Opinions as Incentives*

Yeon-Koo Che† Navin Kartik‡

August 21, 2009

Abstract

We study a model where a decision maker (DM) must rely on an adviser for information about the state of the world relevant for her decision. The adviser has the same underlying preferences as the DM; he may differ, however, in his prior belief about the state, which we interpret as difference of opinion. We derive a tradeoff for the DM: an adviser with greater difference of opinion has greater incentives to acquire information, but reveals less of any information she acquires, via strategic disclosure. The difference of opinion engenders two novel incentives for an agent to acquire information: a "persuasion" motive and a motive to "avoid prejudice." When the DM can choose an adviser from a rich pool of opinion types including a like-minded one, it is optimal to choose an adviser with at least some difference of opinion. Delegation can be demotivating because it eliminates the need for the adviser to persuade and avoid prejudice. We also study the relationship between difference of opinion and difference of preference.

*We would like to acknowledge the input of Jimmy Chan at early stages of this project. We thank Nageeb Ali, Heski Bar-Iassac, Roland Benabou, Oliver Board, Arnaud Costinot, Vince Crawford, Wouter Dessein, Jean Guillaume Forend, Michihiro Kandori, Kohei Kawamura, Li Hao, Bart Lipman, Eric Maskin, Roger Myerson, Carolyn Pitchik, Jennifer Reinganum, Mike Riordan, Ed Schlee, Richard Schmalensee, Joel Sobel, Eric Van den Steen, and various audiences for their opinions. Canice Prendergast and two anonymous referees provided insightful comments that improved the paper. We also received excellent research assistance from David Eil, Chulyoung Kim, Uliana Loginova, and Petra Persson. Che is grateful to the KSEF's World Class University Grant (#R32-2008-000-10056-0) for financial support. Kartik is grateful to the Institute for Advanced Study at Princeton (the Roger W. Ferguson, Jr. and Annette L. Nazareth membership) and the National Science Foundation (Grant SES-0720893) for funding; he also thanks the Institute for its hospitality.

†Columbia University and Yonsei University; yc2271@columbia.edu.

‡Columbia University; nkartik@columbia.edu.

"Difference of opinion leads to enquiry." — Thomas Jefferson

1 Introduction

To an average 17th century (geocentric) person, the emerging idea of the earth moving defied common sense. If the earth revolves, then "why would heavy bodies falling down from on high go by a straight and vertical line to the surface of the earth... [and] not travel, being carried by the whirling earth, many hundreds of yards to the east?" (Galilei, 1953, p. 126). In the face of this seemingly irrefutable argument, Galileo Galilei told a famous story, via his protagonist *Salviati* in *Dialogue Concerning the Two Chief World Systems*, about how an observer locked inside a boat, sailing at a constant speed without rocking, cannot tell whether the boat is moving or not. This story, meant to persuade critics of heliocentrism, became a visionary insight now known as the *Galilean Principle of Relativity*.

This example dramatically illustrates how a different view of the world (literally) might lead to an extraordinary discovery. But the theme it captures is hardly unique. Difference of opinion is valued in many organizations and situations. A prominent rationale for corporations to seek diversity in their workforce is to tap creative ideas. Academic research thrives on the pitting of opposing hypotheses. Government policy failures are sometimes blamed on the lack of a dissenting voice in the cabinet, a phenomenon known as "groupthink" (Janis, 1972). Debates between individuals can be more illuminating when they have differing views; in the absence of any difference, one may try to mimic such an environment by playing "devil's advocate."

Difference of opinion would be obviously valuable if it inherently entails a productive advantage in the sense of bringing new ideas or insights that would otherwise be unavailable. But could it be valuable even when it brings no direct productive advantage? Moreover, are there any costs of people having differing opinions? This paper explores these questions by examining incentive implications of difference of opinion.

We study a setting in which a decision maker, or DM for short, consults an adviser before making a decision. Both individuals' payoff from the decision depends on some exogenous state of the world. We model the decision and the state as real numbers, where the DM's payoff-maximizing decision is equal to the state. At the outset, however, neither the DM nor the adviser knows the state, they only hold some prior views about it. The adviser can exert costly effort to try and discover an informative signal about the state; the probability of observing such a signal is increasing in his effort. The signal could take the form of scientific evidence obtainable by conducting an experiment, witnesses or documents locatable by investigation, a mathematical proof, or a convincing insight that can reveal something about the state. Effort is unverifiable, however, and higher effort imposes a greater cost on the adviser. After the adviser privately observes the information, he strategically communicates with the DM. Communication takes the form of verifiable disclosure: sending a message is costless, but the adviser cannot falsify information, or equivalently, the DM can judge objectively what a signal means. The adviser's strategic choice therefore is whether or not to reveal any information he has acquired. Finally, the DM makes her decision given her updated beliefs after communication with the adviser.

This framework captures common themes encountered by many organizations. For instance, managers solicit information from employees; political leaders seek the opinion of their cabinet members; scientific boards consult experts; and journal editors rely on referees. The model also applies more broadly to some situations where there may be no tightly circumscribed organization or no particular decision to be made: the DM could be any audience such as a political constituency, lower courts in a judicial system, the scientific community or the general public (such as 17th century intelligent laymen), and the "decision" is just the opinion or belief that this audience forms on some matter. Correspondingly, the adviser could be a politician, a supreme court justice, investigator, or a scientist (such as Galileo) who cares about the belief that the audience holds. Such interactions do not involve any contracting relationship between the DM and the adviser; indeed, the DM does not even hire the adviser per se, nor does communication take place in an explicit protocol.

It is often the case, as in the examples mentioned above, that an adviser is interested in the decision made by DM. We assume initially that the adviser has the same fundamental preferences as the DM about which decision to make in each state, but that he may have a difference of opinion about what the unknown state is likely to be. More precisely, the adviser may disagree with the DM about the prior probability distribution of the unknown state, and this disagreement is common knowledge. That is, they "agree to disagree." Although game-theoretic models often assume a common prior, referred to as the *Harsanyi Doctrine*, there is a significant and growing literature that analyzes games with heterogenous priors.[1] Such an open disagreement may arise from various sources: individuals may

[1] Morris (1995) addresses some conceptual issues about non-common prior models and discusses why they can be useful.

simply be endowed with different prior beliefs (just as they may be endowed with different preferences), or they may update certain kinds of public information differently based on psychological, cultural, or other factors (Tversky and Kahneman, 1974; Aumann, 1976; Acemoglu, Chernozhukov, and Yildiz, 2007). Whatever the reason, open disagreement of beliefs are commonplace in practice,[2] and is sometimes more plausible than fundamentally divergent preferences, as has also been argued by Banerjee and Somanathan (2001) and Dixit and Weibull (2007). For instance, consider a firm that must decide which of two technologies to invest in. All employees share the common goal of investing in the better technology, but no one knows which this is. Different employees may hold different beliefs about the viability of each technology, leading to open disagreements about where to invest.

Specifically, we model the adviser's opinion as the mean of his (subjective) prior about the state, normalizing the DM's opinion to mean zero. We suppose that there is a rich pool of possible advisers in terms of their opinion, and advisers are differentiated only by their opinion, meaning that a difference of opinion does not come with better ability or lower cost of acquiring information. This formulation allows us to examine directly whether difference of opinion alone can be valuable to the DM, even without any direct productive benefits.[3]

Our main results concern a tradeoff associated with difference of opinion. To see the intuition, suppose first that effort is not a choice variable for the adviser. In this case, the DM has no reason to prefer an adviser with a differing opinion. In fact, unless the signal is perfectly informative about the state, the DM will strictly prefer a like-minded adviser—i.e., one with the same opinion as she has. This is because agents with different opinions will hold different posteriors about what the right decision is given a partially-informative signal. Consequently, even though he shares the same fundamental preferences, an adviser with a differing opinion will typically withhold some information from the DM. This strategic withholding of information entails a welfare loss for the DM, whereas no such loss will arise if the adviser is like-minded.

When effort is endogenous, the DM is also concerned with the adviser's incentive to exert effort; all else equal, she would prefer an adviser who will exert more effort. We find that differences of opinion create incentives for information acquisition, for two distinct

[2] To mention just two examples, consider very public disagreements about how serious the global warming problem is and how to protect a country against terrorism.

[3] As previously noted, individuals with different backgrounds and experiences are also likely to bring different approaches and solutions to a problem, which may directly improve the technology of production. We abstract from this in order to focus on the incentive implications of difference of opinion.

reasons. First, *an adviser with a difference of opinion is motivated to persuade the DM.* Such an adviser believes that the DM's opinion is wrong, and that by acquiring a signal, he is likely to move the DM's decision towards what he perceives to be the right decision. This motive does not exist for a like-minded adviser. Second, and more subtle, *an adviser with a difference of opinion will exert effort to avoid rational "prejudice."* Intuitively, in equilibrium, an adviser withholds information that is contrary to his opinion, for such information will cause the DM to take an action that the adviser dislikes. Recognizing this, the DM discounts the advice she receives and chooses an action contrary to the adviser's opinion, unless the advice is corroborated by hard evidence—this equilibrium feature of strategic interaction is what we call a "prejudicial effect." Consequently, an adviser with a difference of opinion has incentives to seek out information in order to avoid an adverse inference from the DM, a motive that does not exist for a like-minded adviser.

In summary, we find that *difference of opinion entails a loss of information through strategic communication, but creates incentives for information acquisition.* This tradeoff resonates with common notions that, on the one hand, diversity of opinion causes increased conflict because it becomes harder to agree on solutions—this emerges in our analysis as worsened communication; on the other hand (as was recognized by Jefferson, quoted in our epigraph) it induces increased efforts to convince other individuals, which can lead to improved collective understanding—this emerges here as increased information acquisition. This tradeoff sheds light on the nature of information acquisition and transmission in a general communication setting involving differences of opinion. The positive incentive effect suggests why it was probably not a coincidence that the principle of relativity was developed by an individual such as Galileo Galilei, given his heliocentric view and the appeal of that principle towards making his view credible. At the same time, our theory also suggests why it may have been rational for the general public to be slow in embracing his view.

Equipped with this central tradeoff, we then refine our analysis to obtain two results that apply specifically to organization economics. Suppose first that the DM can indeed choose an adviser from a rich pool of different opinion types (including a like-minded type). Should she select an adviser with a different opinion or a like-minded one? Answering this question requires resolving the tension between information acquisition and transmission. We find that *the DM should select an adviser with some difference of opinion over a perfectly like-minded one.* The reason is that an adviser with a sufficiently small difference of opinion engages in only a negligible amount of strategic withholding of information, so the loss associated with such an adviser is negligible. By the same token, the prejudicial effect and its beneficial impact on information acquisition is also negligible when the difference of opinion is small. In contrast, the persuasion motive that even a slight difference of opinion generates—and thus the benefit the DM enjoys from its impact on increased effort—is non-negligible by comparison. Therefore, the DM derives a net benefit from an adviser with at least a little difference in opinion.

Second, if decision-making can be delegated to the chosen adviser, what are the costs and benefits of delegation, and will it ever be optimal for the DM to cede authority? This question of whether decisions should be made by uninformed principals or delegated to agents with better access to information is of obvious importance. The seminal work of Aghion and Tirole (1997) shows that delegating formal authority to an adviser with a conflict of interest may lead to undesirable decisions ex-post ("loss of control"), but has the benefit of empowering the agent to acquire more information ("increased initiative"). An implication is that it would be better to avoid having a conflict of interest, when possible. Our analysis delivers a complementary perspective. In a nutshell, we argue that lack of congruence (in terms of prior opinions, but also, to a lesser degree, preferences) can be beneficial to an organization, and this benefit can be harnessed only when authority remains in the hands of the principal. The reason is that delegation can be *demotivating* for the adviser because it eliminates both incentive effects we have highlighted: the desire to persuade the DM and to avoid prejudice. The conclusion that emerges is a more nuanced view of how delegation affects initiative from the agent.

While we focus primarily on difference of opinion, we augment the model later in the paper to allow the adviser to also differ from the DM in preferences over decisions. Heterogeneous preferences have a similar effect to difference of opinion on strategic disclosure. This implies that the incentive to acquire information to avoid prejudice is present even when the adviser shares the DM's opinion but has a different preference (hence, the demotivating effect of delegation carries over in part). Nevertheless, there is one crucial distinction between opinions and preferences: while an adviser with a difference of opinion has a persuasion motive to acquire information—he expects to systematically shift the DM's decision closer to his preferred decision—an adviser with only a difference of preference has no such expectation, and thus is less motivated to acquire information. When combined with the loss from communication distortion, this turns out to imply that having an adviser who differs only slightly in preferences need not be a net benefit to the DM, unlike the case of a small difference of opinion.

Nevertheless, we find that a difference of preferences can be valuable in the presence of a difference of opinion. In other words, an adviser with a different opinion has more incentive to acquire information if he also has a preference bias in the direction congruent to his opinion. This complementarity between preference and opinion implies that the incentive effect on information acquisition will be larger when the adviser is a *zealot*—one who believes that evidence is likely to move the DM's action in the direction of his preference bias—than when he is a *skeptic*—one who is doubtful that information about the state of the world will support his preference bias.

Our work builds on the literature on strategic communication, combining elements from the structure of conflicts of interest in Crawford and Sobel (1982) with a verifiable disclosure game following Grossman (1981) and Milgrom (1981). The key difference with much of this literature is that we endogenize the acquisition of information, which allows us to study how information acquisition and transmission is affected by the conflict of interest and differences of prior beliefs.[4] We postpone a detailed discussion of the closely related literature until after a full development of our model and analysis.

The paper is organized as follows. The next Section presents the baseline model with differences of opinion. Section 3 analyzes the disclosure sub-game, identifying the prejudicial effect of strategic communication. Section 4 develops the incentive benefits of this prejudicial effect and also identifies the persuasion motive. In Section 5, we consider heterogeneous preferences. Section 6 focusses on issues of delegation and participation constraints. We discuss robustness to modeling variations in Section 7, and conclude with some broader applications in Section 8. The Appendix contains all proofs that are omitted from the main text.

2 Model

A decision maker (DM) must take an action, $a \in \mathbb{R}$. Her payoff from the action depends on an unknown state of the world, $\omega \in \mathbb{R}$. The DM lacks the necessary expertise, or finds it prohibitively costly to directly acquire information about the state, but may rely on an adviser for information. As discussed in the introduction, there is no presumption initially that the DM hires or selects the adviser, although we shall later consider this choice as well. Throughout, subscripts of DM and A refer to the decision maker and adviser, respectively.

[4]Early papers that also consider endogenous information acquisition include Matthews and Postlewaite (1985) and Shavell (1994); they address different issues.

Prior Beliefs. We allow individuals to have different prior beliefs about the state. Specifically, while all individuals know that the state is distributed according to a Normal distribution with variance $\sigma_0^2 > 0$, individual $i = DM, A$ believes the mean of the distribution to be μ_i. The prior beliefs of each person are common knowledge. We will refer to an adviser's prior belief as his *opinion* or *type*, even though it is not private information. Without loss of generality, we normalize the DM's prior to $\mu_{DM} = 0$. An adviser with $\mu_A = 0$ is said to be *like-minded*; an adviser with $\mu_A \neq 0$ has a *difference of opinion* (with the DM).

Preferences. Each player $i = DM, A$ has the same von Neumann-Morgenstern state-dependent payoff from the DM's decision:

$$u_i(a, \omega) := -(a - \omega)^2.$$

Thus, were the state ω known, players would agree on the optimal decision $a = \omega$. In this sense, there is no fundamental preference conflict. We allow for such conflicts in Section 5. The quadratic loss function is a common specification in the literature: it captures the substantive notion that decisions are progressively worse the further away they are from the true state, and technically, makes the analysis tractable.

Information Acquisition. Regardless of the adviser's type, his investigation technology is the same, described as follows. He chooses the probability that his investigation is successful, $p \in [0, \bar{p}]$, where $\bar{p} < 1$, at a personal cost $c(p)$. The function $c(\cdot)$ is smooth, $c''(\cdot) > 0$, and satisfies the Inada conditions $c'(0) = 0$ and $c'(p) \to \infty$ as $p \to \bar{p}$. We will interchangeably refer to p as an effort level or a probability. With probability p, the adviser obtains a signal about the state, $s \sim N(\omega, \sigma_1^2)$. That is, the signal is drawn from a Normal distribution with mean equal to the true state and variance $\sigma_1^2 > 0$. With complementary probability $1 - p$, he receives no information (or equivalently, a completely uninformative signal), denoted by $\emptyset$. The binary precision levels—either informative or uninformative—simplify the analysis, but do not affect the qualitative nature of our results, as will be discussed in Section 7.

Communication. After privately observing the outcome of his investigation, the chosen adviser strategically discloses information to the DM. The signal s is "hard" or non-falsifiable. Hence, the adviser can only withhold the signal if he has obtained one; if he did not receive a signal, he has no strategic choice to make. Non-manipulability of the signal may represent large penalties against fraud, information being easily verifiable by the DM once received (even though impossible to acquire directly herself), or technological constraints on manipulation. This particular form of manipulating information, while often employed in the literature,[5] is admittedly stark. We discuss in Section 7 why our main insights are robust to allowing for other ways in which the adviser can manipulate his information when communicating to the DM.

Contracts. We adopt the common approach of incomplete contracting Grossman and Hart (1986) by positing that the DM cannot use monetary transfers that are contingent on the information or effort provided by the adviser. Expert performance is non-contractible in a host of settings, particularly when the DM and adviser have no direct relationship (contractual or otherwise), such as with politicians, supreme court justices, lobbyists, or scientists trying to influence public opinion. Contingent transfers may also be infeasible for institutional reasons, e.g. the use of incentive pay is limited in government agencies. To focus on incentive issues, we also postpone participation constraints until Section 7.

Timing. The sequence of events is as follows. First, the adviser's type μ_A is exogenously given (or, later in the paper, chosen by the DM) from an available set of adviser types, $[\underline{\mu}, \overline{\mu}]$, where $\underline{\mu} < 0 < \overline{\mu}$. The adviser then chooses effort and observes the outcome of his investigation, both unobservable to the DM. In the third stage, the adviser either discloses or withholds any information acquired. Finally, the DM takes an action.

As this is a multi-stage Bayesian game, our solution concept is Perfect Bayesian Equilibrium, or for short, *equilibrium* hereafter.[6] We restrict attention to pure strategy equilibria.

[5] See, for example, Shin (1998). The "unraveling" phenomenon that occurs in some models of hard information will not arise here because the adviser does not always have a signal (cf. Shin (1994)).

[6] We acknowledge that learning justifications for equilibrium are more difficult than usual when players have heterogenous priors (Dekel, Fudenberg, and Levine, 2004). Many of our main points remain valid when the adviser's signal is publicly observed, in which case iterated elimination of dominated strategies would suffice. In addition, the equilibrium analysis of strategic disclosure and its implication for information acquisition applies just as well when the conflict is one of preferences rather than opinions, as discussed in Section 5.

2.1 Interim Bias

As a prelude to our analysis, it is useful to identify the players' preferences over decisions when the state is not known. Under the Normality assumptions in our information structure, the signal and state joint distribution can be written, from the perspective of player $i = DM, A$, as

$$\begin{pmatrix} \omega \\ s \end{pmatrix} \sim N\left(\begin{matrix} \mu_i \\ \mu_i \end{matrix}, \begin{pmatrix} \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 + \sigma_1^2 \end{pmatrix} \right).$$

Without a signal about the state, the expected utility of player i is maximized by action μ_i. Suppose a signal s is observed. The posterior of player i is that $\omega|s \sim N\left(\rho s + (1-\rho)\mu_i, \tilde{\sigma}^2\right)$, where $\rho := \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}$ and $\tilde{\sigma}^2 := \frac{\sigma_0^2\sigma_1^2}{\sigma_0^2+\sigma_1^2}$ (Degroot, 1970).[7] Player $i = DM, A$ therefore has the following expected utility from action a given signal s:

$$\begin{aligned} \mathbb{E}[u_i(a,\omega)|s,\mu_i] &= -\mathbb{E}[(a-\omega)^2|s,\mu_i] = -(a-\mathbb{E}[\omega|s,\mu_i])^2 - \text{Var}(\omega|s) \\ &= -\left[a - (\rho s + (1-\rho)\mu_i)\right]^2 - \tilde{\sigma}^2. \end{aligned} \tag{1}$$

Clearly, the expected utility for player i is maximized by an action $\alpha(s|\mu_i) := \rho s + (1-\rho)\mu_i$, where $\alpha(s|\mu)$ is simply the posterior mean for a player with type μ.

Equation (1) shows that so long as signals are not perfectly informative of the state ($\rho < 1$), differences of opinion generate conflicts in preferred decisions given any signal, even though fundamental preferences agree. Accordingly, we define the *interim bias* as $B(\mu) := (1-\rho)\mu$. This completely captures the difference in the two players' preferences over actions given any signal because $\alpha(s|\mu) = \alpha(s|0) + B(\mu)$. Observe that for any $\mu \neq 0$, $sign(B(\mu)) = sign(\mu)$ but $|B(\mu)| < |\mu|$. Hence, while interim bias persists in the same direction as prior bias, it is of strictly smaller magnitude because information about the state mitigates prior disagreement about the optimal decision. This simple observation turns out to have significant consequences. The magnitude of interim bias depends upon how precise the signal is relative to the prior; differences of opinion matter very little once a signal is acquired if the signal is sufficiently precise, i.e. for any μ, $B(\mu) \to 0$ as $\rho \to 1$ (equivalently, as $\sigma_1^2 \to 0$ or $\sigma_0^2 \to \infty$).

Hereafter, since we have normalized $\mu_{DM} = 0$, we will refer to the adviser's type as just μ rather than μ_A to reduce notation.

[7] Since $\sigma_0^2 > 0$ and $\sigma_1^2 > 0$, $\rho \in (0,1)$. However, it will be convenient at points to discuss the case of $\rho = 1$; this should be thought of as the limiting case where $\sigma_1^2 = 0$, so that signals are perfectly informative about the state. Similarly for $\rho = 0$.

3 Equilibrium Disclosure Behavior

This section analyzes the outcome of strategic communication in the disclosure sub-game.[8] For this purpose, it will be sufficient to focus on the interim bias of the adviser, $B(\mu)$, and the DM's belief about the probability p that the adviser observes a signal.[9] Hence, we take the pair (B, p) as a primitive parameter in this section. Our objective is to characterize the set $S \in \mathbb{R}$ of signals that the adviser withholds and the action $a_\emptyset$ the DM chooses when there is no disclosure. Plainly, when s is disclosed, the DM will simply choose her most-preferred action, $\alpha(s|0) = \rho s$.

We start by fixing an arbitrary action $a \in \mathbb{R}$ the DM may choose in the event of nondisclosure, and ask whether the adviser will disclose his signal if he observes it, assuming that $B \geq 0$ (the logic is symmetric when $B < 0$). The answer can be obtained easily with the aid of Figure 1 below. The figure depicts, as a function of the signal, the action most preferred by the DM (ρs) and the action most preferred by the adviser ($\rho s + B$): each is a straight line, the latter shifted up from the former by the constant B. Since the DM will choose the action ρs whenever s is disclosed, the adviser will withhold s whenever the nondisclosure action a is closer to his most-preferred action, $\rho s + B$, than the disclosure action, ρs. This reasoning identifies the nondisclosure interval as the "flat" region of the solid line, which corresponds to the nondisclosure action chosen by the DM.

[8]Strictly speaking, we are abusing terminology in referring to this as a "sub-game," because the DM does not observe the adviser's effort choice, p.

[9]The subsequent analysis will show why it is the DM's belief about the adviser's effort, rather than his actual effort, that matters for disclosure behavior. We will require this belief to be correct when we analyze the information acquisition stage.

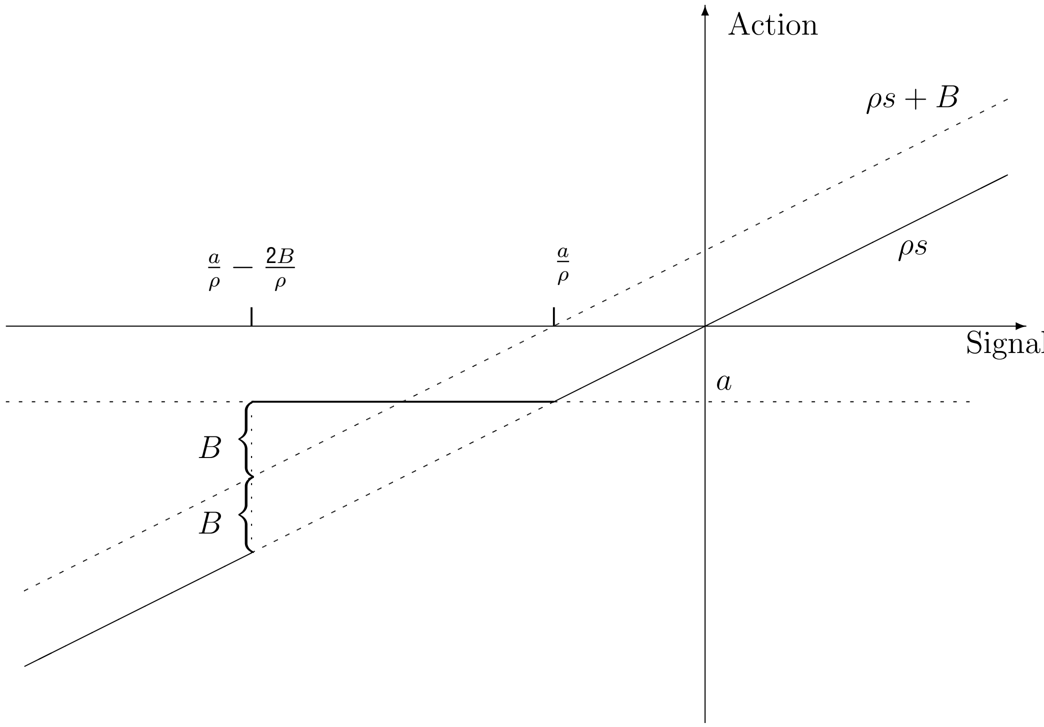

Figure 1: Optimal non-disclosure region

As seen in Figure 1, the adviser's best response is to withhold s (in case he observes s) if and only if $s \in R(B,a) := [l\left(B,a\right), h\left(a\right)]$, where

$$h\left(a\right) := \frac{a}{\rho}, \tag{2}$$

$$l(B,a) := h\left(a\right) - \frac{2B}{\rho}. \tag{3}$$

At $s = h(a)$, the DM will choose $a = \alpha\left(h(a)|0\right)$ whether s is disclosed or not, so the adviser is indifferent. At $s = l(B,a)$, the adviser is again indifferent between disclosure (which leads to $\alpha\left(l(B,a)|0\right) = a - 2B$) and nondisclosure (which leads to a) because they are equally distant from his most preferred action, $a - B$. For any $s \notin [l(B,a), h(a)]$, disclosure will lead to an action closer to the adviser's preferred action than would nondisclosure.[10]

Next, we characterize the DM's best response in terms of her nondisclosure action, for an arbitrary (measurable) set $S \subseteq \mathbb{R}$ of signals that the adviser may withhold. Her best response is to take the action that is equal to her posterior expectation of the state given

[10]We assume nondisclosure when indifferent, but this is immaterial.

nondisclosure, which is computed via Bayes rule:

$$a_N(p,S) = \frac{p\rho \int_S s\gamma\,(s;0)\,ds}{p\int_S \gamma\,(s;0)\,ds + 1 - p}, \tag{4}$$

where $\gamma(s;\mu)$ is a Normal density with mean μ and variance $\sigma_0^2 + \sigma_1^2$. Notice that the DM uses her own prior $\mu_{DM} = 0$ to update her belief. It is immediate that if S has zero expected value, then $a_N(p,S) = 0$. More importantly, for any $p > 0$, $a_N(p,S)$ increases when S gets larger in the strong set order.[11] Intuitively, the DM rationally raises her action when she suspects the adviser of not disclosing larger values of s.

An equilibrium of the disclosure sub-game requires that both the DM and the adviser must play best responses. This translates into a simple fixed point requirement:

$$S = R\,(B,a) \text{ and } a_N(p,S) = a. \tag{5}$$

Given any (B,p), let $(S(B,p), a_\emptyset(B,p))$ be a pair that satisfies (5), and let $\underline{s}(B,p)$ and $\overline{s}(B,p)$ respectively denote the smallest and the largest elements of $S(B,p)$. The following result ensures that these objects are uniquely defined; its proof, and all subsequent proofs not in the text, are in the Appendix.

Proposition 1. *For any (B,p), there is a unique equilibrium in the disclosure sub-game. In equilibrium, both $\underline{s}(B,p)$ and $\overline{s}(B,p)$ are equal to zero if $B = 0$, are strictly decreasing in B when $p > 0$, and strictly decreasing (increasing) in p if $B > 0$ (if $B < 0$). The nondisclosure action $a_\emptyset(B,p)$ is zero if $B = 0$ or $p = 0$, is strictly decreasing in B for $p > 0$, and is strictly decreasing (increasing) in p if $B > 0$ (if $B < 0$).*

It is straightforward that the adviser reveals his information fully to the DM if and only if $B = 0$, i.e. there is no interim bias. To see the effect of an increase in B (when $p > 0$), notice from (2) and (3) that if the DM's nondisclosure action did not change, the upper endpoint of the adviser's nondisclosure region would not change, but he would withhold more low signals with a higher B. Consequently, by (4), the DM must adjust his nondisclosure action downward, which has the effect of pushing down both endpoints of the adviser's nondisclosure region. The new fixed point must therefore feature a smaller nondisclosure set (in the sense of strong set order) and a lower nondisclosure action from

[11] A set S is larger than S' in the strong set order if for any $s \in S$ and $s' \in S'$, $\max\{s,s'\} \in S$ and $\min\{s,s'\} \in S'$.

the DM. We call this the *prejudicial effect*, since a more upwardly biased adviser is in essence punished with a lower inference when he claims not to have observed a signal. The prejudicial effect implies in particular that for any $p > 0$ and $B \neq 0$, $a_\emptyset(B, p)B < 0$.

The impact of p can be traced similarly. An increase in p makes it more likely that nondisclosure from the adviser is due to withholding of information rather than a lack of signal. If $B > 0$ (resp. $B < 0$), this makes the DM put higher probability on the signal being low (resp. high), leading to a decrease (resp. increase) in the nondisclosure action, which decreases (resp. increases) the nondisclosure set in the strong set order.

4 Opinions as Incentives

This section studies how the adviser's opinion affects his incentive to acquire information, and the implications this has on the optimal type of adviser for the DM. As a benchmark, the following Proposition establishes the fairly obvious point that, absent information acquisition concerns, the optimal adviser is a like-minded one.

PROPOSITION 2. *If the probability of acquiring a signal is held fixed at some $p > 0$, the uniquely optimal type of adviser for the DM is like-minded, i.e. an adviser with $\mu = 0$.*

PROOF. For any $p > 0$, $S(\mu, p)$ has positive measure when $\mu > 0$, whereas $S(0, p)$ has measure zero. Hence, the adviser $\mu = 0$ reveals the signal whenever she obtains one, whereas an adviser with $\mu \neq 0$ withholds the signal with positive probability. The result follows from the fact that the DM is strictly better off under full disclosure than partial disclosure. ▌

We now turn to the the case where information acquisition is endogenous. To begin, suppose the DM believes that an adviser with type $\mu \geq 0$ will choose effort p^e. The following Lemma decomposes the payoff for the adviser from choosing effort p, denoted $U_A(p; p^e, B, \mu)$, in a useful manner.[12]

LEMMA 1. *The adviser's expected utility from choosing effort p can be written as*

$$U_A(p; p^e, B, \mu) = K(B, \mu, p^e) + p\Delta(B, \mu, p^e) - c(p),$$

[12] Even though the interim bias B is determined by μ, we write them as separate variables in the funcion $U_A(\cdot)$ to emphasize the two separate effects caused by changes in the difference of opinion: changes in prior beliefs over signal distributions and changes in the interim bias.

where

$$K(B,\mu,p^e) := -\int \left(a_\emptyset\left(B,p^e\right)-(\rho s+B)\right)^2\gamma\left(s;\mu\right)ds-\tilde{\sigma}^2 \tag{6}$$

and

$$\Delta(B,\mu,p^e) := \int_{s\notin S(B,p^e)}\left[\left(a_\emptyset\left(B,p^e\right)-(\rho s+B)\right)^2-B^2\right]\gamma\left(s;\mu\right)ds. \tag{7}$$

The first term in the decomposition, $K(\cdot)$, is the expected utility when a signal is not observed. Equation (6) expresses this utility by iterating expectations over each possible value of s, reflecting the fact that the DM takes action $a_\emptyset(\cdot)$ without its disclosure whereas the adviser's preferred action if the signal were s is $\rho s + B$, and that $\tilde{\sigma}^2$ is the residual variance of the state given any signal. The second term in the decomposition, $p\Delta(\cdot)$, is the probability of obtaining a signal multiplied by the expected gain from obtaining a signal. Equation (7) expresses the expected gain, $\Delta(\cdot)$, via iterated expectations over possible signals. To understand it, note that the adviser's gain is zero if a signal is not disclosed (whenever $s \in S(B,p^e)$), whereas when a signal is disclosed, the adviser's utility (gross of the residual variance) is $-B^2$, because the DM takes action ρs.

We are now in a position to characterize the adviser's equilibrium effort level. Given the DM's belief, p^e, the adviser will choose p to maximize $U_A(p;p^e,B,\mu)$. By the Inada conditions on the effort cost, this choice is in the interior of $[0,\bar{p}]$ and characterized by the first-order condition:

$$\frac{\partial U_A(p;p^e,B,\mu)}{\partial p} = \Delta(B,\mu,p^e) - c'(p) = 0.$$

Equilibrium requires that the DM's belief be correct, i.e. $p^e = p$. Therefore, in equilibrium, we must have

$$\Delta(B,\mu,p) = c'(p). \tag{8}$$

LEMMA 2. *For any (B,μ), there is a solution to (8), and p is an equilibrium effort choice if and only if $p \in (0,\bar{p})$ and satisfies (8).*

In general, we cannot rule out there being multiple equilibrium effort levels for a given type of adviser. The reason is that the DM's action in the event of nondisclosure depends on the adviser's (expected) effort, and the adviser's equilibrium effort in turn depends on the DM's action upon nondisclosure.[13] For the remainder of the paper, for each (B, μ), we focus on the highest equilibrium effort. Since the interim bias B is uniquely determined by $B(\mu) = (1 - \rho)\mu$, we can define the equilibrium probability of information acquisition as a function solely of μ, which we denote by $p(\mu)$. Our first main result is:

PROPOSITION 3. *An adviser with a greater difference of opinion acquires information with higher probability: $p(\mu') > p(\mu)$ if $|\mu'| > |\mu|$.*

To see the intuition, first ignore the strategic disclosure of information, assuming instead that the outcome of the adviser's investigation is publicly observed. In this case, there is no prejudice associated with nondisclosure, so the DM will choose $a_{\emptyset}(B, p) = 0$ independent of B or p. It follows from a mean-variance decomposition that $-\sigma_0^2 - \mu^2$ is the expected utility for the adviser conditional on no signal, and $-\tilde{\sigma}^2 - (B(\mu))^2$ is the expected utility conditional on getting a signal. Hence, the adviser's marginal benefit of acquiring a signal, denoted $\Delta^{pub}(\mu)$, is given by[14]

$$\Delta^{pub}(\mu) = \underbrace{\sigma_0^2 - \tilde{\sigma}^2}_{\text{uncertainty reduction}} + \underbrace{\mu^2 - (B(\mu))^2}_{\text{persuasion}}. \tag{9}$$

Acquiring information benefits the adviser by reducing uncertainty about the true state, as shown by the first part of (9). But in addition, the adviser expects to *persuade* the DM: without information, the adviser views the DM's decision as biased by μ, their ex-ante disagreement in beliefs; whereas with information, the disagreement is reduced to the interim bias, $B(\mu) = (1 - \rho)\mu$. Since $\mu^2 - (B(\mu))^2$ is strictly increasing in $|\mu|$, the persuasion incentive is strictly larger for an adviser with a greater difference of opinion. Hence, such an adviser exerts more effort towards information acquisition. Equivalently, the adviser expects action $\rho\mu$ to be taken with information, but action 0 to be taken without information. Hence, the adviser believes that by acquiring information, he can persuade the DM to take an action that is closer in expectation to his own prior.[15] The benefit of such persuasion is more valuable to an adviser with greater difference of opinion.

[13]Formally, multiplicity emerges when the function $\Delta(B, \mu, \cdot)$ crosses more than once with the strictly increasing function $c'(\cdot)$ over the domain $[0, 1]$. As we will discuss more shortly, if signals are public rather than privately observed by the adviser, there is a unique equilibrium because $\Delta(B, \mu, \cdot)$ is constant. Moreover, we show in the Appendix (in the proof of Proposition 3) that for all μ sufficiently close to 0, there is a unique equilibrium effort level.

[14]Alternatively, one can also verify that equation (7) simplifies to equation (9) if the nondisclosure region $S(\cdot) = \emptyset$ and the nondisclosure action $a_{\emptyset}(\cdot) = 0$, as is effectively the case under public observation of signal.

[15]Of course, the DM does not expect to be persuaded: her expectation of her action conditional on a signal being acquired is 0. Instead, she expects that a signal will cause the adviser's preferred decision to

Now consider the case where information is private, and the adviser strategically communicates. Suppose the DM expects effort p^e from the adviser of type μ. Then she will choose $a_\emptyset(B(\mu), p^e)$ when a signal is not disclosed. Since the adviser always has the option to disclose all signals, his marginal benefit of acquiring information and then strategically disclosing it, as defined by equation (7), is at least as large as the marginal benefit from (sub-optimally) disclosing all signals, which we shall denote $\Delta^{pri}(\mu, a_\emptyset)$ (with the arguments of $a_\emptyset$ suppressed). By mean-variance decomposition again, we have

$$\begin{aligned}\Delta(B(\mu), \mu, p^e) &\geq \Delta^{pri}(\mu, a_\emptyset) \\ &= \underbrace{\sigma_0^2 - \tilde{\sigma}^2}_{\text{uncertainty reduction}} + \underbrace{\mu^2 - (B(\mu))^2}_{\text{persuasion}} + \underbrace{(a_\emptyset)^2 - 2a_\emptyset\mu}_{\text{avoiding prejudice}}.\end{aligned} \tag{10}$$

Recall from Proposition 1 the prejudicial effect: for any $p^e > 0$ and $\mu \neq 0$, $a_\emptyset\mu < 0$. Hence, for any $p^e > 0$ and $\mu \neq 0$, $\Delta^{pri}(\mu, p^e) > \Delta^{pub}(\mu)$: given that information is private, the DM's rational response to the adviser claiming a lack of information affects the adviser adversely—this is the prejudicial effect—and to avoid such an adverse inference, the adviser is even more motivated to acquire a signal than when information is public.

Propositions 1 and 3 identify the tradeoff faced by the DM: an adviser with a greater difference of opinion exerts more effort, but reveals less of any information he may acquire. Does the benefit from improved incentives for information acquisition outweigh the loss from strategic disclosure? We demonstrate below that this is indeed the case for at least some difference in opinion.

Proposition 4. *There exists some $\mu \neq 0$ such that if the DM can choose her adviser, it is strictly better to appoint an adviser of type μ than a like-minded adviser.*

To prove the proposition, we establish that locally around $\mu = 0$, difference of opinion strictly benefits the DM. This is largely due to the persuasion effect. As the difference of opinion μ is raised slightly from zero, the prejudicial effect (which entails both communication loss and information acquisition gain) is negligible, whereas the persuasion motive and the benefit it generates in increased information acquisition is non-negligible by comparison.[16] This can be seen most clearly when the signal is perfectly informative, $\rho = 1$.

shift towards her opinion. This feature that each player expects new information to persuade the other is also central to Yildiz's (2004) analysis of bargaining with heterogeneous priors.

[16] We say "by comparison" because both the adviser's equilibrium effort and the DM's equilibrium non-disclosure action have derivatives of zero with respect to μ at $\mu = 0$, as is established in the proof. Thus, the order of magnitude comparisons are with regard to the second derivatives.

In this case, $B(\mu) = 0$, so there is full disclosure in the communication stage, analogous to a situation where information is public. Hence, by Proposition 3, any adviser with a difference of opinion is preferred to a like-minded adviser. By continuity, for any $\mu \neq 0$, there is a set of ρ's near 1 for which an adviser of type μ is better for the DM than a like-minded adviser. This argument verifies Proposition 4 for all ρ sufficiently close to 1. The proof in the Appendix shows that for any ρ, however far from 1, all adviser types sufficiently close to type 0 are in fact better for the DM than a like-minded adviser.

Remark 1. *The conclusion of Proposition 4 does not depend on selecting the equilibrium with highest effort for a given adviser type. The proof of the Proposition establishes that for all μ sufficiently close to 0, there is a unique equilibrium effort level.*

5 Opinions and Preferences

We have thus far assumed that the DM and the available pool of advisers all have the same fundamental preferences, but differ in opinions. In this section, we augment the space of types to allow for fundamental preference conflicts. This allows us to explore a number of issues, such as: Will the DM benefit from an adviser with different preferences in the same way she will benefit from one with a different opinion? If an adviser can be chosen from a very rich pool of advisers differing both in opinions and preferences, how will the DM combine the two attributes? For instance, for an adviser with a given preference, will she prefer him to be a *skeptic*—one who doubts that discovering information will shift the DM's action in the direction of his preference bias—or a *zealot*—one who believes that his preference will also be "vindicated by the evidence."

To keep matters simple, suppose, as is standard in the literature, that a player's preferences are indexed by a single bias parameter $b \in [\underline{b}, \overline{b}]$, with $\underline{b} < 0 < \overline{b}$, such that his state-dependent von Neumann-Morgenstern utility is $u(a, \omega, b) = -(a - \omega - b)^2$. The adviser therefore now has a two-dimensional type (that is common knowledge), (b, μ). The DM's type is normalized as $(0, 0)$.

Interim but not ex-ante equivalence. Similar to earlier analysis, it is straightforward that an adviser of type (b, μ) desires the action $\alpha(s|b, \mu) := \rho s + (1 - \rho)\mu + b$ when signal s is observed. Hence, such an adviser has an interim bias of $B(b, \mu) := (1 - \rho)\mu + b$. This immediately suggests the interchangeability of the two kinds of biases—preferences and opinions—in the disclosure sub-game. For any adviser with opinion bias μ and no preference bias, there exists a like-minded adviser with only preference bias $b = (1-\rho)\mu$ such that the latter will have precisely the same incentives to disclose the signal as the former. Formally, given the same effort level, the disclosure sub-game equilibrium played by the DM and either adviser is the same.

This isomorphism does not extend to the information acquisition stage. To see this, start with an adviser of type $(0, \mu)$, i.e., with opinion bias μ but no preference bias. When such an adviser does not acquire a signal, he expects the DM to make a decision that is distorted by at least μ from what he regards as the right decision.[17] Consider now an adviser of type $(\mu, 0)$, i.e., with preference bias $b = \mu$ and no opinion bias. This adviser also believes that, absent disclosure of a signal, the DM will choose an action that is at least μ away from his most preferred decision. Crucially, however, their expected payoffs from disclosing a signal are quite different. The former type (opinion-biased adviser) believes that the signal will vindicate his prior and thus bring the DM closer toward his ex-ante preferred decision; whereas the latter type (preference-biased adviser) has no such expectation. One concludes that *the persuasion motive does not exist for an adviser biased in preferences alone.*

Publicly observed signal. To see how the two types of biases can interact in affecting the incentive for information acquisition, it is useful to first consider the case where the adviser's signal (or lack thereof) is publicly observed. This makes the analysis straightforward because there is no strategic withholding of information. Fix any adviser of type (b, μ). If no signal is observed, the DM takes action 0, while the the adviser prefers the action $b + \mu$. Hence, the adviser has expected utility $-\sigma_0^2 - (b+\mu)^2$. If signal s is observed, then the DM takes action ρs; since the adviser prefers action $\rho s + B(b, \mu)$, he has expected utility $-\tilde{\sigma}^2 - (B(b, \mu))^2$. Therefore, the adviser's expected gain from acquiring information is

$$\Delta^{pub}(b, \mu) = \underbrace{\sigma_0^2 - \tilde{\sigma}^2}_{\text{uncertainty reduction}} + \underbrace{\left(2\rho - \rho^2\right)\mu^2}_{\text{persuasion}} + \underbrace{2\rho b \mu.}_{\text{reinforcement}} \tag{11}$$

Suppose first $\mu = 0$, so the adviser is like-minded. In this case, $\Delta^{pub}(b, 0)$ is independent of b. That is, the incentive for a like-minded adviser to acquire information does not depend on his preference, and consequently, there is no benefit from appointing an adviser who differs only in preference. This stands in stark contrast to the case of difference of opinion,

[17] "At least," because the prejudicial effect will cause the DM to take an action even lower than 0, unless information is public or signals are perfectly-informative.

$(0, \mu)$, $\mu \neq 0$, where equation (9) showed that advisers with greater difference of opinion have bigger marginal benefits of acquiring information, and are therefore strictly better for the DM under public information. This clearly shows the distinction between preferences and opinions.

Now suppose $\mu \neq 0$. Then, the persuasion effect reappears, as is captured by the second term of (11). More interestingly, the adviser's preference also matters now, and in fact interacts with the opinion bias. Specifically, *a positive opinion bias is reinforced by a positive preference bias, whereas it is counteracted by a negative preference bias*; this effect appears in (11). The intuition turns on the concavity of the adviser's payoff function, and can be seen as follows. Without a signal, the adviser's optimal action is away from the DM's action by $|b + \mu|$. Concavity implies that the bigger is $|b + \mu|$, the greater the utility gain for the adviser when he expects to move the DM's action in the direction of his ex-ante bias. Therefore, when $\mu > 0$, say, an adviser with $b > 0$ has a greater incentive to acquire information than an adviser with $b < 0$. In fact, if b were sufficiently negative relative to $\mu > 0$, the adviser may not want to acquire information at all, because he expects it to shift the DM's decision *away* from his net bias of $b + \mu$.

Privately observed signal. When the signal is observed privately by the adviser, the prejudicial motive is added to his incentive for information acquisition. The next proposition states an incentive effect of both preference and opinion biases. Extending our previous notation, we use $p(B, \mu)$ to denote the highest equilibrium effort choice of an adviser with interim bias B and prior μ.

PROPOSITION 5. *Suppose* $(|B(b, \mu)|, |\mu|) < (|B(b', \mu')|, |\mu'|)$ *and* $B(b', \mu')\mu' \geq 0$.[18] *Then,* $p(B(b', \mu'), \mu') > p(B(b, \mu), \mu)$.

Proposition 5 nests Proposition 3 as a special case with $b = b' = 0$. Setting $\mu = \mu' = 0$ gives the other special case in which the adviser differs from the DM only in preference. Unlike under public information, a preference bias alone creates incentives for information acquisition when the outcome of the adviser's experiment is private. The reason is that an adviser exerts additional effort to avoid the prejudicial inference the DM attaches to nondisclosure. Of course, from the DM's point of view, this incentive benefit is offset by the loss associated with strategic withholding of information. It turns out that these opposing effects are of the same magnitude locally when $|b| \approx 0$. Hence, in net, a small difference of

[18] We follow the convention that $(x, y) < (x', y')$ if $x \leq x'$ and $y \leq y'$, with at least one strict inequality.

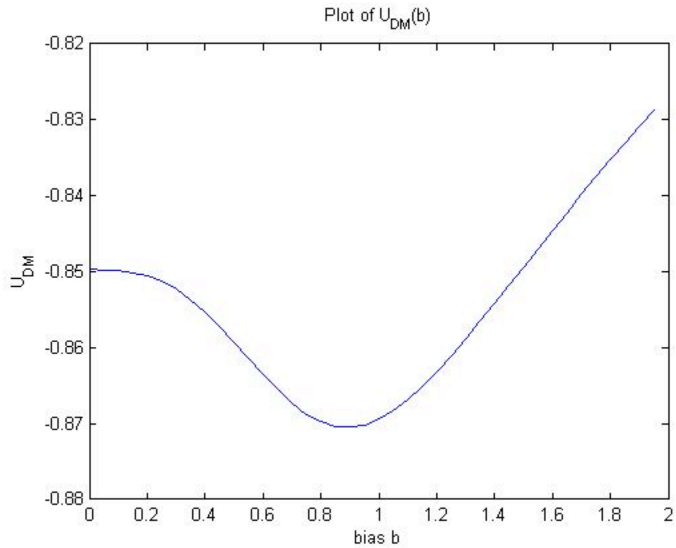

Figure 2: DM's utility as a function of adviser's preference. Parameters: $c(p) = \frac{p^2}{1-p}, \sigma_1^2 = 1$, $\sigma_0^2 = 0.5$.

preference is not unambiguously beneficial to the DM in the way that difference of opinion is. Indeed, a numerical example shows that the DM's utility is decreasing in $|b|$ around $b = 0$, but interestingly, starts increasing when $|b|$ becomes sufficiently large, to the point where it can rise above the utility associated with type $b = 0$. This is shown in Figure 2. In such cases, the DM never prefers an adviser with preference bias unless the bias is sufficiently large, contrasting with difference of opinion. This difference may matter if the space of available adviser types is not sufficiently large (such as $\bar{b} < 1.4$ in the example plotted in Figure 2).

More generally, Proposition 5 reveals how the two types of biases interact with respect to the incentive for information acquisition, yielding some useful corollaries.

Corollary 1. *If $(b', \mu') > (b, \mu) \geq 0$, then an adviser with (b', μ') chooses a higher effort than one with (b, μ).*

Thus, in the domain $(b, \mu) \in \mathbb{R}_+^2$, an increase in either kind of bias—preference or opinion—leads to greater information acquisition.

Corollary 2. *Suppose an adviser has type (b, μ) such that $B(b, \mu) \geq 0$ but that $\mu < 0$. Replacing the adviser with one of type $(b, -\mu)$ leads to a higher effort.*

An adviser of type (b, μ) with $B(b, \mu) \geq 0$ but $\mu < 0$ likes actions higher than the DM would like if the state of the world were publicly known, yet he is a priori pessimistic about obtaining a signal that will shift the DM's action upward. In this sense, he is a *skeptic*, and does not have a strong incentive for information acquisition. Replacing him with a *zealot* who believes that information about the state will in fact lead the DM to take a higher action leads to more information acquisition.

The final corollary shows that having access to a rich pool of advisers on both opinion and preference dimensions endows the DM with enough degree of freedom to eliminate disclosure loss altogether, and yet use the adviser's type as an incentive instrument.

COROLLARY 3. *If $B(b, \mu) = B(b', \mu') \geq 0$ and $\mu' > \mu \geq 0$, then the adviser with (b', μ') chooses a higher effort than the one with (b, μ). Moreover, the DM strictly prefers appointing the former. In particular, if one raises μ and lowers b so as to maintain $B(b, \mu) = 0$, then an higher effort is induced while maintaining full disclosure.*

Choosing an adviser who has opinion $\mu > 0$ but negative preference bias $b = -(1-\rho)\mu$ eliminates interim bias altogether, and thus avoids any strategic withholding of information. If this can be done without any constraints, the DM can raise μ unboundedly and increase his expected utility. However, since the space of available types is likely bounded in practice (as we have assumed), it can be optimal for the DM to choose an expert with an interim bias $B(b, \mu) \neq 0$, analogous to Proposition 4 when available advisers are differentiated by opinions alone.

6 The Allocation of Authority: Initiative and Complacency

An important issue in organizations is the allocation of formal authority, or the choice between delegation and centralization.[19] One prominent view is that delegation of decision-making authority to an agent increases his incentives to become informed because he can use this information to maximize his own interests, but delegation is costly to the DM insofar as the agent's decisions do not maximize the DM's interests. Aghion and Tirole (1997) label this the *initiative* versus *loss of control* tradeoff. Our framework delivers a novel and, we believe, important additional consideration that mitigates the initiative effect. To permit a better comparison with Aghion and Tirole, we extend our model in this section to allow both parties to acquire information.

[19] In keeping with the incomplete contracts approach, we only consider full delegation and no delegation, leaving intermediate possibilities to future work.

Simultaneous Public Information Acquisition. Return to the baseline setting where individuals only differ in their opinions. Let us now think of the DM (of type 0) as a principal who has formal authority over decisions at the outset. She may choose to retain this authority, which we refer to as centralization, or delegate authority over decisions to the adviser or agent of type $\mu \neq 0$, which we refer to as delegation. As in Aghion and Tirole (1997), regardless of the allocation of authority, both individuals can exert costly effort in an attempt to observe the signal, s, about the state. If either party's information acquisition is successful, the signal is publicly observed. Formally, the adviser chooses a probability $p \in [0, \bar{p}]$ of obtaining the signal s at cost $c(p)$, and the DM chooses the probability $P \in [0, \bar{p}]$ of obtaining the signal at cost $C(P)$, where $C(\cdot)$ satisfies the same conditions as we have assumed on $c(\cdot)$.[20] Anticipating the analysis, we also assume

$$C'(P)c'(p) < (1-P)(1-p)C''(P)c''(p) \text{ for all } (P, p) \in [0, \bar{p}]^2, \tag{12}$$

in order to guarantee uniqueness of equilibrium.

The timing is as follows: under centralization, both parties simultaneously exert effort, the signal is either publicly observed or not with probability determined by the efforts, and then the DM makes a decision. A delegation regime proceeds analogously, except that at the last stage, it is the adviser who makes the decision.

There are a few noteworthy differences between the current setting and that of Aghion and Tirole (1997). First, because we are assuming public information, there is no distinction here between "formal authority" and "real authority": whoever has formal authority has real authority. We return to this point later when we consider private information. Second, the conflict of interest here between the two parties stems from a difference of opinion rather than fundamentally different preferences. As already seen, this distinction matters for information acquisition incentives, and will turn out have implications for the allocation of authority. Third, as the two individuals have different ex-ante preferred decisions, the allocation of authority will affect the outcome here even when no information is acquired. By contrast, Aghion and Tirole assume that in the absence of any information, both principal and agent agree on the preferred decision ("no project"), implying that the allocation of authority is irrelevant in this event. While their assumption may be reasonable for some applications, it is also the case that in many situations (formal) authority matters even when—sometimes, especially when—there is no new information. This issue

[20] That is, $C(\cdot)$ is smooth, $C''(\cdot) > 0$, $C'(0) = 0$ and $C'(p) \to \infty$ as $p \to \bar{p}$.

will turn out to be crucial.

Under centralization, the DM makes her preferred decision $\alpha(s|0)$ when s is observed, or else she chooses decision 0. Hence, given effort choices P by the DM and p by the adviser, the expected payoffs for each player are respectively:

$$\mathcal{U}_{DM}^{C}(P,p) = PU_{0}(0) + (1-P)pU_{0}(0) + (1-P)(1-p)\bar{U}_{0}(0) - C(P), \tag{13}$$

$$\mathcal{U}_{A}^{C}(P,p) = PU_{\mu}(0) + (1-P)pU_{\mu}(0) + (1-P)(1-p)\bar{U}_{\mu}(0) - c(p),$$

where $U_{\mu'}(\mu'')$ (resp. $\bar{U}_{\mu'}(\mu'')$) denotes the expected payoff of a player with prior μ' when decisions are made by a player with prior μ'' and a signal is observed (resp. not observed).[21] These expressions account for both the ex-ante and ex-interim conflict of interest stemming from difference of opinion, and that from either player's perspective, all that matters is whether a signal is observed or not, not who observes it.

The first-order conditions for payoff maximization,

$$\begin{aligned} C'(P) &= (1-p)(U_{0}(0) - \bar{U}_{0}(0)), \\ c'(p) &= (1-P)(U_{\mu}(0) - \bar{U}_{\mu}(0)), \end{aligned} \tag{14}$$

imply best response functions for the DM and adviser, denoted $\Phi^{C}(p)$ and $\phi^{C}(P)$ respectively, that are negatively sloped since the cost functions are convex. Thus, effort choices are strategic substitutes. Condition (12) ensures that there is a unique equilibrium, $(\hat{P}^{C}, \hat{p}^{C})$, which satisfies $\hat{P}^{C} = \Phi^{C}(\hat{p}^{C})$ and $\hat{p}^{C} = \phi^{C}(\hat{P}^{C})$.

The characterization under delegation proceeds similarly. In this case, the adviser makes his preferred decision $\alpha(s|\mu)$ when s is observed, and chooses μ otherwise. Expected payoffs are thus given by

$$\mathcal{U}_{DM}^{D}(P,p) = PU_{0}(\mu) + (1-P)pU_{0}(\mu) + (1-P)(1-p)\bar{U}_{0}(\mu) - C(P), \tag{15}$$

$$U_{A}^{D}(P,p) = PU_{\mu}(\mu) + (1-P)pU_{\mu}(\mu) + (1-P)(1-p)\bar{U}_{\mu}(\mu) - c(p).$$

[21] The reader may verify that $U_{\mu'}(\mu'') = -\tilde{\sigma}^{2} - (B(\mu') - B(\mu''))^{2}$ and $\bar{U}_{\mu'}(\mu'') = -\sigma_{0}^{2} - (\mu' - \mu'')^{2}$.

The associated first-order conditions are

$$\begin{aligned} C'(P) &= (1-p)(U_0(\mu) - \bar{U}_0(\mu)), \\ c'(p) &= (1-P)(U_\mu(\mu) - \bar{U}_\mu(\mu)). \end{aligned} \tag{16}$$

These define negatively-sloped best response functions for the DM and adviser, $P = \Phi^D(p)$ and $p = \phi^D(P)$ respectively, and there exists a unique equilibrium $(\hat{P}^D, \hat{p}^D)$ which satisfies $\hat{P}^D = \Phi^D(\hat{p}^D)$ and $\hat{p}^D = \phi^D(\hat{P}^D)$.

We are now in a position to see that the two effects of delegation identified by Aghion and Tirole—loss of control and increase in the agent's initiative—appear here as well. But in addition, there is a novel third effect, *decrease in initiative from complacency*. Let us discuss each one:

- **Loss of control:** Observe that $U_0(\mu) < U_0(0)$. In other words, when a signal is observed, the DM is worse off when decisions are made by interpreting a signal according to the adviser's opinion rather than her own. This is what happens under delegation when there is successful information acquisition. Note that since real authority and formal authority are equivalent here, it is irrelevant who acquires the signal. Moreover, as $\bar{U}_0(\mu) < \bar{U}_0(0)$, loss of control also hurts the DM even absent a signal. Altogether, equations (13) and (15) show that if (P, p) is held fixed, the DM is worse off under delegation.

- **Increased initiative from controlling the use of information:** The ability to control how information should be interpreted, without being overruled by the DM, means that one effect of delegation is to increase the incentive for the adviser to acquire information. Formally, notice from (16) and (14) that because $U_\mu(\mu) > U_\mu(0)$, the first term in the marginal benefit of effort for the adviser goes up under delegation because his expected utility when a signal is observed is higher when he has the power to make decisions.

- **Decreased initiative from complacency:** On the other hand, we also see that under delegation, the adviser suffers less from the lack of information, since he can take his ex-ante preferred action rather than being forced to accept a decision based on the DM's opinion. Formally, this is captured by $\bar{U}_\mu(\mu) > \bar{U}_\mu(0)$. In this sense, delegation makes the agent *complacent* because he does not have to persuade the DM that his opinion is correct. This complacency effect blunts the agent's incentive for information acquisition, as seen from the fact that the terms being subtracted in the agent's marginal benefit of effort is larger in (16) than in (14).

It is the negative effect of delegation on the agent's initiative that is absent in Aghion and Tirole. As already noted, the reason is that in their model, allocation of authority does not affect the decision made without any information, precluding the complacency effect identified above. The net impact on the agent's initiative from delegation depends on the relative magnitude of the two effects. In the current setting, it turns out that this net effect coincides with the persuasion effect we have have identified earlier in equation (9):

$$U_\mu(0) - \bar{U}_\mu(0) - \big(U_\mu(\mu) - \bar{U}_\mu(\mu)\big) = \mu^2 - (B(\mu))^2 > 0. \tag{17}$$

In other words, delegation makes the agent more complacent (when no information is obtained) than it empowers him (when he obtains information), so it ultimately undermines the agent's incentive for information acquisition. Crucial is that the agent's persuasion motive for information acquisition only exists when the DM retains the decision right.

By the same token, delegation harnesses the DM's persuasion motive and strengthen's her incentives to acquire information. Nevertheless, the DM always prefers to retain authority. To see this, first notice that (17) implies that $\phi^C(P) > \phi^D(P)$ for all $P \in [0, \bar{p}]$, and a symmetric argument means that $\Phi^C(p) < \Phi^D(p)$ for all $p \in [0, \bar{p}]$. Since the best response curves are negatively sloped, we must have $\hat{p}^C > \hat{p}^D$ and $\hat{P}^C < \hat{P}^D$. Therefore,

$$\mathcal{U}^C_{DM}(\hat{P}^C, \hat{p}^C) \;\; \geq \;\; \mathcal{U}^C_{DM}(\hat{P}^D, \hat{p}^C) > \mathcal{U}^C_{DM}(\hat{P}^D, \hat{p}^D) > \mathcal{U}^D_{DM}(\hat{P}^D, \hat{p}^D), \tag{18}$$

where the first inequality follows from the fact that $\hat{P}^C$ is a best response to $\hat{p}^C$, the second holds because $\hat{p}^C > \hat{p}^D$, and the third follows from the loss of control observations earlier. The inequalities imply that the DM strictly prefers to retain authority no matter the difference of opinion with the adviser. Of course, if the adviser is like-minded, delegation and centralization are equivalent. We have proved:

Proposition 6. *Consider the simultaneous public information acquisition model with adviser of type $\mu \neq 0$. Relative to centralization, delegation demotivates the adviser in the sense of decreasing his effort, while increasing the DM's effort. Hence, the DM strictly prefers to retain authority.*

Remark 2. *While the first part of the proposition depends on the conflict arising from a difference of opinion, the second part does not. Suppose the agent has the same opinion* *as the DM, but conflicting fundamental preferences (represented as in Section 5 by a bias parameter $b \neq 0$). As discussed before, the agent then has no persuasion motive, so the incentive difference between the two regimes, captured by the expressions in (17), vanishes. In other words, the initiative gain from controlling the use of information and the initiative loss from complacency offset each other completely. This implies that in terms of affecting information acquisition incentives, the DM is indifferent between delegating and retaining authority. Yet, delegation still entails a costly loss of control, so in net, the DM strictly prefers to retain authority. Formally, in (18), the first and second inequalities become equalities (since $(\hat{P}^C, \hat{p}^C) = (\hat{P}^D, \hat{p}^D)$) but the last inequality remains strict.*

Simultaneous Private Information Acquisition. How does the effect of authority change when information is observed only privately by either party? Specifically, consider a modification of the simultaneous public information acquisition model so that both parties' effort choices and whether they observe s (and if so, what the value is) are private information. Under centralization, communication takes place as in our baseline model; under delegation, the principal communicates to the agent.

The main observation that delegation discourages the agent's information acquisition remains valid and even reinforced. This is because for any given level of effort from the DM, the agent loses not only the persuasion motive but also the motive to avoid prejudice when he can make decisions (recall the discussion following Proposition 3). All other effects are similar to those under public information, except that there is some loss of control for the DM even when she has formal authority. For, even under centralization, the agent is now able to influence the DM's decision by choosing when to reveal and when to withhold when he is the only one who has observed the signal. In this sense, even when the DM has complete formal authority, the agent has some real authority, as in Aghion and Tirole (1997) and Dessein (2002).

This additional wrinkle makes it difficult to assess whether delegation or centralization would be optimal for the DM when faced with an adviser of arbitrary type $\mu \neq 0$.[22] Nevertheless, a clear conclusion can be drawn once we endogenize the agent's type. Suppose, as before, that the DM can choose an adviser of type μ from a pool of potential advisers $[\underline{\mu}, \overline{\mu}]$, and moreover, now chooses whether to retain formal authority or delegate decision-

[22] Indeed, in the Crawford and Sobel (1982) model of cheap talk between an uninformed DM and an exogenously informed adviser, Dessein (2002) has shown that delegation is optimal whenever the conflict of interest is sufficiently small.

making to the chosen adviser. Then, the optimality of retaining formal authority follows from three observations. First, if the DM is going to delegate, she would strictly prefer to choose a like-minded adviser.[23] Second, when the adviser is like-minded, centralization and delegation are equivalent in the sense that either regime gives rise to the same incentives and decisions. Third, Proposition 4 implies that under centralization, the DM would strictly prefer an agent with some difference of opinion (i.e., $\mu \neq 0$) to a like-minded one. We summarize as follows:

PROPOSITION 7. *Consider simultaneous but private information acquisition. Under delegation, it is uniquely optimal for the DM to choose a like-minded adviser. However, such an arrangement is strictly worse for the DM than retaining authority and choosing an adviser with a sufficiently small difference of opinion.*

Our results do not argue against the general tenet of Aghion and Tirole (1997) that, given incomplete contracting and limited monetary incentives, allocation of formal authority affects both the decisions that are made and incentives for information acquisition. Our results do, however, caution against the conclusion that granting an agent formal authority will lead to an increase in his initiative; this depends crucially on the details of their model, in particular the feature that allocation of authority becomes irrelevant in the absence of information. When that feature is relaxed, a new effect emerges—initiative decrease from complacency—that offsets the initiative gain the agent acquires from controlling the use of information. This decomposition refines and clarifies Aghion and Tirole's seminal analysis of the initiative effects of delegation.

7 Discussion

We now discuss how our main themes continue to apply in some modifications of the model (for simplicity, return to the baseline case where only the adviser exerts effort), and then connect the current work to the existing literature.

Participation Constraints. In some situations, it may be necessary for the DM to ensure the adviser's participation in the relationship. Given non-contractibility of effort and information, this must be done with a fixed wage payment. Clearly, this does not

[23] Since this claim holds for a fixed level of effort by the DM, it holds when her effort chosen optimally.

affect the fundamental tradeoff between information acquisition and revelation for a given adviser type. But would this affect the optimal choice of adviser, in particular the result that it is not optimal to choose a like-minded adviser? This is potentially a concern because the greater the difference of opinion with the DM, the lower is the welfare for the chosen adviser in our baseline model, and a participation constraint may make the DM internalize, at least partially, the adviser's welfare.

There are two issues that arise in dealing with this question. First, it becomes crucial how much the DM cares about good decision-making relative to potential advisers, in the sense of their tradeoff between decision-utility and money (thus far, this has been irrelevant). Second, because advisers are intrinsically motivated in that they care about good decisions being made, an agent who is faced with the choice of advising the DM must anticipate the quality of decision-making if he chooses not to participate.

In Appendix B, we show that even when the DM must satisfy a participation constraint for the chosen adviser, it remains optimal to select an adviser with a different opinion so long as at least one of two reasonable conditions is satisfied: either the DM places more weight on good decision-making than advisers, or advisers (specifically, those close to like-minded) believe that in their absence, the DM will make decisions by consulting someone closer to the DM's opinion.[24] The intuition for the first case is that if the DM cares more about decision-making relative to money than an adviser, it is optimal to induce additional effort by choosing an adviser with a difference of opinion, even though such an adviser has to be compensated more than a like-minded one. On the other hand, in the second scenario, an adviser with a slightly different opinion does not have to be compensated entirely for the extra effort he would exert over a like-minded adviser, because such an adviser dislikes more the outcome if he does not participate.

Information and Manipulation. We have modeled the adviser's information acquisition and manipulation technology in a simple way. Specifically, the signal is either informative with a fixed precision or completely uninformative (no signal), and the adviser either withholds the information completely or reveals it without any manipulation. The main forces we have identified would hold under various other modeling choices, however. Suppose, for instance, that the adviser's effort determines the precision of his signal, ρ, and this is precision is observable by the DM. An interpretation is that the precision

[24] Alternatively, that if an adviser turns down the job, the DM will make a decision without any adviser, i.e. based only on her prior opinion.

represents the amount of credible evidence or convincing arguments associated with a signal. Naturally, effort cost is increasing in precision.[25] The persuasion effect carries over to this setting, because the DM will put more weight on the signal the higher is the signal's precision. Hence, the adviser expects that by generating a more precise signal, he will be more effective in persuading the DM toward his view of the world (this is in addition to the uncertainty reduction incentive); consequently, an adviser with a greater difference of opinion will have more incentive to exert effort towards raising signal precision.

Similarly, the prejudicial effect and its positive effect on information acquisition will arise when the adviser's manipulation technology takes a richer form than just withholding information. We illustrate by continuing with the variable precision model. Suppose now that, after observing signal s, an adviser can manipulate it to s' at a cost $c(|s'-s|,\rho)$. It is reasonable that the cost is increasing in the degree of manipulation, $|s'-s|$, and that high-precision signals are more difficult to manipulate than low-precision signals, so that the marginal cost of manipulation increases with the precision of signals, ρ. In such a setting, the degree of manipulation an adviser engages in after observing his signal will be higher when signals are less precise and, moreover, when he has a greater difference of opinion with the DM. Therefore, the DM optimally responds more prejudicially to the information delivered by an adviser with a greater difference of opinion (by choosing an action closer to her prior belief), but the extent of this prejudicial response decreases with the signal precision. In turn, this will give an adviser with a different opinion an incentive to raise precision in order to mitigate the prejudicial inference (aside from the persuasion motive), an incentive that is absent for the like-minded adviser.

Discounting Information. We have assumed that agents with different opinions agree on the conditional distribution of signals and update in a Bayesian fashion. Our results confirm the intuition that individuals should discount what is said by those with a different opinion, as the prejudicial effect increases in the difference of opinion. Nevertheless, one may argue that even setting aside strategic communication issues, individuals tend to discount information more when it is further from their prior views. Such behavior could arise for a variety of reasons. A Bayesian explanation would be uncertainty about the quality of signal, or otherwise put, about the adviser's competence. A different possibility is that individuals simply have a behavioral or cognitive bias, whereby they have a tendency to

[25]Our baseline model corresponds to a case in which the signal has zero precision ("no information") or some fixed positive precision, and effort determines precision stochastically rather than deterministically.

interpret inconclusive information as confirming current beliefs. This is referred to in the psychology literature as *confirmatory bias*.

Our model can be extended to accommodate such features, providing an additional insight about why individuals may not seek advice from those with overly different opinions. For simplicity, focus on the case of public information. Following Rabin and Schrag (1999), a confirmatory bias for the DM can be captured by assuming that she ignores the true signal (or misperceives it to be signal 0) with some probability that is increasing in the degree to which the signal is non-confirmatory, i.e. in the magnitude of the signal. Naturally, the DM is not aware of her bias, but to make the setting interesting, suppose the adviser is aware of the DM's bias.[26] A novel effect arises: an adviser with a greater difference of opinion perceives a higher chance of getting information that does not conform to the DM's prior, and thus anticipates being ignored with a higher likelihood than a less extreme adviser would anticipate. Thus, for any adviser type, the marginal benefit of exerting effort is smaller than in the setting without confirmatory bias for DM; moreover, this effect is stronger when the difference of opinion is larger. This implies it will generally not be optimal for the DM to choose an adviser with an extremely different opinion, in contrast to our baseline model with public information where the DM would choose the most extreme adviser available. On the other hand, because the persuasion effect still exists, it will remain optimal to choose an adviser with some difference of opinion, at least when the extent of confirmatory bias is not too large.[27]

A similar logic would also apply in a fully Bayesian setting where the quality of signal is unknown. In this case, upon observing a signal, the DM will update her belief about not only the state of the world but also the adviser's competence. In particular, when the DM observes information she regarded as ex-ante unlikely, she will update only partially about the state of world, ascribing higher probability to the adviser's incompetence. Understanding this, an adviser with a very different opinion anticipates observing signals that will lead the DM to conclude that he is of low competence, which mitigates the incentive to acquire information. Again, while this can cause the DM to choose an adviser who is less extreme than in the model with known competence, the persuasion effect implies that it will be sub-optimal to choose a like-minded adviser, at least when there is not too much uncertainty about competence. More broadly, this extension suggests that in situations

[26] It does not matter whether the adviser suffers from his own confirmatory bias, so long as he is not aware of it.

[27] A formal statement and proof is available on request.

where expertise of potential advisers (or the usefulness of their expertise for the particular decision problem) is less well established, we should observe appointments with smaller difference of opinion.

In reality, the DM's will vary in the degree to which they discount non-confirmatory information because of the heterogeneity in behavioral bias, in their knowledge of advisers' competence, or in the precision they assign to their own priors relative to the new information. For instance, DM's who are more confident about their own priors will more likely ascribe non-confirming information to advisers' incompetence. These differences may explain the observed variation in the willingness leaders exhibit in embracing differences of opinions in their teams.[28]

Related Literature. Let us relate our findings to a few recent papers that have explored related and complementary themes to ours. Dur and Swank (2005) show that when faced with a *binary decision*, an open-minded or moderate adviser values information more than an ex-ante biased adviser. Thus, an ex-ante biased DM may choose a more moderate adviser to encourage information acquisition, even though this adviser will engage in strategic communication. However, an ex-ante unbiased DM will find it optimal to choose an adviser of the same type. In contrast, in our model, the decision space is continuous and unbounded, hence there is no notion of being moderate; instead what the DM benefits from is a difference of opinion, and this is beneficial regardless of the DM's type.

Gerardi and Yariv (2007) study a jury model with binary decisions and public information. They show that it can be beneficial to appoint a juror with a different preference from the DM (in terms of the threshold of reasonable doubt), and in fact, the optimal juror is typically one who is extremely biased in the opposite direction of the DM; see also Dur and Swank (2005, Section 5.2). The intuition is transparent: a juror who is ex-ante biased in the opposite direction of the DM faces an unfavorable status quo. Hence, he can only gain from new information, which will at worst not change the status quo or, with luck, shift the DM's decision to his favored alternative.[29] With non-binary decisions, however,

[28]One example is that in appointing cabinet members "an Obama White House will brim with big personalities and far more spirited debate than occurred among the largely like-minded advisers who populated President [George W.] Bush's first term." Obama's cabinet appointment has been referred to as following the "violin model: Hold power with the left hand, and play the music with your right," by David J. Rothkopf, Deputy Undersecretary of Commerce during the Bill Clinton administration. ("Obama Tilts to Center, Inviting a Clash of Ideas," New York Times, November 22, 2008.)

[29]A related point is made by Hori (2008) who considers an agent hired to advise a principal on the adoption of a new project. The cost of providing incentives for information acquisition may be lower for

new information can lead to a decision that is less favorable than the status quo. In fact, an adviser may not expect new information to shift the DM's decision on average when the conflict is one of preferences alone. Consequently, as was shown in Section 5, in the current framework a preference conflict alone with public information does not motivate information acquisition. Gerardi and Yariv (2007) discuss the appointment of multiple jurors and sequential consultation.

The persuasion effect that we find from difference of opinion is also noted by Van den Steen (2004).[30] His model differs from ours, however, in a number of ways; for example he assumes both binary signals and decisions. Substantively, he does not study the prejudicial effect that occurs when information is privately observed by the adviser, and its implications for information acquisition. There is, therefore, no analogue of our finding that some difference of opinion is optimal for the DM even accounting for the cost of strategic disclosure. Instead, Van den Steen (2004) studies coordination issues between multiple individuals' action choices.

The optimality of having an adviser whose interim preferences are different from the DM's is reminiscent of Dewatripont and Tirole (1999) and Prendergast (2007), but the frameworks and forces at work are quite different.[31] Dewatripont and Tirole (1999) study the optimality of giving monetary rewards for "advocacy" from agents when the central problem is that of multitasking between conflicting tasks. In contrast, the effects in our model are derived solely from a single effort choice. Prendergast (2007) shows that a society may prefer to appoint bureaucrats with preferences different from its own, because such agents may have greater intrinsic motivation to exert effort. For example, in a setting with no conflicts of interest at all, society would appoint bureaucrats who are more altruistic towards their clients than the average member of society, simply because such bureaucrats care more about making good decisions and thus will exert more effort.

Finally, the current paper is consistent with the literature showing that ex-post suboptimal mechanisms are sometimes ex-ante optimal because they can provide greater incentives for agents to acquire information. Examples in settings with some similarities to ours are Li

an agent who is biased toward adoption, if the project is a priori unlikely to be profitable (in which case an unbiased agent may simply recommend non-adoption without learning about the project).

[30] This was brought to our attention after the first draft of our paper was complete; we thank Eric Van den Steen.

[31] We also note the work of Landier, Sraer, and Thesmar (2008), who show that when there is a complementarity of effort between a DM and an agent, it can be beneficial for an organization if they have heterogeneous intrinsic preferences over decisions. Again, the driving force is somewhat different from that of the current paper.

(2001) and Szalay (2005). The prejudicial effect in our model and its incentive implication on the adviser's effort can be interpreted in this light. In our model, the DM does not need commitment power to enforce suboptimal decisions: the prejudicial effect emerges in the unique equilibrium of the communication sub-game. One can view the DM as committing via her adviser choice. That said, we should emphasize that even when information is public—in which case there is no ex-post distortion for the DM—the persuasion motive still implies optimality of difference of opinion.

8 Concluding Remarks

This paper has explored incentives for information acquisition and transmission when there is a difference of opinion. While there are direct implications in the field of organizational economics—such as selecting an optimal type of agent and delegation of decision-making authority—the central tradeoff has broader applications. We conclude by mentioning a few.

Maverick sciences. The history of sciences is rich with episodes in which "maverick science,"—scientific inquiry that departs from mainstream or orthodox views—has led to important discoveries. Some of these discoveries are typically dismissed by the majority of the scientific community and/or general public early on and only gradually become accepted as a part of mainstream science. Galileo's struggle for heliocentricism against the church and community at large is particularly famous, as mentioned in our introduction, but there are many other well-known cases.[32] The role of maverick science in furthering scientific knowledge would be consistent with our analysis. Interpreting the public or general scientific community as the DM in our model, one can ask how well-served she is by a particular scientist (adviser), even if she does not volitionally choose him. Our results imply that new discoveries are more likely to come from those with different views about the world (than would be explainable by their population composition), and that public interest can be served by such individuals. Moreover, our theory also suggests that establishments are right to dismiss non-conformist theories until rigorously proven, both because scientists with unorthodox views may not reveal evidence that is contrary to their views, and because it is often the need to persuade and process of persuasion itself that

[32]For example, the germ theory of disease, the theory of continental drift, and black holes were all held in suspicion for long periods of time before they became commonly accepted.

leads to scientific breakthroughs.

Debate as a persuasion game. In the spirit of our model, debates can be seen as games of persuasion being played by individuals with possibly different opinions. This perspective sheds light on different aspects of debating. For instance, organizers of a debate in a public forum tend to invite individuals with opposing views on the issue to participate in the debate. An obvious reason may be to facilitate formation of balanced (posterior) opinion, but there is also a view that debates by people with opposing views tend to bring out more informed and fuller examination of the issue at hand. This latter idea is very much at the heart of our theoretical prediction that a difference of opinion heightens the efforts to persuade. The same idea also sheds light on the "art" of debating. People often play the role a *devil's advocate* by creating or accentuating opposition if they feel that another debater has put forth insufficient, inadequate or incomplete arguments. Playing a devil's advocate could be an effective way to induce a persuasion effort from a debating partner, so long as the view of the devil's-advocate-player is ex-ante unknown. Given uncertainty of a person's view, playing a devil's advocate will force the debating partner to take the opposing view seriously, and induce effort via the persuasion motive, according to our theory. We are developing this point more formally in ongoing work.

Persuasion in judiciaries. Our analysis also offers a useful perspective on legal systems and the behavior of judiciaries. An important role of a common law judge is to make a law by setting a new precedent. When a judge issues a ruling that seemingly goes against existing precedents or the prevailing general view on the matter, s/he is compelled to persuade not only future courts but also the general public. This may explain why a new legal theory or "doctrine" emerges often as a way to either rule against, or reconcile the current decision with, an existing precedent. A case in point is the well-cited common law doctrine on liability. When the *Davies v. Mann* court faced the challenge of ruling in favor of a "negligent" plaintiff—seemingly in violation of the *contributory negligence doctrine* set as an earlier precedent—, the court developed a new doctrine, *last clear chance*, which holds that a defendant would be liable, even when a plaintiff is negligent, if the former had the last clear chance to avoid the harm (and didn't). The persuasive appeal of the last clear chance doctrine is that it makes the new ruling appear not so much as an outright overturning of the earlier precedent but rather as a circumspect and informed qualification of its application (Cooter and Ulen, 2008, pp. 68–73).

In a similar vein, the organization and proceedings of the U.S. Supreme Court can be seen as harnessing persuasion incentives. Nine Supreme justices preside over the cases brought to them and produce written opinions. When there are serious dissensions among them, opinions representing the majority view and the dissenting view(s) are both written. It is reasonable to regard justices as persuading an audience (which could be lower courts or future courts, or the general public) with their views. Obviously, their views also reflect to some extent the views of their audience. Whenever there is a dissension within the justices, it is more likely that the views of their audience are divided, and hence—according to the persuasion effect—the incentives to convince the audience are stronger. This implies that justices are induced to deliberate more and otherwise exert more effort on issues with greater differences of opinions amongst them. This improves efficiency in the allocation of justices' attention and efforts. Consistent with this hypothesis, we find in a simple regression analysis that the written majority opinion of the U.S. Supreme Court tends to be longer the more dissenting judges there are.[33]

[33] Details are available from the authors.

A Appendix: Proofs

Proof of Proposition 1: By assumption, $p < 1$. We provide a proof for the case where $B \geq 0$. A symmetric argument can be used to establish the result for the opposite case of $B < 0$. We start by deriving an equation whose solution will constitute the equilibrium condition (5). First, it follows from (5) that

$$a_{\emptyset}(B,p) = a_N\left(p, \left[\underline{s}(B,p), \overline{s}(B,p)\right]\right).$$

Substituting in from (2), (3), and (4) gives the main equation:

$$\overline{s}(B,p) = \frac{p}{p\int_{\overline{s}(B,p)-\frac{2B}{\rho}}^{\overline{s}(B,p)} \gamma(s;0)\,ds + 1 - p} \int_{\overline{s}(B,p)-\frac{2B}{\rho}}^{\overline{s}(B,p)} s\gamma(s;0)\,ds. \tag{19}$$

We will show that there is a unique solution to (19) in two steps below; this implies that there is a unique disclosure equilibrium.

Step 1. *For any p, $\overline{s}(0,p) = \underline{s}(0,p) = a_{\emptyset}(0,p) = 0$.*

Proof: Immediate from the observation that $l(0,a) = h(a)$, and $a_N(p,S) = 0$ if S has measure 0. ‖

Step 2. *For any (B,p), there is a unique equilibrium in the disclosure game.*

Proof: Step 1 proves the result for $B = 0$, so we need only that show that there is a unique solution to (19) when $B > 0$. This latter is accomplished by showing that there is a unique solution to

$$\Upsilon(\overline{s}, B, p) := -p\int_{\overline{s}-\frac{2B}{\rho}}^{\overline{s}} (\overline{s} - s)\,\gamma(s;0)\,ds - \overline{s}(1-p) = 0. \tag{20}$$

Without loss of generality, we can restrict attention to $\overline{s} < 0$, because there is no solution to (20) with $\overline{s} \geq 0$ when $B > 0$. To see that there is at least one solution, apply the intermediate value theorem with the following observations: $\Upsilon(\overline{s}, B, p)$ is continuous in $\overline{s}$, and satisfies $\Upsilon(0, B, p) < 0$ and $\Upsilon(\overline{s}, B, p) \to \infty$ as $\overline{s} \to -\infty$ (because the integral term in (20) is bounded).

To prove uniqueness, observe

$$\begin{aligned}
\frac{\partial}{\partial \overline{s}}\Upsilon\left(\overline{s},B,p\right) &= p\left(1+2(B/\rho)\gamma\left(\overline{s}-2(B/\rho);0\right)-\int_{\overline{s}-2(B/\rho)}^{\overline{s}}\gamma\left(s;0\right)ds\right)-1 \\
&< p\left(1+2(B/\rho)\gamma\left(\overline{s}-2(B/\rho);0\right)-\int_{\overline{s}-2(B/\rho)}^{\overline{s}}\gamma\left(\overline{s}-2(B/\rho);0\right)ds\right)-1 \\
&= p-1<0,
\end{aligned}$$

where the inequality uses the fact that $\gamma\left(\cdot;0\right)$ is strictly increasing on the negative Reals. Consequently, there can only be one solution to (20). ‖

The comparative statics results are established in several steps again for the case $B \geq 0$ (a symmetric argument is applicable for the opposite case $B < 0$).

Step 3. *For any* $B \geq 0$*,* $p > 0$*,* $\frac{\partial}{\partial B}\overline{s}(B,p) < (=) 0$ *if* $B > (=0) 0$.

Proof: We showed earlier that $\frac{\partial}{\partial \overline{s}}\Upsilon\left(\overline{s},B,p\right) < 0$. We also have

$$\frac{\partial}{\partial B}\Upsilon\left(\overline{s},B,p\right) = -\frac{4Bp}{\rho^2}\gamma\left(\overline{s}-2(B/\rho);0\right) < 0.$$

By the implicit function theorem,

$$\frac{\partial \overline{s}\left(B,p\right)}{\partial B} = -\frac{\frac{\partial \quad \left(\overline{s}(B,p),B,p\right)}{\partial B}}{\frac{\partial \quad \left(\overline{s}(B,p),B,p\right)}{\partial \overline{s}}} \leq 0,$$

with equality if and only if $B=0$. ‖

Step 4. *For any* $B \geq 0, p > 0$*,* $\frac{\partial}{\partial B}\underline{s}(B,p) < 0$.

Proof: The result follows from Step 3, upon noting that $\underline{s}(B,p) = \overline{s}(B,p) - \frac{2B}{\rho}$. ‖

Step 5. *For any* $p > 0$*,* $\frac{\partial}{\partial p}\overline{s}(B,p) < (=)\, 0$ *if* $B > (=)\, 0$*; and* $\frac{\partial}{\partial p}\underline{s}(B,p) < (=)\, 0$ *if* $B > (=)\, 0$.

Proof: We showed earlier that $\frac{\partial}{\partial \overline{s}}\Upsilon\left(\overline{s},B,p\right) < 0$. We also have

$$\frac{\partial}{\partial p}\Upsilon\left(\overline{s},B,p\right) = -\int_{\overline{s}-2(B/\rho)}^{\overline{s}}\left(\overline{s}-s\right)\gamma\left(s;0\right)ds + \overline{s}.$$

By the implicit function theorem, $\frac{\partial \overline{s}(B,p)}{\partial p} = -\frac{\frac{\partial \Upsilon(\overline{s}(B,p),B,p)}{\partial p}}{\frac{\partial \Upsilon(\overline{s}(B,p),B,p)}{\partial \overline{s}}}$. The first statement is proven by noting that $\overline{s}(0,p) = 0$ and that, for any $B > 0$, $\overline{s}(B,p) \leq 0$. The second statement follows from the first statement, since $\underline{s}(B,p) = \overline{s}(B,p) - \frac{2B}{\rho}$. ‖

Step 6. *The nondisclosure action $a_\emptyset(B,p)$ is zero if $B=0$ or $p=0$, is strictly decreasing in B for $p>0$, and is strictly decreasing (increasing) in p if $B>0$ (if $B<0$).*

Proof: The result follows from inspection of (4), combined with the preceding Steps.‖

▮

Proof of Lemma 1: The adviser's expected payoff from choosing p given the DM's belief p^e is given by:

$$U_A(p;p^e,B,\mu) \;=\; p\Big[\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_A(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right]+\mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|s,\mu]\mid\mu\right]\Big]$$
$$+(1-p)\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|\mu]-c(p).$$

The first term decomposes the adviser's payoff when he obtains the signal (with probability p): he reveals the signal if $s\notin S(B,p^e)$, which leads to the action $\alpha_0(s)$ by the DM; and he withholds the signal when $s\in S(B,p^e)$, which leads to the action $a_\emptyset(B,p^e)$ by the DM. The second term is the payoff when the adviser does not observe the signal (which arises with probability $1-p$), in which case the DM picks $a_\emptyset(B,p^e)$. The last term is the cost of information acquisition.

The conclusion of the Lemma follows from manipulating terms:

$$\begin{aligned}
&U_A(p;p^e,B,\mu)\\
=\;& p\Big[\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_A(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right]+\mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|s,\mu]\mid\mu\right]\Big]\\
&\quad+(1-p)\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|\mu]-c(p)\\
=\;& p\left(\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_A(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right]+\mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|s,\mu]\mid\mu\right]\right)\\
&\quad+(1-p)\mathbb{E}_s[\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|s,\mu]]-c(p)\\
=\;& p\left(\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_A(\alpha_0(s),\omega)-u_A(a_\emptyset(B,p^e),\omega)\mid s,\mu]\mid\mu\right]\right)\\
&\quad+\mathbb{E}_s\left[\mathbb{E}_\omega[u_A(a_\emptyset(B,p^e),\omega)|s,\mu]\mid\mu\right]-c(p)\\
=\;& p\left(\int_{s\notin S(B,p^e)}\left[\left(a_\emptyset\left(B,p^e\right)-\rho s-B\right)^2-B^2\right]\gamma\left(s;\mu\right)ds\right)\\
&\quad-\int\left(a_\emptyset\left(B,p^e\right)-\left(\rho s+B\right)\right)^2\gamma\left(s;\mu\right)ds-\tilde{\sigma}^2-c(p),
\end{aligned}$$

where the last expression is obtained from substituting (1). ▮

Proof of Lemma 2: By Proposition 1, $a_\emptyset(B,0)=0$; hence we evaluate $\Delta(B,\mu,0)=\rho^2(\sigma_0^2+\sigma_1^2+\mu^2)+2\rho B\mu>0$ from (7). For any (B,μ), $\Delta(B,\mu,\cdot):[0,\bar{p}]\to\mathbb{R}$ is a bounded mapping. Therefore, by the Inada conditions, we have $c'(0) = 0 < \Delta(B, \mu, 0)$ and $c'(p) > \Delta(B, \mu, p)$ for large enough p. Since both sides of (8) are continuous in p, there exists $p \in (0, 1)$ that satisfies (8). It also follows that any equilibrium p must be interior, so it must satisfy (8). Finally, if p satisfies (8), we have

$$\frac{\partial U_A(\tilde{p}; p, B, \mu)}{\partial \tilde{p}} = \Delta(B, \mu, p) - c'(\tilde{p}) \gtreqless 0 \text{ if } \tilde{p} \lesseqgtr p,$$

due to the convexity of $c(\cdot)$, so p is an equilibrium effort choice. ▌

Proof of Proposition 3: This is a special case of Proposition 5. ▌

Proof of Proposition 4: Let $U(\mu)$ be the expected utility for the DM of appointing an adviser of type μ. We can write

$$U(\mu) := p(\mu)W(\mu) + (1 - p(\mu))V(\mu),$$

where

$$\begin{aligned} W(\mu) &:= w(\mu, p(\mu)), \\ V(\mu) &:= v(\mu, p(\mu)), \end{aligned}$$

where

$$w(\mu, p) := -\tilde{\sigma}^2 - \int_{\underline{s}(B(\mu),p)}^{\overline{s}(B(\mu),p)} \left(a_\emptyset(B(\mu), p) - s\rho\right)^2 \gamma(s; 0) ds,$$

and

$$v(\mu, p) := -\tilde{\sigma}^2 - \int_{-\infty}^{\infty} \left(a_\emptyset(B(\mu), p) - s\rho\right)^2 \gamma(s; 0) ds.$$

We shall prove that $U'(0) = 0$ but the right second derivative of $U(\mu)$ evaluated at $\mu = 0$, denoted $U''(0^+)$, is strictly positive.[34] This will imply that the DM prefers an adviser with some $\mu > 0$ to an adviser with $\mu = 0$.

Step 1. $p''(0) > p'(0) = 0$.

Proof: Rewrite the equilibrium condition (8) for the adviser's effort choice as:

$$\mathcal{A}(\mu, p(\mu)) - c'(p(\mu)) = 0, \tag{21}$$

[34] A symmetric argument would prove that the left second derivative of U at $\mu = 0$ is strictly negative.

where $\mathcal{A}(\mu,p) := \Delta(B(\mu),\mu,p)$. Substituting in $a_{\emptyset}(B,p) = \rho\overline{s}(B,p)$ into (7) and manipulating terms yields

$$\begin{aligned}\mathcal{A}(\mu,p) &= \int_{s\notin S(B(\mu),p)} \rho(s-\overline{s}(B(\mu),p))\left(2(1-\rho)\mu+\rho(s-\overline{s}(B(\mu),p))\right)\gamma(s;\mu)ds \\ &= 2\mu(1-\rho)\rho\int_{s\notin S(B(\mu),p)}(s-\overline{s}(B(\mu),p))\gamma(s;\mu)ds \\ &\quad +\rho^2\int_{s\notin S(B(\mu),p)}(s-\overline{s}(B(\mu),p))^2\gamma(s;\mu)ds. \qquad (22)\end{aligned}$$

Using subscripts on $\mathcal{A}$ to denote partial derivatives in the usual way, a routine albeit tedious calculation (available in a supplementary appendix) yields

$$\mathcal{A}_{11}(0,p) > \mathcal{A}_1(0,p) = \mathcal{A}_2(0,p) = 0. \qquad (23)$$

There is a unique solution to (21) at $\mu = 0$ because $\mathcal{A}(0,p)$ is a positive constant independent of p (since $\overline{s}(0,p) = \underline{s}(0,p) = 0$) and $c'(p)$ is strictly increasing with range $[0,\infty)$. Since $\mathcal{A}_2(0,p) = 0 < c''(p)$, the implicit function theorem implies that there is some neighborhood of $\mu = 0$ where $p(\mu)$ is unique, continuously differentiable, and satisfies

$$p'(\mu) = \frac{\mathcal{A}_1(\mu,p(\mu))}{c''(p(\mu)) - \mathcal{A}_2(\mu,p(\mu))}. \qquad (24)$$

It follows from (23) and (24) that $p'(0) = 0$. That $p''(0) > 0$ is seen as follows:

$$\begin{aligned}p''(0) &= \lim_{\mu\to 0}\frac{p'(\mu)-p'(0)}{\mu} \\ &= \lim_{\mu\to 0}\frac{\mathcal{A}_1(\mu,p(\mu))}{\mu\left[c''(p(\mu)) - \mathcal{A}_2(\mu,p(\mu))\right]} \\ &= \lim_{\mu\to 0}\frac{\mathcal{A}_{11}(\mu,p(\mu)) + \mathcal{A}_{12}(\mu,p(\mu))p'(\mu)}{\mu\left[c'''(p(\mu))p'(\mu) - \mathcal{A}_{21}(\mu,p(\mu)) - \mathcal{A}_{22}(\mu,p(\mu))p'(\mu)\right] + 1\left[c''(p(\mu)) - \mathcal{A}_2(\mu,p(\mu))\right]} \\ &= \frac{\mathcal{A}_{11}(0,p(0))}{c''(p(0))} \\ &> 0,\end{aligned}$$

where the second line uses (24) and $p'(0) = 0$, the third line applies L'Hopital's rule, the fourth line again uses $p'(0) = 0$, and the last line uses $\mathcal{A}_{11}(0,p(0)) > 0$.

STEP 2. $W(0) - V(0) > 0$ and $W'(0) = V'(0) = W''(0) = V''(0) = 0$.

PROOF: Recall $B(0) = 0$ and from Proposition 1 that $\overline{s}(0,p) = \underline{s}(0,p) = 0$. Hence, for any p,

$$w(0,p) - v(0,p) = \int_{-\infty}^{\infty} (a_{\emptyset}(0,p) - \rho s)^2 \gamma(s;0)ds > 0,$$

from which it follows that $W(0) - V(0) > 0$. For any p, direct computation (available in the supplementary appendix) yields

$$v_1(0,p) = v_2(0,p) = v_{11}(0,p) = w_1(0,p) = w_2(0,p) = w_{11}(0,p) = 0. \tag{25}$$

It follows from (25) and Step 1 that

$$\begin{aligned} W'(0) &= w_1(0,p(0)) + p'(0)\, w_2(0,p(0)) = 0, \\ V'(0) &= v_1(0,p(0)) + p'(0)\, v_2(0,p(0)) = 0, \\ W''(0) &= w_{11}(0,p(0)) + w_{12}(0,p(0))\, p'(0) \\ &\quad + p'(0)\left(w_{12}(0,p(0)) + w_{22}(0,p(0))\, p'(0)\right) + p''(0)\, w_2(0,p(0)) \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} V''(0) &= v_{11}(0,p(0)) + v_{12}(0,p(0))\, p'(0) \\ &\quad + p'(0)\left(v_{12}(0,p(0)) + v_{22}(0,p(0))\, p'(0)\right) + p''(0)\, v_2(0,p(0)) \\ &= 0. \parallel \end{aligned}$$

From Step 1 and Step 2, we obtain

$$U'(0^+) = p'(0)(W(0) - V(0)) + p(0)W'(0) + (1 - p(0))V'(0) = 0,$$

and

$$\begin{aligned} U''(0^+) &= p''(0)(W(0) - V(0)) + 2p'(0)(W'(0) - V'(0)) + p(0)W''(0) + (1 - p(0))V''(0) \\ &= p''(0)(W(0) - V(0)) > 0. \end{aligned}$$

Combined, $U'(0^+) = 0$ and $U''(0^+) > 0$ imply that there exists $\mu > 0$ such that $U(\mu) > U(0)$. ∎

Proof of Proposition 5: Consider any pair (B,μ) and (B',μ') satisfying the hypothesized condition. It is without loss to assume $B'\geq 0$ and $\mu'\geq 0$. Further, since $(B(b,\mu),\mu)$ and $(B(-b,-\mu),-\mu)$ are payoff equivalent and thus generate the same incentive for the adviser, it is without loss to assume $\mu\geq 0$. The condition then reduces to $(0,0)\leq(|B|,\mu)<(B',\mu')$. We focus on the case in which $B\geq 0$. As we will argue later, the case of $B<0$ can be treated by the same argument applied twice, one for a shift from (B,μ) to $(-B,\mu)$, and another for a shift from $(-B,\mu)$ to (B',μ').

Let $p(B,\mu)$ be the (largest) p supported in equilibrium given an adviser with (B,μ). Suppose now an adviser with (B',μ') is chosen, but the DM *believes* that the adviser will continue to choose $p=p(B,\mu)$. We prove below that, given such a belief, the adviser of type (B',μ') will choose strictly higher $p'>p(B,\mu)$. It will then follow that, since the adviser's best response correspondence is upper-hemicontinuous in the DM's belief (by the Theorem of the Maximum), there must exist $p''>p(B,\mu)$ such that p'' is supported under (B',μ'), which would imply that $p(B',\mu')>p(B,\mu)$.

To prove the statement, suppose to the contrary that the adviser with (B',μ') will find it optimal to choose $p'\leq p(B,\mu)$ given the DM's belief that the adviser will choose $p(B,\mu)$. The disclosure sub-game following the effort choice p' is characterized by the pair $(S(B',p(B,\mu)),\ a_\emptyset(B',p(B,\mu)))$. By the first-order condition, we must then have

$$\Delta(B',\mu',p(B,\mu))=c'(p')\leq c'(p(B,\mu))=\Delta(B,\mu,p(B,\mu)). \tag{26}$$

For notational simplicity, let $S(\tilde{B}):=S(\tilde{B},p(B,\mu))$, $\overline{s}(\tilde{B}):=\overline{s}(\tilde{B},p(B,\mu))$, and $\underline{s}(\tilde{B}):=\underline{s}(\tilde{B},p(B,\mu))$, and let $\overline{s}:=\overline{s}(B)$ and $\overline{s}':=\overline{s}(B')$.

The proof follows several steps.

Step 1. The following inequality holds:

$$\begin{aligned}\Delta(B',\mu',p(B,\mu)) &\geq \Pi(B',\mu')\\ &:= \int_{s\notin S(B)}\left[\left(a_\emptyset(B',p(B,\mu))-\rho s-B'\right)^2-B'^2\right]\gamma(s;\mu')ds.\end{aligned}$$

Proof: By picking a nondisclosure interval S, given his type μ, effort p, and the DM's nondisclosure action $a_\emptyset$, the adviser's expected utility is

$$\pi\left(S;B,\mu,p,a_\emptyset\right):=p\int_{s\notin S}\left[\left(a_\emptyset-\rho s-B\right)^2-B^2\right]\gamma\left(s;\mu\right)ds-\left(a_\emptyset-\mu-b\right)^2-\sigma_0^2.$$

Thus, since the adviser chooses $S(B', p(B,\mu))$ rather than $S(B, p(B,\mu))$, it must be that

$$\pi\left(S(B', p(B,\mu)); B', \mu', p, a_\emptyset\left(B', p\left(B,\mu\right)\right)\right) \geq \pi\left(S(B, p(B,\mu)); B', \mu', p, a_\emptyset\left(B', p\left(B,\mu\right)\right)\right),$$

which implies the desired inequality by the definition of π. ‖

STEP 2. $\Pi(B',\mu') > \Pi(B,\mu)$.

PROOF: By substituting $a_\emptyset(B'; p(B,\mu)) = \rho\overline{s}(B'; p(B,\mu))$, we can write

$$\begin{aligned}\Pi(B',\mu') &= (2B'\rho - 2\rho^2\overline{s}') \int_{s\notin S(B)} s\gamma(s;\mu')ds + \rho^2 \int_{s\notin S(B)} s^2\gamma(s;\mu')ds \\ &\quad + \left(\rho^2\overline{s}'^2 - 2\rho B'\overline{s}'\right) \int_{s\notin S(B)} \gamma(s;\mu')ds. \end{aligned} \tag{27}$$

The proof is completed by showing that

$$\begin{aligned}\Pi(B',\mu') &= \Pr\{s\notin S(B) \mid \mu'\}\left[2\rho(B'-\rho\overline{s}')\mathbb{E}[s \mid s\notin S(B),\mu'] + \rho\overline{s}'\left(\rho\overline{s}' - 2B'\right)\right] \\ &\quad +\rho^2 \int_{s\notin S(B)} s^2\gamma(s;\mu')ds \\ &> \Pr\{s\notin S(B) \mid \mu\}\left[2\rho(B-\rho\overline{s})\mathbb{E}[s \mid s\notin S(B),\mu'] + \rho\overline{s}\left(\rho\overline{s} - 2B\right)\right] \\ &\quad +\rho^2 \int_{s\notin S(B)} s^2\gamma(s;\mu')ds \qquad (28) \\ &\geq \Pr\{s\notin S(B) \mid \mu\}\left[2\rho(B-\rho\overline{s})\mathbb{E}[s \mid s\notin S(B),\mu'] + \rho\overline{s}\left(\rho\overline{s} - 2B\right)\right] \\ &\quad +\rho^2 \int_{s\notin S(B)} s^2\gamma(s;\mu)ds \qquad (29) \\ &= \Pi(B,\mu). \end{aligned}$$

The two equalities above are immediate from (27); so we must show the two inequalities, (28) and (29).

Inequality (28) follows from the fact that $(B',\mu') > (B,\mu)$ and that $B \geq 0$. That $B' \geq B \geq 0$ implies $0 \geq \overline{s} = \overline{s}(B) \geq \overline{s}(B') = \overline{s}'$ (by Proposition 1), which in turn implies that

$$B' - \rho\overline{s}' \geq B - \rho\overline{s} \geq 0. \tag{30}$$

From (30) and $B' \geq B \geq 0 \geq \overline{s} \geq \overline{s}'$, it further follows that

$$\overline{s}'\left(\rho\overline{s}' - 2B'\right) \geq \overline{s}(\rho\overline{s} - 2B). \tag{31}$$

Next, since the Normal density $\gamma(\cdot;\mu')$ dominates in likelihood ratio the density $\gamma(\cdot;\mu)$,

$$\mathbb{E}[s \mid s \notin S(B), \mu'] \geq \mathbb{E}[s \mid s \notin S(B), \mu]. \tag{32}$$

Since $S(B) \subseteq \mathbb{R}_-$ (by Proposition 1) and $\mu' \geq 0$, we also have

$$\mathbb{E}[s \mid s \notin S(B), \mu'] \geq 0. \tag{33}$$

Finally, $\mu' \geq \mu$ implies that

$$\Pr\{s \notin S(B) \mid \mu'\} \geq \Pr\{s \notin S(B) \mid \mu\}. \tag{34}$$

(30)–(34) and $\rho > 0$ imply inequality (28) in weak form. The inequality holds strictly, however, since either $B' > B$ or $\mu' > \mu$, which means that one of (30)–(34) is strict.

Inequality (29) is established as follows:

$$\begin{aligned}
\int_{s \notin S(B)} s^2 \gamma(s;\mu') ds &= \mathbb{E}[s^2|\mu'] - \Pr\{s \in S(B) \mid \mu'\}\, \mathbb{E}[s^2|s \in S(B), \mu'] \\
&= {\mu'}^2 - \sigma_1^2 - \Pr\{s \in S(B) \mid \mu'\}\, \mathbb{E}[s^2|s \in S(B), \mu'] \\
&\geq \mu^2 - \sigma_1^2 - \Pr\{s \in S(B) \mid \mu\}\, \mathbb{E}[s^2|s \in S(B), \mu] \\
&= \int_{s \notin S(B)} s^2 \gamma(s;\mu) ds,
\end{aligned}$$

where the inequality follows from $|\mu'| \geq |\mu|$, since $\mathbb{E}[s^2|\mu'] \leq \mathbb{E}[s^2|\mu]$ (which follows from the fact that $\gamma(\cdot;\mu')$ likelihood-ratio dominates $\gamma(\cdot;\mu)$ and that s^2 is decreasing in s for $s \in S(B) \subseteq \mathbb{R}_-$), and since $\Pr\{s \in S(B) \mid \mu'\} \leq \Pr\{s \in S(B) \mid \mu\}$. ‖

Combining Step 1 and Step 2, we have

$$\Delta(B', \mu', p(B, \mu)) > \Pi(B, \mu). \tag{35}$$

By definition, it also follows that

$$\Pi(B, \mu) = \Delta(B, \mu, p(B, \mu)). \tag{36}$$

Combining (35) and (36) yields

$$\Delta(B', \mu', p(B, \mu)) > \Delta(B, \mu, p(B, \mu)),$$

which contradicts (26). We have thus proven the statement of the proposition.

The case of $B < 0$ can be treated by applying the same sequence of arguments twice, one for a shift from (B, μ) to $(-B, \mu)$, and then another for a shift from $(-B, \mu)$ to (B', μ'). The second step satisfies the hypothesized condition, so the same argument works. The first step poses a slightly novel situation with Step 2. Yet, the same inequality works with $(B', \mu') := (-B, \mu)$. ▌

B Participation Constraints

As discussed in the text, there are two issues with formulating participation constraints in our setting. First, it becomes crucial how much the DM cares about good decision-making relative to advisers, in the sense of their trade off between decision-utility and money (thus far, this has been irrelevant). Second, because advisers are intrinsically motivated (they care about good decisions being made), an agent who is faced with the option of advising the DM must consider the quality of decision-making if he chooses not to participate. To flexibly deal with both these considerations, we posit the following problem for the DM:

$$\max_{\mu \in [\underline{\mu}, \overline{\mu}], t \geq 0} \lambda U_0(\mu) - t \text{ s.t.}$$

$$U_\mu(\mu) - c(p(\mu)) + t \geq U_\mu(f(\mu)). \tag{IR}$$

Here, $U_\mu(\tilde{\mu})$ is the expected utility for an agent of opinion μ solely from decision-making when the DM hires an agent of type $\tilde{\mu}$ — this function is computed from our previous analysis. Note that in the absence of participation constraints, the DM just maximizes $U_0(\mu)$. The variable t is the monetary transfer or wage paid by the DM to the chosen adviser; $p(\mu)$ is the effort exerted by a chosen adviser of type μ; and $f(\mu)$ is a function indicating the "default" adviser type if an adviser of type μ who is approached by the DM chooses not to participate.[35] We suppose that $f(\cdot)$ is twice differentiable, and that for all μ, $|f(\mu)| \leq |\mu|$. This inequality captures the substantive notion that an adviser who turns down the job believes that the DM will hire someone who is (weakly) closer to the DM's opinion.[36] Note that an implication is that $f(0) = 0$ and $f'(0) \leq 1$. We have scaled the DM's decision-utility, $U_0(\mu)$, by a parameter λ: this allows for the DM to value decision-making differently from advisers, since λ affects the rate of substitution between decision-utility and money. It is reasonable to suppose that the DM cares at least as much about making good decisions as other agents, so we assume that $\lambda \geq 1$.

To focus clearly on the impact of a participation constraint, we consider only the case of publicly observed or perfectly informative signals, so that strategic communication issues are suppressed. A similar logic applies even when information is strategically communi-

[35]It is without loss of generality to not also include a fixed outside option utility for all adviser types as that can just be normalized to zero.

[36]Another possibility would be to assume that if the adviser does not participate, the DM makes a decision without any adviser, in which case there is no signal drawn and she just chooses $a = 0$. The Proposition below holds under this specification as well.

cated.

PROPOSITION 8. *Assume public or perfectly informative signals, and the participation constraint (IR). If $\lambda > 1$ or $f'(0) < 1$, it is optimal for the DM to choose an adviser with a different opinion.*

Proof of Proposition 8: Under public or perfectly informative signals,

$$U_\mu(\tilde{\mu}) = p(\tilde{\mu})(-\tilde{\sigma}^2 - (B(\mu))^2) + (1 - p(\tilde{\mu}))(-\sigma_0^2 - \mu^2).$$

Hence, $U_\mu(\mu) - U_\mu(f(|\mu|))$ is continuous in μ and converges to 0 as $\mu \to 0$. If some adviser type $\mu \neq 0$ can be hired at a zero wage, the result is immediate because $p(\mu)$ is strictly increasing in $|\mu|$. So suppose that any $\mu \neq 0$ must be paid a strictly positive wage to ensure participation. This requires t to solve (IR) with equality in any solution (including at $\mu = 0$, by continuity). Thus, the DM's problem reduces to

$$\max_{\mu \in [\underline{\mu}, \overline{\mu}]} g(\mu), \tag{37}$$

where

$$g(\mu) := \lambda(-p(\mu)\tilde{\sigma}^2 - (1 - p(\mu))\sigma_0^2) - c(p(\mu)) + (p(\mu) - p(f(|\mu|)))(\sigma_0^2 - \tilde{\sigma}^2 + \mu^2 - (B(\mu))^2).$$

It suffices to show that the solution to (37) is not 0 when either $\lambda > 1$ or $f'(0) < 1$. Recall from earlier that $p''(0) > 0 = p'(0)$. Using also the facts that $f(0) = 0$ and $c'(p(0)) = \sigma_0^2 - \tilde{\sigma}^2$ (from the first-order condition), a routine but tedious calculation (available in the supplementary appendix) shows that $g'(0) = 0$ while

$$g''(0) = p''(0)\left(\sigma_0^2 - \tilde{\sigma}^2\right)\left(\lambda - (f'(0))^2\right).$$

Since $p''(0) > 0$, $(\sigma_0^2 - \tilde{\sigma}^2) > 0$, $\lambda \geq 1 \geq f'(0)$, the desired conclusion follows.∎

References

ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2007): "Learning and Disagreement in an Uncertain World," mimeo, Massachusetts Institute of Technology. 4

Aghion, P., and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29. 6, 22, 23, 27, 28

Aumann, R. J. (1976): "Agreeing to Disagee," *Annals of Statistics*, 4, 1236–1239. 4

Banerjee, A., and R. Somanathan (2001): "A Simple Model Of Voice," *The Quarterly Journal of Economics*, 116(1), 189–227. 4

Cooter, R., and T. Ulen (2008): *Law and Economics*. Pearson: Addison Wesley, 5th edition edn. 35

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451. 7, 27

Degroot, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York. 10

Dekel, E., D. Fudenberg, and D. K. Levine (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46(2), 282–303. 9

Dessein, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69(4), 811–838. 27

Dewatripont, M., and J. Tirole (1999): "Advocates," *Journal of Political Economy*, 107(1), 1–39. 33

Dixit, A. K., and J. W. Weibull (2007): "Political Polarization," mimeo, Princeton University and Stockholm School of Economics. 4

Dur, R., and O. H. Swank (2005): "Producing and Manipulating Information," *Economic Journal*, 115(500), 185–199. 32

Galilei, G. (1953): *Dialogue Concerning the Two Chief World Systems, Ptolemaic & Copernican*. Berkeley, University of California Press. 2

Gerardi, D., and L. Yariv (2007): "Costly Expertise," mimeo, Yale University and Caltech. 32, 33

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law & Economics*, 24(3), 461–483. 7

Grossman, S. J., and O. D. Hart (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(4), 691–719. 9

Hori, K. (2008): "The Role of Private Benefits in Information Acquisition," *Journal of Economic Behavior and Organization*, 68(3–4), 626–631. 32

Janis, I. L. (1972): *Victims of Groupthink: A Psychological Study of Foreign-policy Decisions and Fiascoes*. Houghton Mifflin. 2

Landier, A., D. Sraer, and D. Thesmar (2008): "Optimal Dissent in Organizations," forthcoming in *Review of Economic Studies*. 33

Li, H. (2001): "A Theory of Conservatism," *Journal of Political Economy*, 109(3), 617–636. 33

Matthews, S., and A. Postlewaite (1985): "Quality Testing and Disclosure," *RAND Journal of Economics*, 16(3), 328–340. 7

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391. 7

Morris, S. (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 11, 227–253. 3

Prendergast, C. (2007): "The Motivation and Bias of Bureaucrats," *American Economic Review*, 97(1), 180–196. 33

Rabin, M., and J. L. Schrag (1999): "First Impressions Matter: A Model of Confirmatory Bias," *Quarterly Journal of Economics*, 114(1), 37–82. 31

Shavell, S. (1994): "Acquisition and Disclosure of Information Prior to Sale," *RAND Journal of Economics*, 25(1), 20–36. 7

Shin, H. S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64(1), 253–264. 9

——— (1998): "Adversarial and Inquisitorial Procedures in Arbitration," *RAND Journal of Economics*, 29(2), 378–405. 9

Szalay, D. (2005): "The Economics of Clear Advice and Extreme Options," *Review of Economic Studies*, 72, 1173–1198. 34

Tversky, A., and D. Kahneman (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131. 4

Van den Steen, E. (2004): "The Costs and Benefits of Homogeneity, with an Application to Culture Clash," mimeo, MIT Sloan. 33

Yildiz, M. (2004): "Waiting to Persuade," *Quarterly Journal of Economics*, 119(1), 223–248. 17