are sufficiently aligned, i.e.,  $\max\{0, \frac{2k-1}{\alpha k^2}\} = \frac{2k-1}{\alpha k^2}$  only if  $k > \frac{1}{2}$ .) Given that the manager is indifferent between disseminating the signal only to a subset of users (and allowing spillover) or to everyone, she publicly disseminates the signal. Hence in equilibrium the users never incur delay costs.

A comparison of the expected utilities of the players reveals that, for any  $k$  and  $D$ , while the manager and the informed users at least weakly prefer UD, the uninformed users at least weakly prefer MD. Hence no Pareto ranking can be made.<sup>46</sup> To consider efficiency, I examine the aggregate users’ welfare under regime  $r = \text{UD, MD}$ . By Proposition 7 the delay cost is not paid in equilibrium and so

$$\begin{aligned} W_r^A &\equiv \int_0^1 E_{t=1}[-(a_i - \omega)^2 + \pi(\hat{A}(\cdot) - \omega)^2 - \mathbb{1}_{t_i=3.2}D]di \\ &= \underbrace{-\int_0^1 E_{t=1}[(a_i - \omega)^2]di}_{\text{base aggregate welfare, } \leq 0} + \pi \underbrace{E_{t=1}[(\hat{A}(\cdot) - \omega)^2]}_{\text{average deviation, } \geq 0}, \quad r = \text{UD, MD}. \end{aligned} \quad (15)$$

The ranking of the base welfare (first term) under UD and MD follows directly from Proposition 2. The relative performance concerns give rise to a gain from average deviation (second term) which affects the overall welfare in the opposite direction. Specifically, the more accurate the users’ actions are on an aggregate level, the larger the base aggregate welfare (first term), but at the same the smaller the deviation of the average users’ action (second term). The following result evaluates the tradeoff between these countervailing forces.

**Corollary 3** *There exists a unique  $\bar{\pi} \in (0, 1)$  such that the aggregate users’ welfare under UD is strictly greater than the aggregate users’ welfare under MD if  $\pi < \bar{\pi}$  and  $k \leq \frac{1}{2}$ .*

---

<sup>46</sup>I thank an anonymous referee for pointing this out.

*Otherwise, the aggregate users' welfare under UD is weakly lower than the aggregate user's welfare under MD.*

The case  $D < \frac{2k-1}{\alpha k^2}$  is feasible only when the preferences of the manager and the users are sufficiently aligned ( $k > \frac{1}{2}$ ). Then, by Proposition 7, under UD, the manager publicly disseminates the signal and so the aggregate users' welfare is the same as the one under MD. If  $D > \max\{0, \frac{2k-1}{\alpha k^2}\}$  and  $k > \frac{1}{2}$ , the base welfare in (15) under MD is larger than the one under UD (by Proposition 2) because under MD all users are informed and take an action that is perfectly aligned with the state of nature while under UD only a fraction  $k$  of the users is informed (Proposition 7). This however also means that the average action is perfectly aligned with the state of nature (because under MD all users are “in the know”), resulting in zero average deviation, which is smaller than the deviation under UD (as then some users remain in “the dark”). The first effect always dominates the second. Hence, when  $k > \frac{1}{2}$ , the users' welfare under UD is weakly lower than the one under MD. Reverse arguments hold for the case of sufficiently misaligned preferences ( $k \leq \frac{1}{2}$ ). Then, the base welfare in (15) is larger under UD (by Proposition 2) but the gain due to average deviation is larger under MD (because then all users are uninformed). For sufficiently low  $\pi$ , the first effect outweighs the second, and so the users' welfare under UD is strictly larger than the one under MD.

In this subsection, I assume that the informed users act early for exogenous reasons. Even with this stylized assumption, the analysis is not trivial. If the informed users could choose the timing of their actions, they might strategically act late (and incur a delay cost) to avoid the spillover and retain their relative performance advantage. Furthermore, if the number of users were finite, the users' incentives to act early would depend on the fraction of informed users, in addition to the precision. Then, the manager could avoid spillover by strategically using either the fraction  $x$ , the precision  $\beta$ , or both. It will be interesting to investigate these and other related questions in future work.

### **5.3.2 Actions Observed With Noise**

Often the users of information do not perfectly observe the actions of their peers. To analyze such settings, I return to the baseline assumption that the users care only about their own accuracy and do not incur delay costs. I assume that the users' actions are observed with

an additive random noise  $\xi_i \sim \mathcal{N}\left(0, \frac{1}{\gamma}\right)$ ,  $\gamma \geq 0$ , where  $\xi_i$  is not correlated with any other noise term. For tractability reasons, I assume that the total number of users is  $N > 0$  and sufficiently large, where users  $i = 1, 2, \dots, z$  are informed and their fraction is  $x = \frac{z}{N}$ .

As before, the informed users act early for exogenous reasons. Their actions are given by  $E_{t=3.1}[\omega|s, \beta] = \mu(s|\beta)$  and are observed by their peers as  $\mu(s|\beta) + \xi_i$ ,  $i = 1, 2, \dots, z$ . The uninformed users prefer to act late, hoping to observe the actions of their informed peers. Let  $Z$  be the vector of the actions of the informed users as observed by their peers. Upon observing  $Z$ , the uninformed users update their beliefs about  $\omega$  and, at date 3.2, take actions  $E_{t=3.2}[\omega|Z]$ . If the actions of the users are very noisy, i.e., if  $\gamma \rightarrow 0$ , then the uninformed users will not learn anything from observing the actions of the informed users, and all results under UD from Section 3 hold. For strictly positive finite  $\gamma$ , the uninformed users update their beliefs but cannot perfectly reveal the signal.

**Proposition 8** *There exists a threshold  $\bar{\gamma} > 0$ , such that, if  $\gamma \in (0, \bar{\gamma})$ , the manager implements an information system that perfectly reveals the state of nature to a fraction  $x(\gamma) \in (0, k)$  of users. The fraction  $x(\gamma)$  is decreasing in  $\gamma$ .*

The manager's objective is to regulate the information spillover and minimize the differential between the actual aggregate action and her preferred aggregate action. To limit the spillover, the manager can either provide opaque information or reduce the fraction of informed users (because the fewer the informed users, the less the uninformed users can learn from observing  $Z$ ). Providing opaque information introduces additional noise in the aggregate action, unrelated to the realized state  $\omega$ . Therefore the manager prefers to decrease the fraction  $x$  below  $k$  and provide as precise as possible information.

As  $\gamma$  increases, the uninformed users learn more from the actions of the users “in the know,” and it becomes very costly for the manager to control the spillover by decreasing  $x$ . In the limit case, as  $\gamma \rightarrow \infty$ , the uninformed users perfectly reveal the signal, and so the optimal solution will be identical to the one under MD regime. This implies that noise can be welfare increasing. To illustrate suppose that  $k \leq \frac{1}{2}$ . Then, if  $\gamma \rightarrow \infty$ , all users know the signal (either observe the signal or learn it from the actions of their peers) and the manager prefers not to implement an information system. As  $\gamma$  decreases, the uninformed users cannot reveal the signal from the actions of their informed peers, and this provides

incentives to the manager to provide socially beneficial information. The comparison of UD and MD remains the same as in Proposition 2.

## 6 Concluding Remarks

In this article, I demonstrate the tension between the size of the set of users to whom information is provided and the quality of that information chosen by a manager. The results call into doubt the commonly held belief that fairness and more overall information go hand in hand. I find that, when managers can selectively disseminate information, they may gather more precise information. Hence an unintended consequence of rules requiring public dissemination (such as Regulation FD) could be a reduction of information in the public domain. Ironically, this effect is especially strong when the incentive conflict between managers and information users is severe. This article thus calls into doubt the conventional wisdom that regulating information dissemination is especially needed when the incentive conflicts are severe.

## Appendix

**Proof of Proposition 1:** Note that  $V(x) \equiv V(x, \hat{\beta}(x))$  is single-peaked at  $x = k$ . To see this start with  $x = k$ . By Lemma 1, the manager implements a perfectly revealing information system so that  $s = \omega$ ,  $\hat{A}(\cdot) = k\omega + (1 - k)\mu_0$  and

$$\begin{aligned} V(k) &= -E_{t=1}[(k\omega + (1 - k)\mu_0 - k\omega - (1 - k)\bar{\omega})^2] \\ &= -(1 - k)^2(\mu_0 - \bar{\omega})^2. \end{aligned}$$

Now consider  $x = k + \eta_1$ , where  $\eta_1 \in (0, k)$ . By Lemma 1, the manager implements a perfectly revealing information system so that  $s = \omega$ ,  $\hat{A}(\cdot) = (k + \eta_1)\omega + (1 - k - \eta_1)\mu_0$  and

$$\begin{aligned} V(k + \eta_1) &= -E_{t=1}[((k + \eta_1)\omega + (1 - k - \eta_1)\mu_0 - k\omega - (1 - k)\bar{\omega})^2] \\ &= -(1 - k)^2(\mu_0 - \bar{\omega})^2 - \eta_1^2 E_{t=1}[\varepsilon^2] \\ &< -(1 - k)^2(\mu_0 - \bar{\omega})^2 \\ &= V(k). \end{aligned}$$

Similarly,  $V(k - \eta_1) < V(k)$ . Now consider  $x = k + \eta_2$ , where  $\eta_2 \geq k$ . By Lemma 1, the manager does not implement an information system so  $\hat{A}(\cdot) = \mu_0$  and

$$\begin{aligned} V(k + \eta_2) &= -E_{t=1}[(\mu_0 - k\omega - (1 - k)\bar{\omega})^2] \\ &= -(1 - k)^2(\mu_0 - \bar{\omega})^2 - k^2 E_{t=1}[\varepsilon^2] \\ &< -(1 - k)^2(\mu_0 - \bar{\omega})^2 - \eta_1^2 E_{t=1}[\varepsilon^2] \\ &= V(k + \eta_1) \\ &< V(k). \end{aligned}$$

It follows that the manager will implement a perfectly revealing information system and reveal the signal to a fraction  $k$  of users. ■

**Proof of Lemma 2:** Simplifying,

$$\begin{aligned} V(x, \beta) &= -E_{t=1}[(x\mu(s|\beta) + (1 - x)\mu_0 - k\omega - (1 - k)\bar{\omega})^2] \\ &= -E_{t=1} \left[ \left( x(\varepsilon + \delta) \frac{\beta}{\alpha + \beta} - k\varepsilon + (1 - k)(\mu_0 - \bar{\omega}) \right)^2 \right] \\ &= -E_{t=1} \left[ \left( x(\varepsilon + \delta) \frac{\beta}{\alpha + \beta} - k\varepsilon \right)^2 \right] - (1 - k)^2(\mu_0 - \bar{\omega})^2 \end{aligned}$$

Then, differentiating the manager's objective in (8) with respect to  $\beta$ ,

$$\begin{aligned}
\frac{\partial V(x, \beta)}{\partial \beta} - \frac{\partial C(c, \beta)}{\partial \beta} &= -\frac{\partial}{\partial \beta} E_{t=1} \left[ \left( x(\varepsilon + \delta) \frac{\beta}{\alpha + \beta} - k\varepsilon \right)^2 \right] - \frac{\partial C(c, \beta)}{\partial \beta} \\
&= \frac{\partial}{\partial \beta} \left( \frac{x(2k - x)\beta}{\alpha(\alpha + \beta)} - \frac{k^2}{\alpha} \right) - \frac{\partial C(c, \beta)}{\partial \beta} \\
&= \frac{(2k - x)x}{(\alpha + \beta)^2} - \frac{\partial C(c, \beta)}{\partial \beta}.
\end{aligned} \tag{16}$$

If  $k \leq \frac{x}{2}$ , then (16) is negative for any precision in the domain, and the manager does not implement an information system (equivalently, sets  $\beta = 0$ ). If  $k > \frac{x}{2}$ , then there exist  $\beta^c(x)$  that satisfies FOC:

$$\left. \frac{(2k - x)x}{(\alpha + \beta)^2} - \frac{\partial C(c, \beta)}{\partial \beta} \right|_{\beta = \beta^c(x)} = 0 \tag{17}$$

Applying the Implicit Function Theorem,

$$\begin{aligned}
\frac{\partial \beta^c(x)}{\partial c} &= -\frac{\frac{\partial^2 C(c, \beta)}{\partial \beta \partial c}}{\frac{2x(2k - x)}{(\alpha + \beta)^3} + \frac{\partial^2 C(c, \beta)}{\partial \beta^2}} < 0 \\
\frac{\partial \beta^c(x)}{\partial \alpha} &= -\frac{\frac{2(2k - x)x}{(\alpha + \beta)^3}}{\frac{2x(2k - x)}{(\alpha + \beta)^3} + \frac{\partial^2 C(c, \beta)}{\partial \beta^2}} < 0 \\
\frac{\partial \beta^c(x)}{\partial k} &= \frac{\frac{2x}{(\alpha + \beta)^2}}{\frac{2x(2k - x)}{(\alpha + \beta)^3} + \frac{\partial^2 C(c, \beta)}{\partial \beta^2}} > 0,
\end{aligned}$$

because  $k > \frac{x}{2}$ . Lastly, note that  $\beta^c(x)$  is single peaked at  $x = k$  because, by the Implicit Function Theorem,

$$\frac{\partial \beta^c(x)}{\partial x} = -\frac{\frac{2(x - k)}{(\alpha + \beta)^2}}{\frac{2x(2k - x)}{(\alpha + \beta)^3} + \frac{\partial^2 C(c, \beta)}{\partial \beta^2}} \begin{cases} > 0 & \text{if } x < k, \\ = 0 & \text{if } x = k, \\ < 0 & \text{if } x > k. \end{cases}$$

(The last inequality holds because the manager chooses  $\beta = \beta^c(x)$  if and only if  $k > \frac{x}{2}$ . Otherwise, she does not implement an information system, i.e.,  $\beta = 0$ ). ■

**Proof of Proposition 2:** By revealed preference, the manager always prefers UD over MD. Given their quadratic loss payoff, the users prefer more precise information. By Lemma 1 and Proposition 1, if  $k \leq \frac{1}{2}$ , then the manager implements an information system only under UD.

The informed users are better off under UD (because it is the only regime under which the manager implements an information system). The uninformed users are indifferent (because they do not observe information under either regime). As a result, UD Pareto dominates MD. However, if  $k > \frac{1}{2}$ , the manager implements a perfectly revealing information system under both regimes. The informed users are indifferent (because they perfectly observe  $\omega$  under both regimes), while the users who are uninformed under UD are better off under MD (because it allows them to observe information). ■

**Proof of Proposition 3:** Differentiating the manager's objective in (9),

$$\begin{aligned}
\frac{\partial V(x, \beta)}{\partial \beta} - \frac{\partial C(c, \beta)}{\partial \beta} &= -\frac{\partial}{\partial \beta} E_{t=1} \left[ \left( x(\varepsilon + \delta) \frac{\beta}{\alpha + \beta} - k\varepsilon \right)^2 \right] - \frac{\partial C(c, \beta)}{\partial \beta} \\
&= \frac{\partial}{\partial \beta} \left( \frac{x(2k - x)\beta}{\alpha(\alpha + \beta)} - \frac{k^2}{\alpha} \right) - \frac{\partial C(c, \beta)}{\partial \beta} \\
&= \frac{x(2k - x)}{(\alpha + \beta)^2} - \frac{\partial C(c, \beta)}{\partial \beta};
\end{aligned} \tag{18}$$

$$\begin{aligned}
\frac{\partial V(x, \beta)}{\partial x} - \frac{\partial C(c, \beta)}{\partial x} &= -\frac{\partial}{\partial x} E_{t=1} \left[ \left( x(\varepsilon + \delta) \frac{\beta}{\alpha + \beta} - k\varepsilon \right)^2 \right] \\
&= \frac{\partial}{\partial x} \left( \frac{x(2k - x)\beta}{\alpha(\alpha + \beta)} - \frac{k^2}{\alpha} \right) \\
&= \frac{2\beta(k - x)}{\alpha(\alpha + \beta)}.
\end{aligned} \tag{19}$$

The critical points satisfying (18) and (19) simultaneously are

$$(x, \beta) \in \{(k, \tilde{\beta}), (0, 0), (2k, 0)\},$$

where  $\tilde{\beta}$  satisfies  $\frac{k^2}{(\alpha + \beta)^2} - \frac{\partial C(c, \beta)}{\partial \beta} \Big|_{\beta = \tilde{\beta}} = 0$ . To verify SOC, I examine the Hessian:

$$H = \begin{bmatrix} -\frac{2\beta}{\alpha(\alpha + \beta)} & \frac{2(k - x)}{(\alpha + \beta)^2} \\ \frac{2(k - x)}{(\alpha + \beta)^2} & -\frac{2(2k - x)x}{(\alpha + \beta)^3} - \frac{\partial^2 C(c, \beta)}{\partial \beta^2} \end{bmatrix}.$$

At  $(\tilde{x}, 0)$ , where  $\tilde{x} \in \{0, 2k\}$ , the Hessian is

$$H = \begin{bmatrix} 0 & \frac{2(k - \tilde{x})}{\alpha^2} \\ \frac{2(k - \tilde{x})}{\alpha^2} & -\frac{2(2k - \tilde{x})\tilde{x}}{(\alpha + \beta)^3} \end{bmatrix},$$

so  $|H_1| = 0$  and  $|H_2| < 0$ . However, at  $(k, \tilde{\beta})$ , the Hessian is

$$H = \begin{bmatrix} -\frac{2\tilde{\beta}}{\alpha(\alpha+\tilde{\beta})} & 0 \\ 0 & -\frac{2k^2}{(\alpha+\tilde{\beta})^3} - \frac{\partial^2 C(c, \tilde{\beta})}{\partial \beta^2} \end{bmatrix},$$

so  $|H_1| < 0$  and  $|H_2| > 0$ . Therefore  $(x_{UD}^c, \beta_{UD}^c) = (k, \tilde{\beta})$ . Next note that  $\tilde{\beta} = \beta^c(x = k)$ . Hence, if the manager can choose the fraction of informed users, she will set  $x_{UD}^c = k$  and  $\beta_{UD}^c \equiv \beta^c(x = k)$ . ■

**Proof of Proposition 4:** The result for the manager follows by revealed preference. To evaluate the users' preferences, note that a user who observes a signal incurs in expectation the posterior variance and his payoff is increasing in the signal precision  $\beta$ :

$$\begin{aligned} \frac{\partial}{\partial \beta} E_{t=1}[u(\hat{a}(s, \beta), \omega)|s, \beta] &= -\frac{\partial}{\partial \beta} E_{t=1} \left[ \left( \mu_0 + \frac{\beta(\varepsilon + \delta)}{\alpha + \beta} - \omega \right)^2 \right] \\ &= -\frac{\partial}{\partial \beta} E_{t=1} \left[ \left( \frac{\beta\delta}{\alpha + \beta} - \frac{\alpha\varepsilon}{\alpha + \beta} \right)^2 \right] \\ &= -\frac{\partial}{\partial \beta} \left( \frac{\beta^2}{(\alpha + \beta)^2} E_{t=1}[\delta^2] - \frac{\alpha^2}{(\alpha + \beta)^2} E_{t=1}[\varepsilon^2] \right) \\ &= -\frac{\partial}{\partial \beta} \left( \frac{1}{\alpha + \beta} \right) \\ &= \frac{1}{(\alpha + \beta)^2} > 0. \end{aligned} \tag{20}$$

It immediately follows that

$$E_{t=1}[u(\hat{a}(s, \beta), \omega)|s, \beta] \geq E_{t=1}[u(\hat{a}(\mu_0), \omega)], \forall \beta \geq 0. \tag{21}$$

*Case  $k \leq \frac{1}{2}$ :* By Lemma 2 and Proposition 3, the manager implements an information system only under UD. The uninformed users are indifferent, because they do not observe a signal under either of the regimes. By (21), the informed users are better off under UD.

*Case  $k > \frac{1}{2}$ :* By Lemma 2 and Proposition 3, the manager implements an information system under both regimes. By (21), the uninformed users are strictly better off under MD (because they observe a signal, while under UD they do not). By Lemma 2 and (20), the informed users strictly prefer UD. ■

**Proof of Corollary 1:** Note that the information gain is increasing in the precision, because

$$\frac{\partial}{\partial \beta} (E_{t=1}[u(\hat{a}(s, \beta), \omega)|s, \beta] - E_{t=1}[u(\hat{a}(\mu_0), \omega)]) = \frac{\partial}{\partial \beta} E_{t=1}[u(\hat{a}(s, \beta), \omega)|s, \beta] > 0$$

by (20). Then, by Lemma 2,  $\phi^c(x) \equiv E_{t=1}[u(\hat{a}(s, \beta), \omega)|s, \beta] - E_{t=1}[u(\hat{a}(\mu_0), \omega)]$  is decreasing in  $\Delta x \equiv |x - k|$ . ■

**Proof of Proposition 5:** Consider the differential of the welfare terms under the two regimes:

$$\Delta W \equiv W_{UD}^c - W_{MD}^c = k\phi_{UD}^c - \phi_{MD}^c.$$

The sign of  $\Delta W$  is nontrivial (because  $k < 1$  by assumption but  $\phi_{UD}^c > \phi_{MD}^c$  by Corollary 1). Note that  $\lim_{c \rightarrow 0} \beta_{MD}^c = \lim_{c \rightarrow 0} \beta_{UD}^c \rightarrow \infty$ . Note that for  $j = UD, MD$ ,

$$\begin{aligned} \phi_j^c &= E_{t=1}[u(\hat{a}(s, \beta_j^c), \omega)|s, \beta_j^c] - E_{t=1}[u(\hat{a}(\mu_0), \omega)] \\ &= -\frac{1}{\alpha + \beta_j^c} + \frac{1}{\alpha} \\ &= \frac{\beta_j^c}{\alpha(\alpha + \beta_j^c)}. \end{aligned}$$

It follows that  $\lim_{c \rightarrow 0} \phi_{MD}^c = \lim_{c \rightarrow 0} \phi_{UD}^c = \frac{1}{\alpha}$  and hence

$$\lim_{c \rightarrow 0} \Delta W \propto k - 1 < 0.$$

Further,  $\lim_{c \rightarrow \infty} \beta_{MD}^c = \lim_{c \rightarrow \infty} \beta_{UD}^c = 0$ , so  $\lim_{c \rightarrow \infty} \phi_{MD}^c = \lim_{c \rightarrow \infty} \phi_{UD}^c = 0$  and hence

$$\lim_{c \rightarrow \infty} \Delta W = 0.$$

However,  $\phi_{MD}^c$  reaches zero weakly faster than  $k\phi_{UD}^c$  if  $k \leq \frac{1}{2}(\sqrt{5} - 1)$ . To see why consider:

$$\Phi \equiv \frac{k\phi_{UD}^c}{\phi_{MD}^c} = \frac{\frac{k\beta_{UD}^c}{\alpha(\alpha + \beta_{UD}^c)}}{\frac{\beta_{MD}^c}{\alpha(\alpha + \beta_{MD}^c)}}.$$

Note that, if  $C(c, \beta) = \frac{c\beta^2}{2}$ , then  $\beta_{UD}^c$  and  $\beta_{MD}^c$  satisfy

$$\begin{aligned} \beta_{UD}^c &= \frac{k^2}{c(\alpha + \beta_{UD}^c)^2}, \\ \beta_{MD}^c &= \frac{2k - 1}{c(\alpha + \beta_{MD}^c)^2}. \end{aligned}$$

and hence

$$\Phi = \frac{\frac{k^3}{\alpha c(\alpha + \beta_{UD}^c)^3}}{\frac{2k - 1}{\alpha c(\alpha + \beta_{MD}^c)^3}} = \frac{k^3(\alpha + \beta_{MD}^c)^3}{(2k - 1)(\alpha + \beta_{UD}^c)^3}.$$

Now note that

$$\lim_{c \rightarrow \infty} \Phi = \frac{k^3}{2k-1} \geq 1,$$

if  $k \leq \frac{1}{2}(\sqrt{5}-1)$ . In other words,  $\phi_{MD}^c$  reaches zero weakly faster than  $k\phi_{UD}^c$  if  $k \leq \tilde{k} \equiv \frac{1}{2}(\sqrt{5}-1) \in (\frac{1}{2}, 1)$ . It follows that for  $c$  sufficiently high and  $k \leq \tilde{k}$ ,  $\Delta W \geq 0$ . ■

**Proof of Corollary 2:** Let

$$\Pi^c(x) \equiv \lambda[V(x, \beta^c(x)) - C(c, \beta^c(x))] + W^c(x)$$

denote the objective function of the planner with costly implementation, where

$$\beta^c(x) = \arg \max_{\beta} V(x, \beta) - C(c, \beta).$$

The optimal  $\hat{x}_{RD}^c$  satisfies  $\frac{d\Pi^c(x)}{dx} \Big|_{x=\hat{x}_{RD}^c} = 0$ . Differentiating with respect to  $x$ ,

$$\begin{aligned} \frac{d\Pi^c(x)}{dx} &= \lambda \underbrace{\left( \frac{\partial V(x, \beta)}{\partial \beta} - \frac{\partial C(c, \beta)}{\partial \beta} \right)}_{=0 \text{ by FOC}} \Big|_{\beta=\beta^c(x)} \frac{d\beta^c(x)}{dx} \\ &\quad + \lambda \frac{\partial V(x, \beta^c(\cdot))}{\partial x} + \phi^c(x) + x \frac{d\phi^c(x)}{dx} \\ &= \frac{1}{\alpha + \beta^c(x)} \left[ \frac{\beta^c(x)(1 + 2\lambda(k-x))}{\alpha} + \frac{x}{(\alpha + \beta^c(x))} \frac{\partial \beta^c(x)}{\partial x} \right]. \end{aligned}$$

Let  $g(x) \equiv \frac{\beta^c(x)(1+2\lambda(k-x))}{\alpha} + \frac{x}{(\alpha+\beta^c(x))} \frac{\partial \beta^c(x)}{\partial x}$ . Given that  $\frac{1}{\alpha+\beta^c(x)} > 0$ ,  $\hat{x}_{RD}^c$  satisfies  $g(x)|_{x=\hat{x}_{RD}^c} = 0$ . Next observe that, if  $C(c, \beta) = \frac{c\beta^2}{2}$ , then  $\beta^c(x)$  satisfies

$$\frac{(2k-x)x}{(\alpha + \beta^c(x))^2} - c\beta^c(x) = 0 \tag{22}$$

and applying the Implicit Function Theorem,

$$\frac{d\beta^c(x)}{dx} = \frac{\frac{2(k-x)}{(\alpha+\beta^c(x))^2}}{\frac{2x(2k-x)}{(\alpha+\beta^c(x))^3} + c} = \frac{(k-x)(\alpha + \beta^c(x))}{(2k-x)x}.$$

Substituting,

$$g(x) = \frac{\beta^c(x)(1 + 2\lambda(k-x))}{\alpha} + \frac{k-x}{2k-x}.$$

Note that  $\hat{x}_{RD}^c \geq k$  because

$$g(x=k) = \frac{\beta^c(x=k)}{\alpha} \geq 0.$$

Observe that  $\hat{x}_{RD}^c < 2k$  (to satisfy the implementation constraint), because by Lemma 2, if  $x \geq 2k$ , then  $\beta^c(x) = 0$ . If  $k \leq \frac{1}{2}$ , then it follows directly that  $\hat{x}_{RD}^c < 1$  (because  $2k < 1$ ). However, if  $k > \frac{1}{2}$ , then  $\hat{x}_{RD}^c < 1$  only if  $c$  sufficiently high. To see why first consider  $c = 0$ .

$$\lim_{c \rightarrow 0} g(x = 1) = \frac{\lim_{c \rightarrow 0} \beta^c(x = 1)(1 + 2\lambda(k - 1))}{\alpha} + \frac{k - 1}{2k - 1} > 0$$

because when  $k > \frac{1}{2}$  then  $1 + 2\lambda(k - 1) > 1 + 2\lambda(\frac{1}{2} - 1) = 1 - \lambda \geq 0$  (because  $\lambda \leq 1$  by assumption) and  $\lim_{c \rightarrow 0} \beta^c(x = 1) \rightarrow \infty$ . In other words, if  $c = 0$  and  $k > \frac{1}{2}$ , the planner sets  $x_{RD}^c = 1$ . However,

$$\begin{aligned} g(x = 1) &= \frac{\beta^c(x = 1)(1 + 2\lambda(k - 1))}{\alpha} + \frac{k - 1}{2k - 1} \\ &\propto \beta^c(x = 1)(2k - 1)(1 + 2\lambda(k - 1)) + (k - 1)\alpha < 0, \end{aligned}$$

if  $\beta^c(x = 1)$  is sufficiently low, which by Lemma 2 occurs when  $c$  is sufficiently high (recall that  $\lim_{c \rightarrow \infty} \beta^c(x = 1) = 0$ ). It follows that  $x_{RD}^c < 1$  if  $c$  is sufficiently high for any  $k$ .

To show the comparative statics of  $x_{RD}^c$  with respect to  $c$ , I apply the Implicit Function Theorem:

$$\begin{aligned} \frac{\partial x_{RD}^c}{\partial c} &= - \frac{\frac{(1 + 2\lambda(k - x_{RD}^c))}{\alpha} \frac{\partial \beta^c(x = x_{RD}^c)}{\partial c}}{\frac{-\alpha k + (x_{RD}^c - 2k)^2 (-2\lambda \beta^c(x = x_{RD}^c) + (1 + 2\lambda(k - x_{RD}^c)) \frac{\partial \beta^c(x = x_{RD}^c)}{\partial x})}{\alpha (x_{RD}^c - 2k)^2}} \\ &\propto \frac{1 + 2\lambda(k - x_{RD}^c)}{-\alpha k + (x_{RD}^c - 2k)^2 (-2\lambda \beta^c(x = x_{RD}^c) + (1 + 2\lambda(k - x_{RD}^c)) \frac{\partial \beta^c(x = x_{RD}^c)}{\partial x})} < 0 \end{aligned}$$

because

- (i)  $\alpha \geq 0$  by assumption and  $(x_{RD}^c - 2k)^2 > 0$ ;
- (ii)  $\frac{\partial \beta^c(x = x_{RD}^c)}{\partial c} < 0$  by Lemma 2;
- (ii)  $\frac{\partial \beta^c(x = x_{RD}^c)}{\partial x} = \frac{(k - x_{RD}^c)(\alpha + \beta^c(x = x_{RD}^c))}{(2k - x_{RD}^c)x_{RD}^c} \propto \frac{(k - x_{RD}^c)}{(2k - x_{RD}^c)} < 0$ , because  $x_{RD}^c \in (k, 2k)$ ;
- (iii)  $1 + 2\lambda(k - x_{RD}^c) > 0$  if  $\lambda \in [0, 1]$ . To see this, recall from the preceding discussion in the proof that  $x_{RD}^c < \text{Min}\{2k, 1\}$ . Therefore, if  $k \leq \frac{1}{2}$ , then  $\text{Min}\{2k, 1\} = 2k$  and  $1 + 2\lambda(k - x_{RD}^c) > 1 + 2\lambda(k - 2k) = 1 - 2\lambda k > 1 - 2k > 0$  (because  $\lambda \leq 1$  by assumption). If  $k > \frac{1}{2}$ , then  $\text{Min}\{2k, 1\} = 1$  and  $1 + 2\lambda(k - x_{RD}^c) > 1 + 2\lambda(k - 1) > 1 + 2\lambda(\frac{1}{2} - 1) = 1 - \lambda \geq 0$  (because  $\lambda \leq 1$  by assumption).  $\blacksquare$

**Proof of Lemma 3:** *Part (i):* At date 3, the informed users take actions that equal their posterior expectation  $\mu(s|\beta)$ . The maintained assumption is that the users (i) are not ordered on  $[0, 1]$  and (ii) realize they are included “in the know” only after receiving the signal. Hence if a user ends up being uninformed, he only infers that the signal was such that the manager did not choose  $x = 1$ . Hence, at date 3, the uninformed users take actions that equal their posterior belief  $\hat{\mu} \equiv E[\omega|\hat{x} \neq 1]$ , where  $\hat{x}$  is the uninformed users’ conjecture of the manager’s dissemination strategy. In equilibrium, this conjecture coincides with the manager’s choice of dissemination strategy at date 2:

$$x_{t=2}(s|\beta) = \arg \max_{x \in [0,1]} E_{t=2}[v(\hat{A}(x, s, \beta), \omega)|s, \beta]$$

and the posterior beliefs of the uninformed users (upon realizing they are not included “in the know”) satisfy

$$\hat{\mu} = E[\omega|x_{t=2}(s|\beta) \neq 1].$$

Suppose  $\hat{\mu} < \bar{\omega}$  (I show below that this indeed is the case). To avoid clutter let  $\mu_s \equiv \mu(s|\beta)$ . Simplifying,

$$\begin{aligned} E_{t=2}[v(\hat{A}(x, s, \beta), \omega)|s, \beta] &= - (E_{t=2}[x\mu_s + (1-x)\hat{\mu} - k\omega - (1-k)\bar{\omega}|s, \beta])^2 \\ &\quad - \text{Var}(x\mu_s + (1-x)\hat{\mu} - k\omega - (1-k)\bar{\omega})|s, \beta) \\ &= - ((x-k)\mu_s + (1-x)\hat{\mu} - (1-k)\bar{\omega})^2 - k^2 \text{Var}(\omega|s, \beta), \end{aligned} \quad (23)$$

because  $\text{Var}(\mu_s|s, \beta) = 0$ . The manager’s optimization problem at date 2 is:

$$\begin{aligned} \max_x \hat{V}(x|s, \beta) &\equiv - ((x-k)\mu_s + (1-x)\hat{\mu} - (1-k)\bar{\omega})^2 - k^2 \text{Var}(\omega|s, \beta) \\ \text{s.t.} \quad &0 \leq x \leq 1. \end{aligned}$$

The Lagrangian is:

$$\mathcal{L}^R = - ((x-k)\mu_s + (1-x)\hat{\mu} - (1-k)\bar{\omega})^2 - k^2 \text{Var}(\omega|s, \beta) + \lambda_1^R x + \lambda_2^R (1-x).$$

The Karush-Kuhn-Tucker stationarity condition is:

$$\frac{\partial \mathcal{L}^R}{\partial x} = -2(\hat{\mu} - \mu_s)(k(\mu_s - \bar{\omega}) + \bar{\omega} - \hat{\mu}(1-x) - x\mu_s) + \lambda_1^R - \lambda_2^R = 0.$$

The Karush-Kuhn-Tucker feasibility conditions are:

$$x \geq 0 \quad \text{and} \quad 1 - x \geq 0.$$

The Karush-Kuhn-Tucker complementarity slackness conditions are:

$$\lambda_1^R x = 0 \quad \text{and} \quad \lambda_2^R (1 - x) = 0.$$

The case  $\lambda_1^R = 0$  and  $\lambda_2^R = 0$  can be immediately ruled out because it implies  $x = 0$  and  $x = 1$ , which cannot hold simultaneously. There are three remaining cases to be considered:

**Case 1:**  $\lambda_1^R = 0$ ,  $x > 0$ ,  $\lambda_2^R > 0$ ,  $1 - x = 0$

This case implies  $x = 1$  and  $\lambda_2^R = 2(1 - k)(\mu_s - \hat{\mu})(\bar{\omega} - \mu_s)$ . This is feasible (i.e.,  $\lambda_2^R > 0$ ) if  $\mu_s \in (\hat{\mu}, \bar{\omega})$ .

**Case 2:**  $\lambda_1^R > 0$ ,  $x = 0$ ,  $\lambda_2^R = 0$ ,  $1 - x > 0$

This case implies  $x = 0$  and  $\lambda_1^R = 2(\hat{\mu} - \mu_s)(k(\mu_s - \bar{\omega}) + \bar{\omega} - \hat{\mu})$ . This is feasible (i.e.,  $\lambda_1^R > 0$ ) if  $\mu_s \in (m, \hat{\mu})$ , where  $m \equiv \frac{\hat{\mu} - (1-k)\bar{\omega}}{k} < \hat{\mu} < \bar{\omega}$ .

**Case 3:**  $\lambda_1^R = 0$ ,  $x > 0$ ,  $\lambda_2^R = 0$ ,  $1 - x > 0$

This case implies  $x = k + (1 - k)\frac{(\hat{\mu} - \bar{\omega})}{(\hat{\mu} - \mu_s)}$ , and it is feasible (i.e.,  $0 \leq x \leq 1$ ) if  $\mu_s \notin (m, \bar{\omega})$ .

Summarizing, the solution to the optimization program is:

$$x_{t=2}(s|\beta) = \begin{cases} k + (1 - k)\frac{(\hat{\mu} - \bar{\omega})}{(\hat{\mu} - \mu_s)} \in (0, k) & \text{if the signal is such that } \mu_s < m, \\ 0 & \text{if the signal is such that } \mu_s \in [m, \hat{\mu}), \\ 1 & \text{if the signal is such that } \mu_s \in (\hat{\mu}, \bar{\omega}], \\ k + (1 - k)\frac{(\hat{\mu} - \bar{\omega})}{(\hat{\mu} - \mu_s)} \in (k, 1) & \text{if the signal is such that } \mu_s > \bar{\omega}. \end{cases}$$

If the signal is such that  $\mu_s = \hat{\mu}$  any  $x$  is a solution.

As a next step, note that  $\mu_s$  is a normally distributed random variable. Let  $f(\mu_s)$  denote the p.d.f. of  $\mu_s$ . In equilibrium, the beliefs of the uninformed users satisfy:

$$\hat{\mu} = E[\omega | x_{t=2}(s|\beta) \neq 1] = \frac{\int_{-\infty}^{\hat{\mu}} \mu_s f(\mu_s) d\mu_s + \int_{\bar{\omega}}^{\infty} \mu_s f(\mu_s) d\mu_s}{\int_{-\infty}^{\hat{\mu}} f(\mu_s) d\mu_s + \int_{\bar{\omega}}^{\infty} f(\mu_s) d\mu_s}.$$

Rearranging the above expression yields:

$$\int_{-\infty}^{\hat{\mu}} (\hat{\mu} - \mu_s) f(\mu_s) d\mu_s = \int_{\bar{\omega}}^{\infty} (\mu_s - \hat{\mu}) f(\mu_s) d\mu_s. \quad (24)$$

Observe that the LHS is increasing in  $\hat{\mu}$ , while the RHS is decreasing in  $\hat{\mu}$ . Next, note that, as  $\hat{\mu} \rightarrow -\infty$ , the LHS becomes zero and hence lower than the RHS, which is positive. Furthermore, as  $\hat{\mu} \rightarrow \mu_0$ , the LHS exceeds the RHS as  $\mu_0 < \bar{\omega}$ . Lastly, note that (24) never holds for  $\hat{\mu} < \bar{\omega}$ . Summarizing, there exists a unique  $\hat{\mu} \in (-\infty, \mu_0)$  that satisfies (24).

*Part (ii):* The proof follows a similar logic to the one in the proof of part (i). Suppose  $\hat{\mu} > \bar{\omega}$ . (I show below that this indeed is the case.) Following similar steps, the solution to the optimization program is:

$$x_{t=2}(s|\beta) = \begin{cases} k + (1-k) \frac{(\hat{\mu}-\bar{\omega})}{(\hat{\mu}-\mu_s)} \in (k, 1) & \text{if the signal is such that } \mu_s < \bar{\omega}, \\ 1 & \text{if the signal is such that } \mu_s \in [\bar{\omega}, \hat{\mu}), \\ 0 & \text{if the signal is such that } \mu_s \in (\hat{\mu}, m], \\ k + (1-k) \frac{(\hat{\mu}-\bar{\omega})}{(\hat{\mu}-\mu_s)} \in (0, k) & \text{if the signal is such that } \mu_s > m, \end{cases}$$

where  $m \equiv \frac{\hat{\mu} - (1-k)\bar{\omega}}{k} > \hat{\mu}$ . If the signal is such that  $\mu_s = \hat{\mu}$  any  $x$  is a solution (in a continuous distribution the probability of this event is zero). Lastly, in equilibrium, the beliefs of the uninformed users satisfy:

$$\hat{\mu} = E[\omega | x_{t=2}(s|\beta) \neq 1] = \frac{\int_{-\infty}^{\bar{\omega}} \mu_s f(\mu_s) d\mu_s + \int_{\hat{\mu}}^{\infty} \mu_s f(\mu_s) d\mu_s}{\int_{-\infty}^{\bar{\omega}} f(\mu_s) d\mu_s + \int_{\hat{\mu}}^{\infty} f(\mu_s) d\mu_s}.$$

Rearranging the above expression yields:

$$\int_{-\infty}^{\bar{\omega}} (\hat{\mu} - \mu_s) f(\mu_s) d\mu_s = \int_{\hat{\mu}}^{\infty} (\mu_s - \hat{\mu}) f(\mu_s) d\mu_s. \quad (25)$$

Observe that the LHS is increasing in  $\hat{\mu}$ , while the RHS is decreasing in  $\hat{\mu}$ . Next, note that, as  $\hat{\mu} \rightarrow \infty$ , the LHS is positive and hence higher than the RHS, which becomes zero. Furthermore, as  $\hat{\mu} \rightarrow \mu_0$ , the LHS is lower than the RHS as  $\mu_0 > \bar{\omega}$ . Lastly, note that (24) never holds for  $\hat{\mu} > \bar{\omega}$ . Summarizing, there exists a unique  $\hat{\mu} \in (\mu_0, \infty)$  that satisfies (24). ■

**Proof of Proposition 6:** Consider  $\mu_0 < \bar{\omega}$ . Again, to avoid clutter, let  $\mu_s \equiv \mu(s|\beta)$  and  $m \equiv \frac{\hat{\mu} - (1-k)\bar{\omega}}{k}$ . Substituting for  $x_{t=2}(s|\beta)$  (as derived in the proof of Lemma 3), and noting that  $Var(\omega|s, \beta) = -(E[\omega|s, \beta])^2 + E[\omega^2|s, \beta] = -\mu_s^2 + E[\omega^2|s, \beta]$ , the manager's expected

payoff at date 2 at the revised fraction,  $\hat{v}(\mu_s) \equiv E_{t=2}[v(\hat{A}(x_{t=2}(s|\beta), s, \beta), \omega|s, \beta)]$ , is:

$$\hat{v}(\mu_s) = \begin{cases} k^2\mu_s^2 - k^2E[\omega^2|s, \beta] & \text{if } \mu_s < m, \\ -k^2(\hat{\mu} - k\mu_s - (1-k)\bar{\omega})^2 + k^2\mu_s^2 - k^2E[\omega^2|s, \beta] & \text{if } \mu_s \in [m, \hat{\mu}), \\ -(1-k)^2(\mu_s - \bar{\omega})^2 + k^2\mu_s^2 - k^2E[\omega^2|s, \beta] & \text{if } \mu_s \in (\hat{\mu}, \bar{\omega}), \\ k^2\mu_s^2 - k^2E[\omega^2|s, \beta] & \text{if } \mu_s > \bar{\omega}. \end{cases}$$

Let  $Y(\mu_s)$  be the concave closure of  $\hat{v}(\mu_s)$ , i.e., the smallest concave function that is everywhere weakly greater than  $\hat{v}(\mu_s)$ . By the analysis in Kamenica and Gentzkow (2011), the manager will implement a perfectly informative system ( $\beta \rightarrow \infty$ ), at date 1, if  $Y(\mu_s) > \hat{v}(\mu_s)$ . Now note that (i)  $\hat{v}(\mu_s) \leq k^2\mu_s^2 - k^2E[\omega^2|s, \beta]$  and (ii) the expectation of  $E[\omega^2|s, \beta]$  is constant across all Bayes plausible distributions of posteriors so that  $k^2\mu_s^2 - k^2E[\omega^2|s, \beta]$  can be treated as a constant plus the term  $k^2\mu_s^2$ , which is strictly convex in the posterior beliefs. It follows that  $Y(\mu_s) > k^2\mu_s^2 - k^2E[\omega^2|s, \beta] \geq \hat{v}(\mu_s)$ . The proof for the case  $\mu_0 > \bar{\omega}$  follows similar logic and is therefore omitted. ■

#### Formal proof of argument in footnote 42:

**Claim:** *Suppose that the manager can delay (or revise) the choice of  $x$  to date 2 (after observing  $s$ ) and that  $x$  is publicly observable. Then, there does not exist a fully revealing equilibrium.*

**Proof:** The proof is by contradiction. Suppose that there exists a fully revealing equilibrium in which, after observing a signal realization  $s$ , the manager chooses a fraction of informed users  $x_{t=2}(s|\beta)$  that is invertible in  $s$  everywhere; the informed users take actions that equal  $\mu(s|\beta)$ ; and the uninformed users take actions that equal  $E[\omega|x = x_{t=2}(s|\beta)] = E[\omega|s, \beta] = \mu(s|\beta)$ . For the purpose of developing a contradiction suppose that the manager observes a signal realization of  $s = s^o$  such that hers and the informed users' posterior expectation is  $\mu(s = s^o|\beta) < \hat{\omega}$  (similar arguments hold for the case of  $\mu(s = s^o|\beta) > \hat{\omega}$ ). If the manager chooses  $x^o \equiv x_{t=2}(s = s^o|\beta)$ , the uninformed users infer the signal, i.e.,  $E[\omega|x = x^o] = E[\omega|s = s^o, \beta] = \mu(s = s^o|\beta)$ . The actual aggregate action collapses to

$$\hat{A}(x = x^o, s = s^o, \beta) = x^o\mu(s = s^o|\beta) + (1 - x^o)E[\omega|x = x^o] = \mu(s = s^o|\beta),$$

whereas the aggregate action that the manager prefers is

$$A^*(s = s^o|\beta) \equiv E_{t=2}[A^*(\omega)|s = s^o, \beta] = k\mu(s = s^o|\beta) + (1 - k)\hat{\omega} > \mu(s = s^o|\beta).$$

Recall that the manager wants to eliminate the gap between the actual aggregate action and her most preferred action.<sup>47</sup> Next, note that there exists  $s^{oo}$ , such that (i)  $\mu(s = s^{oo}|\beta) \in (\mu(s = s^o|\beta), A^*(s = s^o|\beta))$ , (ii)  $x^{oo} \equiv x_{t=2}(s = s^{oo}|\beta) \neq x^o$  and (iii)  $E[\omega|x = x^{oo}] = E[\omega|s = s^{oo}, \beta] = \mu(s = s^{oo}|\beta)$ . Given the uninformed users strategy to take actions that equal  $E[\omega|x = x_{t=2}(s|\beta)] = E[\omega|s, \beta] = \mu(s|\beta)$  for any  $s$ , if the manager observes  $s^o$ , she prefers choosing  $x^{oo}$  and pooling with  $s^{oo}$  because doing so results in an aggregate action that is closer to her most preferred action:

$$\begin{aligned} A^*(s = s^o|\beta) &\geq \mu(s = s^{oo}|\beta) \\ &\geq x^{oo}\mu(s = s^o|\beta) + (1 - x^{oo})\mu(s = s^{oo}|\beta) \\ &= \hat{A}(x = x^{oo}, s = s^o, \beta) \\ &\geq \mu(s = s^o|\beta) \\ &= \hat{A}(x = x^o, s = s^o, \beta). \end{aligned}$$

This contradicts the existence of a fully revealing equilibrium. ■

**Proof of Lemma 4:** The users' actions are equal to the expected state given their information at the date they act because: (i) there is a continuum of users over unity and so the action of a single user does not affect the aggregate action  $A$ , and (ii) the delay cost is independent of the user's action. The actions of the uninformed users depend on whether they observed the actions of the informed users (and by doing so learned the signal  $s$ ) or did not observe the actions of the informed users. In the former case, their actions equal the posterior  $\mu(s|\beta)$  and in the latter case equal the prior  $\mu_0$ . The informed users' actions always equal the posterior  $\mu(s|\beta)$ . Consider a representative uninformed user  $i$  who conjectures that a fraction  $d \in [0, 1]$  of uninformed users acts late. Simplifying, the ex post payoff

---

<sup>47</sup>To see why, note that the manager's expected payoff conditional on the signal realization  $s$  at date 2 can be presented as  $E_{t=2}[v(\hat{A}(x, s, \beta), \omega)|s, \beta] = -\left(E_{t=2}[\hat{A}(x, s, \beta) - k\omega - (1 - k)\bar{\omega})|s, \beta]\right)^2 - \text{Var}(\hat{A}(x, s, \beta) - k\omega - (1 - k)\bar{\omega})|s, \beta) = -(\hat{A}(x, s, \beta) - E_{t=2}[A^*(\omega)|s, \beta])^2 - k^2\text{Var}(\omega|s, \beta)$ , where the last term is constant in  $x$ .

of a representative uninformed user  $i$  is

$$u_i(\cdot) = -(\mu_0 - \omega)^2 + \pi(x\mu(s|\beta) + (1-x)d\mu(s|\beta) + (1-x)(1-d)\mu_0 - \omega)^2$$

if he acts early and

$$u_i(\cdot) = -(\mu(s|\beta) - \omega)^2 + \pi(x\mu(s|\beta) + (1-x)d\mu(s|\beta) + (1-x)(1-d)\mu_0 - \omega)^2 - D$$

if he acts late. Then, for given  $\beta$ , at date 3.1, the representative uninformed user  $i$  compares his expected payoffs and acts early iff:

$$\begin{aligned} \Delta u_i(\beta) &\equiv E_{t=3.1}[(\mu(s|\beta) - \omega)^2 - (\mu_0 - \omega)^2] + D \\ &= -\frac{\beta}{\alpha(\alpha + \beta)} + D \\ &> 0, \end{aligned}$$

which happens when  $D \geq \frac{\beta}{\alpha(\alpha + \beta)}$ . Given that all uninformed users are identical, then for given  $\beta$ , a symmetric equilibrium arises in which all uninformed users act early if  $D \geq \frac{\beta}{\alpha(\alpha + \beta)} \equiv D(\beta)$ . Otherwise they act late. The comparative statics of  $D(\beta)$  is straightforward. ■

**Proof of Proposition 7:** Using the proof of Lemma 4, the manager compares the value of the following three optimization programs:

**Program  $\mathcal{P}(1)$ :**

$$\begin{aligned} &\max_{\beta \geq 0, x \in [0,1]} V(x, \beta) \\ \text{s.t. } &D \geq \frac{\beta}{\alpha(\alpha + \beta)} \end{aligned}$$

**Program  $\mathcal{P}(2)$ :**

$$\max_{\beta \geq 0} V(x = 1, \beta)$$

If  $D \geq \frac{1}{\alpha}$  the constraint in program  $\mathcal{P}(1)$  is always satisfied for any  $\beta$ . Hence, by the proof of Proposition 1, under  $\mathcal{P}(1)$ , the manager chooses  $(k, \infty)$ , which is identical to the solution of the baseline case under UD. The solution of  $\mathcal{P}(2)$  is identical to the solution of the baseline

case under MD. By Proposition 2, the manager prefers UD over MD. Hence if  $D \geq \frac{1}{\alpha}$ , the manager chooses  $(k, \infty)$  and the uninformed users act early (as well as the informed users by assumption).

If  $D < \frac{1}{\alpha}$  the constraint of program  $\mathcal{P}(1)$  is not satisfied for the solution of the unconstrained program  $(k, \infty)$ . To satisfy the constraint the manager can set a precision level  $\beta(D)$  that just satisfies the constraint, i.e.,  $\beta(D) = \frac{\alpha^2 D}{1 - \alpha D}$ . Note that  $\frac{d\beta(D)}{dD} > 0$ . Recall that  $V(x, \beta) = -(1 - k)^2(\mu_0 - \bar{\omega})^2 - \frac{k^2}{\alpha} + x(2k - x)\frac{\beta}{\alpha(\alpha + \beta)}$ . Substituting for  $\beta = \beta(D)$ ,

$$V(x, \beta = \beta(D)) = -(1 - k)^2(\mu_0 - \bar{\omega})^2 - \frac{k^2}{\alpha} + x(2k - x)\frac{\alpha^2 D}{\alpha(1 - \alpha D)(\alpha + \frac{\alpha^2 D}{1 - \alpha D})}.$$

Then taking derivative with respect to  $x$ , I get  $\frac{\partial V(x, \beta = \beta(D))}{\partial x} \propto k - x$  and so the manager sets  $x = k$  and the value of program  $\mathcal{P}(1)$  is  $-(1 - k)^2(\mu_0 - \bar{\omega})^2 - \frac{k^2}{\alpha} + k^2\frac{\alpha^2 D}{\alpha(1 - \alpha D)(\alpha + \frac{\alpha^2 D}{1 - \alpha D})}$ . The solution of  $\mathcal{P}(2)$  is identical to the solution of the baseline case under MD. The value of  $\mathcal{P}(2)$  is  $-(1 - k)^2(\mu_0 - \bar{\omega})^2 - \frac{k^2}{\alpha} + \max\{\frac{2k-1}{\alpha}, 0\}$ . It is straightforward to see that the value of program  $\mathcal{P}(1)$  exceeds that of program  $\mathcal{P}(2)$  if  $D \geq \max\{0, \frac{2k-1}{\alpha k^2}\}$ ; and vice versa. ■

**Proof of Corollary 3:** By Proposition 7, for any  $D < \max\{0, \frac{2k-1}{\alpha k^2}\}$  (which is feasible only if  $k > \frac{1}{2}$ ), all users learn the signal and the manager chooses  $\beta \rightarrow \infty$  under both regimes. Hence  $W_{UD}^A - W_{MD}^A = 0$ . I next consider the case  $D \geq \max\{0, \frac{2k-1}{\alpha k^2}\}$ . Under UD, the manager disseminates the signal to a fraction  $k$  of users and chooses  $\beta^A = \beta(D)$  if  $D \in [\frac{2k-1}{\alpha k^2}, \frac{1}{\alpha}]$  or  $\beta^A \rightarrow \infty$  if  $D > \frac{1}{\alpha}$ . The users' welfare under UD is then:

$$\begin{aligned} W_{UD}^A &= kE_{t=1} \left[ -(\mu(s|\beta^A) - \omega)^2 + \pi(k\mu(s|\beta^A) + (1 - k)\mu_0 - \omega)^2 \right] \\ &\quad + (1 - k)E_{t=1} \left[ -(\mu_0 - \omega)^2 + \pi(k\mu(s|\beta^A) + (1 - k)\mu_0 - \omega)^2 \right] \\ &= -\frac{k}{\alpha + \beta^A} - \frac{(1 - k)}{\alpha} + \pi \text{Var}(k\mu(s|\beta^A) - \omega) \\ &= -\frac{k}{\alpha + \beta^A} - \frac{(1 - k)}{\alpha} + \pi \left( \frac{1}{\alpha} - (2 - k)k \frac{\beta^A}{\alpha(\alpha + \beta^A)} \right) \\ &= -\frac{\alpha(1 - \pi) + \beta^A(1 - k)(1 - \pi(1 - k))}{\alpha(\alpha + \beta^A)} \\ &< 0. \end{aligned}$$

Under MD the manager publicly disseminates the signal and chooses  $\beta \rightarrow \infty$  ( $\beta = 0$ ) if

$k > \frac{1}{2}$  ( $k \leq \frac{1}{2}$ ). Hence the users' welfare under MD is

$$W_{MD}^A = \begin{cases} E_{t=1} [-(\omega - \omega)^2 + \pi(\omega - \omega)^2] = 0 & \text{if } k > \frac{1}{2}, \\ E_{t=1} [-(\mu_0 - \omega)^2 + \pi(\mu_0 - \omega)^2] = -\frac{(1-\pi)}{\alpha} & \text{if } k \leq \frac{1}{2}. \end{cases}$$

If  $k > \frac{1}{2}$ , then  $W_{UD}^A < W_{MD}^A$ . Now consider the case  $k \leq \frac{1}{2}$ . Then,  $W_{UD}^A - W_{MD}^A \propto 1 - \pi(2-k)$ . Hence  $W_{UD}^A > W_{MD}^A$  if  $\pi < \frac{1}{2-k} \equiv \bar{\pi}$  and  $W_{UD}^A \leq W_{MD}^A$  if  $\pi \geq \bar{\pi}$ . ■

**Proof of Proposition 8:** Let  $o_i(s, \beta, \xi_i)$  denote the action of informed user  $i$  as observed by his peers. Upon observing  $Z$ , the uninformed users update their beliefs about  $\omega$  and, at date 3.2, take actions

$$a^N(\mu_0, Z) = E_{t=3.2}[\omega|Z] = \frac{\alpha\mu_0 + \theta \sum_{i=1}^z o_i(s, \beta, \xi_i)}{\alpha + z\theta}$$

where the superscript "N" denotes noisy action and  $\theta \equiv \frac{(\alpha+\beta)\gamma}{\alpha+\beta+\gamma}$  is the precision with which the observed actions of the informed users reflect the state  $\omega$ . To see why, note that  $o_i(s, \beta, \xi_i) = \omega + \frac{\beta\delta - \alpha\varepsilon}{\alpha+\beta} + \xi_i$  and so  $Var(\frac{\beta\delta - \alpha\varepsilon}{\alpha+\beta} + \xi_i) = \frac{1}{\alpha+\beta} + \frac{1}{\gamma} = \frac{\alpha+\beta+\gamma}{(\alpha+\beta)\gamma}$ . In line with the previous analysis,  $A$  is the average action:

$$\begin{aligned} A^N(x, s, \beta) &= \frac{z}{N}\mu(s|\beta) + \left(1 - \frac{z}{N}\right) a^N(\mu_0, Z) \\ &= \underbrace{x\mu(s|\beta) + (1-x)\mu_0}_{\text{baseline average action}} + (1-x) \underbrace{\frac{xN\theta}{(\alpha + xN\theta)} \left( \frac{\beta}{(\alpha + \beta)}(\varepsilon + \delta) + \eta_z \right)}_{\equiv n(x, \beta) \quad (\text{additional term due to spillover})} \end{aligned}$$

where  $\eta_z \equiv \frac{\sum_{i=1}^z \xi_i}{z} \sim \mathcal{N}\left(0, \frac{1}{z\gamma}\right)$ . Now, in addition to the average action that the users take in the baseline setting, there is an additional term  $n(x, \beta)$  that arises due to spillover. At date 1 the manager maximizes her expected utility:

$$\begin{aligned} V^N(x, \beta) &= -E_{t=1}[(A^N(x, s, \beta) - k\omega - (1-k)\bar{\omega})^2] \\ &= \underbrace{x \frac{\beta}{\alpha(\alpha + \beta)} \left(1 + \frac{(1-x)N\theta}{\alpha + xN\theta}\right) \left(2k - x \left(1 + \frac{(1-x)N\theta}{\alpha + xN\theta}\right)\right) - \frac{(1-x)^2 x N \theta^2}{\gamma(\alpha + xN\theta)^2}}_{\equiv A1(x, \beta)} \\ &\quad - \underbrace{\frac{k^2}{\alpha} - (1-k)^2(\mu_0 - \bar{\omega})^2}_{\equiv A2 \quad (\text{independent of } x \text{ and } \beta)}. \end{aligned}$$

Given that  $A2$  is a constant, the manager's choice of  $x$  and  $\beta$  satisfies

$$(x^N, \beta^N) = \arg \max_{\beta \geq 0, x \in [0,1]} A1(x, \beta).$$

First, note that

$$\frac{\partial A1(x, \beta)}{\partial \beta} = \frac{xT}{\gamma(\alpha + \beta)^2(\alpha + \theta Nx)^3} \propto T,$$

where

$$\begin{aligned} T \equiv & \underbrace{\gamma(\alpha + N\theta)(\alpha + Nx\theta)}_{\geq 0}(\alpha(2k - x) + (2k - 1)Nx\theta) \\ & + 2 \underbrace{\frac{\partial \theta}{\partial \beta}(\alpha + \beta)N(1 - x)\beta Nx\theta(\theta(1 - x) - \gamma(1 - k))}_{\geq 0} \\ & + 2 \underbrace{\frac{\partial \theta}{\partial \beta}(\alpha + \beta)N(1 - x)\alpha(N\theta^2(1 - x)x + \beta\gamma(k - x))}_{\geq 0}. \end{aligned}$$

Sufficient conditions for  $T > 0$  are

$$\alpha(2k - x) + (2k - 1)Nx\theta > 0$$

$$\theta(1 - x) - \gamma(1 - k) > 0$$

$$N\theta^2(1 - x)x + \beta\gamma(k - x) > 0,$$

which are all satisfied if  $x \leq \bar{x} \equiv \min \left\{ \frac{2\alpha k}{\alpha + (1 - 2k)N\theta}, 1 - \frac{\gamma(1 - k)}{\theta}, k \right\}$ . This is feasible if  $\bar{x} > 0$ .

Recall that  $\theta = \frac{(\alpha + \beta)\gamma}{\alpha + \beta + \gamma}$  and note that:

- 1)  $\frac{2\alpha k}{\alpha + (1 - 2k)N\theta}$  is decreasing in  $\gamma$  and  $\lim_{\gamma \rightarrow 0} \frac{2\alpha k}{\alpha + (1 - 2k)N\theta} = 2k > 0$ ;
- 2)  $1 - \frac{\gamma(1 - k)}{\theta} > 0$  is decreasing in  $\gamma$  and  $\lim_{\gamma \rightarrow 0} 1 - \frac{\gamma(1 - k)}{\theta} = k > 0$ ;
- 3)  $k > 0$  by assumption.

Hence there exist  $\gamma^o > 0$  and  $\gamma^{oo} > 0$  such that  $x \leq \bar{x}$  is feasible if  $\gamma < \min\{\gamma^o, \gamma^{oo}\}$ .

Substituting for  $\beta \rightarrow \infty$ , the optimal fraction satisfies

$$L(x, \gamma) \equiv \lim_{\beta \rightarrow \infty} \frac{\partial A1(x, \beta)}{\partial x} = \frac{2\alpha(k - x) - \gamma N(1 - 2k + x^2)}{(\alpha + \gamma Nx)^2} = 0, \quad (26)$$

which yields  $x(\gamma) = \frac{\sqrt{(\alpha + \gamma N)(\alpha - \gamma(1 - 2k)N) - \alpha}}{\gamma N}$ . I note that  $x(\gamma) < k$  because  $L(x = k, \gamma) = -\frac{\gamma N(1 - k)^2}{(\alpha + \gamma Nk)^2} < 0$  (as  $\gamma > 0$ ,  $N > 0$  and  $k < 1$ ) and  $x(\gamma) > 0$  because  $L(x = 0, \gamma) = \frac{2\alpha k - \gamma(1 - 2k)N}{\alpha^2} > 0$  if  $\gamma < \gamma_o$  (as  $\frac{2\alpha k - \gamma(1 - 2k)N}{\alpha^2}$  is decreasing in  $\gamma$  and  $\lim_{\gamma \rightarrow 0} \frac{2\alpha k - \gamma(1 - 2k)N}{\alpha^2} =$

$\frac{2k}{\alpha} > 0$ ). Next, I note that  $x(\gamma)$  is a decreasing function, because, using the Implicit Function Theorem,

$$\begin{aligned}
\frac{dx(\gamma)}{d\gamma} &\propto \frac{\partial L(x, \gamma)}{\partial \gamma} \\
&= \frac{N(\gamma N x^N (1 - 2k + (x^N)^2) + \alpha(-1 + k(2 - 4x^N) + 3(x^N)^2))}{(\alpha + \gamma N x^N)^3} \\
&= -\frac{2\alpha(\alpha + \gamma k N - \sqrt{(\alpha + \gamma N)(\alpha - \gamma(1 - 2k)N))}}{\gamma^2 N(\alpha + \gamma N)(\alpha - \gamma(1 - 2k)N)} \quad \text{by (26)} \\
&< 0 \quad \text{if } \gamma < \gamma_o.
\end{aligned}$$

As a last step I verify that  $x(\gamma) < \bar{x}$ . This is because (i)  $\lim_{\beta \rightarrow \infty} \bar{x} = \frac{2\alpha k}{\alpha + \gamma N(1 - 2k)}$  and (ii)  $L(x = \frac{2\alpha k}{\alpha + \gamma N(1 - 2k)}, \gamma) = \frac{-2\alpha^2 k - \alpha \gamma (1 - 2k)^2 N - \gamma^2 (1 - 2k)^3 N^2}{\alpha^2 (\alpha + \gamma N)} < 0$  if  $\gamma < \gamma_{oo}$ .

Let  $\bar{\gamma} \equiv \min\{\gamma^o, \gamma^{oo}, \gamma_o, \gamma_{oo}\}$ . It follows that, if  $\gamma < \bar{\gamma}$ , the optimal solution is  $\beta^N \rightarrow \infty$  and  $x(\gamma) \in (0, k)$  where  $x(\gamma)$  is decreasing in  $\gamma$ . ■

## References

- Agrawal, A., S. Chadha, and M. Chen. 2006. Who is afraid of Reg FD? The behavior and performance of sell-side analysts following the SEC's fair disclosure rules. *Journal of Business* 79(6): 2811-2834.
- Alonso, R., and O. Camara. 2014. *Persuading voters*. Working paper.
- Arya, A., J. Glover, and K. Sivaramakrishnan. 1997. Commitment issues in budgeting. *Journal of Accounting Research* 35: 273-278.
- Arya, A., J. Glover, B. Mittendorf, and G. Narayanamoorthy. 2005. Unintended consequences of regulating disclosures: the case of Regulation Fair Disclosure. *Journal of Accounting and Public Policy* 24(3): 243-252.
- Baiman, S. 1975. The evaluation and choice of internal information systems within a multi-person world. *Journal of Accounting Research* 13: 1-15.
- Beyer, A., and I. Guttman. 2011. The effect of trading volume on analysts' forecast bias. *The Accounting Review* 86(2): 451-481.
- Beyer, A., D. Cohen, T. Lys, and B. Walther. 2010. The financial reporting environment: review of the recent literature. *Journal of Accounting and Economics* 50: 296-343.
- Bushman, R. 1991. Public disclosure and the structure of private information markets. *Journal of Accounting Research* 29(2): 261-276.
- Che, Y., and N. Kartik. 2009. Opinions as incentives. *Journal of Political Economy* 117 (5): 815-860.
- Chen, Q., T. Lewis, K. Schipper, and Y. Zhang. 2016. Uniform versus Discretionary Regime in Reporting Information with Unverifiable Precision and a Coordination Role. *Journal of Accounting Research*. Forthcoming.
- Dutta, S. 1996. Private and public disclosures and the efficiency of stock prices. *Review of Accounting Studies* 1 (4): 285-307.
- Demski, J., and G. Feltham. 1994. Market response to financial reports. *Journal of Accounting and Economics* 17: 3-40.
- Dye, R. 2001. An evaluation of "Essays on disclosure" and the disclosure literature in accounting. *Journal of Accounting and Economics* 32: 181-235
- Edmans, A., M. Heinle, and C. Huang. 2016. The Real Costs of Financial Efficiency When Some Information Is Soft. *Review of Finance* 20(6): 2151-2182.
- Fischer, P., and P. Stocken. 2010. Analyst information acquisition and communication. *The Accounting Review* 85(6): 1985-2009.

- Francis, J., D. Nanda, X., and Wang. 2006. Re-examining the effects of Regulation Fair Disclosure using foreign listed firms to control for concurrent shocks. *Journal of Accounting and Economics* 41 (3): 271-292.
- Gao, P., and P. Liang. 2013. Informational feedback effect, adverse selection, and the optimal disclosure policy. *Journal of Accounting Research* 51(5): 1133-1158.
- Gentzkow, M., and E. Kamenica. 2014. Costly persuasion. *American Economic Review* 104(5): 457-62.
- Göx, R., and A. Wagenhofer. 2009. Optimal impairment rules. *Journal of Accounting and Economics* 48(1): 2-16.
- Harsanyi, J. 1955. Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy* 63(4): 309-321.
- Heflin, F., K. R. Subramanyam, and Y. Zhang. 2003. Regulation FD and the financial information environment: Early evidence. *The Accounting Review* 78: 1-37.
- Herrmann, D., O. Hope, and W. Thomas. 2008. International diversification and forecast optimism: the effects of Reg FD. *Accounting Horizons* 22 (2): 179-197.
- Hughes, J., and S. Pae, 2004. Voluntary disclosure of precision information. *Journal of Accounting and Economics* 37(2): 261-289.
- Hong, H., J. Kubik, and A. Solomon. 2000. Security analysts' career concerns and herding of earnings forecasts. *The RAND Journal of Economics* 31: 121-144.
- Jorgensen, B., J. Li, and N. Melumad. 2013. *Why do firms gravitate to selective disclosure*. Working paper.
- Kamenica, E., and M. Gentzkow. 2011. Bayesian persuasion. *American Economic Review* 101: 2590-2615.
- Kim, O., and R. Verrecchia. 1991. Market reaction to anticipated announcements. *Journal of Financial Economics* 30: 273-309.
- McNichols, M., and B. Trueman. 1994. Public disclosure, private information collection, and short-term trading. *Journal of Accounting and Economics* 17: 69-94.
- Melumad, N., and T. Shibano. 1991. Communication in settings with no transfers. *The RAND Journal of Economics* 22(2): 173-198.
- O'Brien, P., M. McNichols, and H. Lin. 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43: 623-650.
- Pae, S. 1999. Acquisition and discretionary disclosure of private information and its implications for firms' productive activities. *Journal of Accounting Research* 37: 465-474.

- Penno, M. 1996. Unobservable precision choices in financial reporting. *Journal of Accounting Research* 34: 141-149.
- Sabino, J. 1993. *Analysts incentives to sell information*. Working paper.
- Shavell, S. 1991. Acquisition and disclosure of information prior to sale. *RAND Journal of Economics* 25(1): 20-36.
- Stocken, P. 2012. Strategic accounting disclosure. *Foundations and Trends in Accounting* 7(4): 197-291.
- Titman, S., and B. Trueman. 1986. Information quality and the valuation of new issues. *Journal of Accounting and Economics* 8: 159-72.
- Verrecchia, R. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32: 97-180.
- Wang, Y. 2014. *Bayesian persuasion with multiple receivers*. Working paper.
- Wu, J., and A. Zang. 2009. What determines financial analysts' career outcomes during mergers. *Journal of Accounting and Economics* 47: 59-86.