Opinions as Incentives*

Yeon-Koo Che† Navin Kartik‡

November 6, 2007

Abstract

We study a model where a decision maker (DM) must select an adviser to advise her about an unknown state of the world. There is a pool of available advisers who all have the same underlying preferences as the DM; they differ, however, in their prior beliefs about the state, which we interpret as differences of opinion. We derive a tradeoff faced by the DM: an adviser with a greater difference of opinion has greater incentives to acquire information, but reveals less of any information she acquires, via strategic disclosure. Nevertheless, it is optimal to choose an adviser with at least some difference of opinion. The analysis reveals two novel incentives for an agent to acquire information: a "persuasion" motive and a motive to "avoid prejudice." Delegation is costly for the DM because it eliminates both these motivations. We also study the relationship between difference of opinion and difference of preference.

*We would like to acknowledge the active input of Jimmy Chan at early stages of this project. We thank Nageeb Ali, Arnaud Costinot, Vince Crawford, Jean Guillaume Forend, Michihiro Kandori, Kohei Kawamura, Li Hao, Bart Lipman, Eric Maskin, Carolyn Pitchik, Jennifer Reinganum, Ed Schlee, Joel Sobel, and various seminar audiences for their opinions. David Eil, Chulyoung Kim, and Uliana Loginova provided excellent research assistance. Kartik is grateful to the Institute for Advanced Study at Princeton and the National Science Foundation for funding; he also thanks the Institute for its hospitality.

†Economics Department, Columbia University; yc2271@columbia.edu.

‡Department of Economics, University of California, San Diego; nkartik@ucsd.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*.

The above 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. Indeed, difference of opinion is valued in many organizations and situations. Corporations seek diversity in their workforce allegedly 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 coined "groupthink" by psychologists (e.g. Janis, 1972). Debates between individuals can be more illuminating when they take different views; in their absence, debaters often create an artificial difference 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 develop a model in which a decision maker, or DM for short, consults an adviser before making a decision. There is an unknown state of the world that affects both individuals' payoff from the decision. We model the DM's decision and the state of the world as real numbers, and assume the DM's optimal decision coincides with the state. Initially, neither the DM nor the adviser has any information about the state beyond their prior views. The adviser can exert effort to try and produce an informative signal about the state, which occurs with probability that 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 can, nevertheless, choose not to disclose the information he acquires. Finally, the DM takes her decision optimally given her updated beliefs after communication with the adviser.

This framework captures common situations 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. But the model permits broader interpretations: the DM could be the general public (such as 17th century intelligent laymen), and its decision is simply the posterior belief on some matter. In turn, the adviser could be a scientist (such as Galileo), investigator, special counsel, a lobbying group, or a debater trying to sway that belief.

It is often the case, as in the examples mentioned above, that an adviser is interested in the decision taken by DM. We assume initially that the adviser has the same fundamental preferences as the DM about which decision to take 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.[1] Such disagreements abound in many circumstances, as has also been argued by, for example, Banerjee and Somanathan (2001). 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. As another example, a general and her advisors may agree on the objective of winning a war at minimum cost. They may have

[1] That is, they "agree to disagree." Such an open disagreement may arise from various sources: individuals may 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 (Acemoglu, Chernozhukov, and Yildiz, 2007; Aumann, 1976; Tversky and Kahneman, 1974). Whatever the reason, open disagreements do exist and often persist even after extensive debates and communication.

different beliefs, however, about the strength of the opposition troops, leading to disagreements about how many of their own troops should be sent into combat—disagreements that do not change even when told each other's views. Many political disagreements also seem best viewed through the lens of different prior beliefs rather than different fundamental preferences (Dixit and Weibull, 2007).[2]

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 this, 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, despite having the same preference, will generally arrive at different posteriors about what the right decision is given partially-informative signals. Consequently, 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 as much effort as possible. We find that differences of opinion provide incentives for information acquisition, for two distinct reasons. First, *an adviser with a difference of opinion is motivated to persuade the DM.* Such an adviser believes that the DM's prior opinion is wrong, and 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 the like-minded adviser. Second, and more subtle, *an adviser with difference of opinion will exert effort to avoid "prejudice."* Intuitively,

[2] To mention just two examples, consider disagreements about how serious the global warming problem is (if it exists all, to some) 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 these in order to focus on the incentive implications of difference of opinion.

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 a hard evidence—this equilibrium feature of strategic interaction is what we call a "prejudicial effect." Hence, an adviser with difference of opinion will have incentives to seek out information in order to avoid prejudice, a motivation 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 clearly recognized by Jefferson) it also leads to increased efforts to understand and convince other individuals—this emerges here as increased information acquisition.

How should the DM resolve this tradeoff between information acquisition and transmission? We find that *the DM prefers an adviser with some difference of opinion to a perfectly like-minded one.* The reason is that an adviser with sufficiently small difference of opinion engages in only a negligible amount of strategic withholding of information, so the loss associated with such a difference 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 strictly benefits from an adviser with at least a little difference in opinion, and would not optimally choose a like-minded adviser from a rich pool of available individuals. Applied to a broader context, our result implies 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 the public interest can be served by them even though they will not reveal what they know all the time.

Sections 2–4 formalize our model and the above logic. Section 5 then augments the model to allow the adviser to differ from the DM in both his opinion and his fundamental preferences over decisions given the state of the world. Heterogeneous preferences have a similar effect as difference of opinion on strategic disclosure. But this similarity does *not* extend to the adviser's choice of effort, because the two attributes are fundamentally distinct in terms of how they motivate the adviser. While an adviser with difference of opinion has a persuasion motive for acquiring a signal—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 has no persuasion motive. For this reason, having an advisor who differs only in preferences yields no clear benefit for the DM.

Nevertheless, we find the difference of preferences to be valuable in the presence of difference of opinion. In other words, an adviser with a different opinion has more incentive to acquire information if he also has an 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.

We explore some other issues in Section 6. Of particular interest, we find that the benefit from difference of opinion is lost when the DM delegates the decision authority to the adviser. This observation sheds new light on the merit of delegation in organizational settings (cf. Aghion and Tirole, 1997). We also discuss implications of the adviser's perception of the precision of his own information, or his *confidence*, finding that more confident advisers exert more effort.

Our paper builds on the literature on strategic communication, combining elements from the structure of conflicts of interest in Crawford and Sobel (1982) with the verifiable disclosure game first introduced by Grossman (1981) and Milgrom (1981). The key innovation in this regard is that we endogenize information acquisition and focus on the effects of difference of prior beliefs. It is best to postpone a discussion of the most closely related literature to Section 7, after a full development of our model and analysis. Section 8 then concludes with a discussion of our modeling choices and possible extensions. The Appendix contains omitted proofs.

2 Model

A decision maker (DM) must take a decision, $a \in \mathbb{R}$. The appropriate decision 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 can choose a single adviser from a pool of available agents to advise her.

Prior Beliefs. We allow individuals—available advisers and the DM—to have different prior beliefs about the state. Specifically, while all individuals know the state is distributed according to a Normal distribution with variance $\sigma_0^2 > 0$, individual i believes the mean of the distribution is μ_i. The prior beliefs of each person are common knowledge.[4] We will refer to an adviser's prior belief as his *opinion* or *type*, even though it is not private information. Two individuals, i and j, have differences of opinion if $\mu_i \neq \mu_j$. Without loss of generality, we normalize the DM's prior to $\mu = 0$. An adviser with $\mu = 0$ is said to be *like-minded*.

Full-information preferences. All players have 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 we use is a common specification in the literature: it captures the substantive notion that decisions are progressively worse the further they are from the true state, and technically, makes the analysis tractable.

Information Acquisition. Regardless of the chosen adviser's type, his investigation technology is the same, described as follows. He chooses the probability that his investigation is successful, $p \in [0, 1]$, at a 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 $p \to 1$. We will interchangeably refer to p as an effort level or a probability.[5] 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 signal, denoted by $\emptyset$. Thus, effort is success-enhancing in the sense of Green and Stokey (2007) and increases information in the sense of Blackwell (1951).

[4] 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. Banerjee and Somanathan (2001) and Spector (2000) do so in communication models with exogenous information; in other contexts, examples are Eliaz and Spiegler (2006), Harrington (1993), Van den Steen (2005), and Yildiz (2003). For a general discussion about non-common priors, see Morris (1995).

[5] This is justified because our formulation is equivalent to assuming the adviser chooses some effort e at cost $c(e)$, which maps into a probability $p(e)$, where $p(0) = 0$ and $p(\cdot)$ is a strictly increasing function.

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 choice to make. The signal may be non-manipulable because there are large penalties against fraud, information is easily verifiable by the DM once received (even though impossible to acquire herself), or information is technologically hard to manipulate.[6]

Timing. The sequence of events is as follows. First, the DM picks the adviser's prior belief, $\mu \in \mathbb{R}$, which we interpret as choosing which adviser to consult from a rich pool. Then the adviser 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 a decision.

As this is multi-stage Bayesian game, it is appropriate to solve it using the concept of perfect Bayesian equilibrium (Fudenberg and Tirole, 1991), or for short, *equilibrium* hereafter. We restrict attention to pure strategy equilibria.

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. Throughout, we use subscripts DM and A for the decision maker and adviser, respectively. 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 $\omega|s \sim N\left(\rho s + (1-\rho)\mu_i, \tilde{\sigma}^2\right)$,

[6] Our formulation follows, for example, Shin (1998). Alternatively, we could assume that the adviser must make an assertion that the signal lies in some compact set, $\mathcal{S}$, or $\mathbb{R}$, with the only constraint that $s \in \mathcal{S}$, as formulated by Milgrom (1981). In this case, when the signal is not observed, the adviser has to report $\mathbb{R}$. By endowing the DM with a "skeptical posture" (Milgrom and Roberts, 1986) when the adviser claims any set $\mathcal{S} \neq \mathbb{R}$, our analysis can be extended to this setting.

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 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 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 $\sigma_0^2 \to \infty$).

3 Equilibrium Disclosure Behavior

In this section, we analyze behavior of adviser and DM in the disclosure sub-game. For this purpose, it will be sufficient to focus on the interim bias $B(\mu)$ of the adviser and the probability $p < 1$ with which he observes a signal.[8] Hence, we take the pair (B, p) with $p < 1$ as a primitive parameter in this section. Our objective is to characterize the set $S \in \mathbb{R}$ of signals that the adviser (with (B, p)) withholds and the action $a_\emptyset$ the DM chooses when there is no disclosure. Obviously, 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

[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 $\rho = 0$.

[8] We can restrict attention to $p < 1$ because, by the Inada conditions, the adviser will never choose p=1.

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 for the DM (ρs) and the action most preferred for 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.

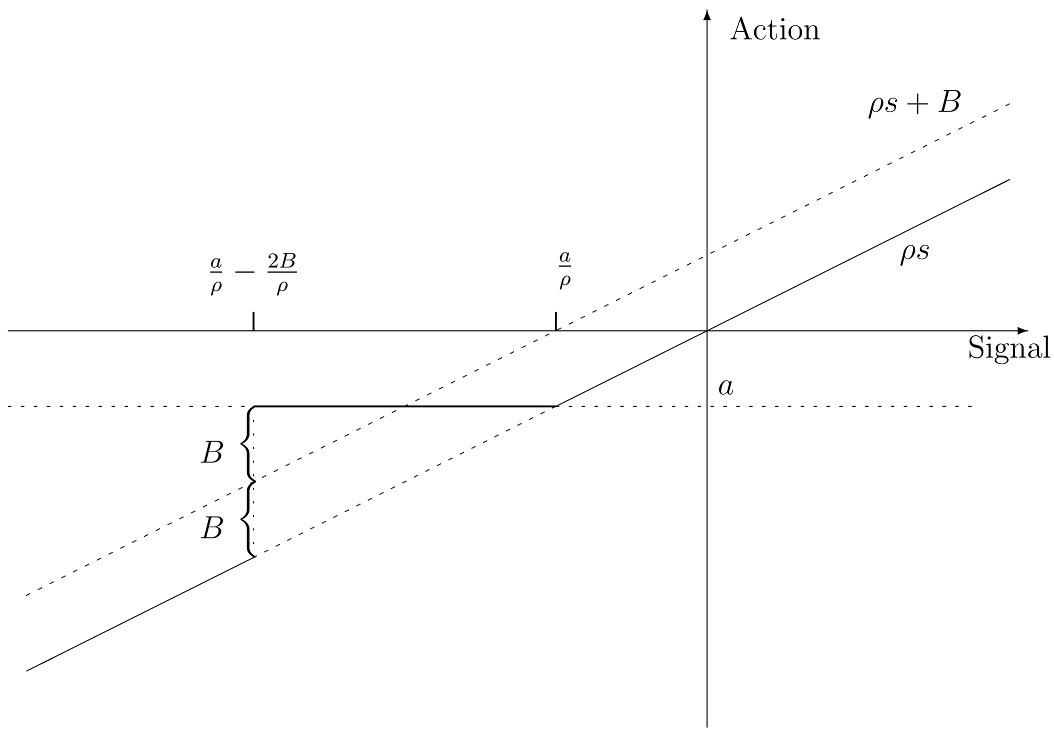

Figure 1: Optimal non-disclosure region

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

$$h(a) = \frac{a}{\rho}, \quad (2)$$

$$l(B, a) = h(a) - \frac{2B}{\rho}. \quad (3)$$

At $s = h(a)$, the DM will choose $a = \alpha(h(a)|0)$ whether s is disclosed or not, so the adviser is indifferent. At $s = l(a)$, the adviser is again indifferent between disclosure, which leads to $\alpha\left(l(a)|0\right)=a-2B$, 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.[9]

Next, we characterize the DM's best response in terms of her nondisclosure action, for an arbitrary (measurable) set $S \subset \mathbb{R}$ 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 nondisclosure, which is computed via Bayes rule:

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

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 as S gets large in the strong set order.[10] 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:

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

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. (Disclosure Equilibrium) *For any (B,p) with $p<1$, 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 $p=0$, and is strictly decreasing in B for $p>0$.*

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

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

[10] 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 $\min\{s,s'\}\in S'$.

$p > 0$), 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. 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 the DM. We call this the *prejudicial effect*, since a more upward biased adviser is in essence punished with a lower inference when he claims not to have observed a signal.[11] 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.

Finally, it is worth emphasizing that the adviser's optimal disclosure behavior only depends directly on his interim bias, B, and the DM's nondisclosure action, $a_{\emptyset}$. In particular, it does not depend directly on the probability of acquiring a signal, p, although p does directly affect the DM's optimal nondisclosure action, and thereby has an indirect effect in equilibrium on the adviser's disclosure choice. An implication that we exploit in the sequel is that what determines play in the disclosure sub-game is not actually the true probability of acquiring a signal, but rather the DM's belief about this probability. Put differently, the p we have taken as a primitive in this section need not be the probability with which the adviser acquires a signal; it is instead the probability that the DM ascribes to him getting a signal.

4 Opinions as Incentives

We now turn to the implications of the adviser's type on his incentive to acquire information, using the characterization of disclosure behavior from the previous section. Thereafter, we identify the optimal type of adviser for the DM in light of these implications. It is useful

[11] As an application, consider the inference that would have been drawn about Bill Clinton's likelihood of guilt in the Monica Lewinksy scandal should special counsel Ken Starr have not turned up any evidence (Starr is generally perceived to be Right-leaning); contrast this with the public's inference had a Left-leaning special counsel claimed the same.

to establish as a benchmark the fairly obvious point that, absent information acquisition concerns, the optimal adviser is a like-minded one.

PROPOSITION 2. (EXOGENOUS EFFORT) *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 DM is strictly better off under full disclosure than partial disclosure. ▌

To begin the analysis of endogenous information acquisition, suppose the DM believes that an adviser with $\mu \geq 0$, with induced interim bias $B(\mu)$, will choose effort p^e. The following Lemma decomposes the payoff for the adviser from choosing effort p, denoted $U_1(p; p^e, B, \mu)$, in a useful manner.[12]

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

$$U_1(p; p^e, B, \mu) = \Delta(B, \mu, p^e) + pA(B, \mu, p^e) - c(p),$$

where

$$A(B, \mu, p^e) := \left(\int_{s \notin S(B,p^e)} \left[\left(a_\emptyset\left(B, p^e\right) - (\rho s + B)\right)^2 - B^2 \right] \gamma\left(s; \mu\right) ds \right) \tag{6}$$

and

$$\Delta(B, \mu, p^e) := -\int \left(a_\emptyset\left(B, p^e\right) - (\rho s + B)\right)^2 \gamma\left(s; \mu\right) ds - \tilde{\sigma}^2. \tag{7}$$

The first term in the decomposition, $\Delta(\cdot)$, is the expected utility when a signal is not observed. Equation (7) expresses this utility by iterating expectations over each possible value of s, reflecting the fact that the DM takes decision $a_\emptyset(\cdot)$ 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, $pA(\cdot)$, is the

[12]Even though the interim bias B is determined by μ, we write them as separate variables in the funcion $U_1(\cdot)$ 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.

probability of a obtaining a signal multiplied by the expected gain from obtaining a signal. Equation (6) expresses the expected gain, $A(\cdot)$, 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 decision 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 $U_1(p; p^e, B, \mu)$. By the Inada conditions on effort costs, this choice is interior and characterized by the first-order condition:

$$\frac{\partial U_1(p; p^e, B, \mu)}{\partial p} = A(B, \mu, p^e) - c'(p) = 0.$$

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

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

Lemma 2. *For any (B, μ), p is an equilibrium effort choice if and only if $p \in (0, 1)$ and satisfies (8). For any (B, μ), an equilibrium exists.*

We cannot rule out that there may be multiple equilibrium effort levels for a given type of adviser. The reason is that the DM's action in the event to nondisclosure depends on adviser's (expected) effort, and the adviser's equilibrium effort in turn depends upon the DM's action upon nondisclosure.[13] For the remainder of the paper, for each (B, μ), we focus on the highest equilibrium effort, denoted $\bar{p}(B, \mu)$. This is somewhat analogous to the standard practice of allowing the DM to select the equilibrium she prefers in principal-agent models. 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, $p(\mu) := \bar{p}(B(\mu), \mu)$. The following is a central result of the paper.

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

To elucidate the intuition behind the result, first ignore the strategic disclosure of information, assuming instead that the outcome of the adviser's investigation is publicly

[13]Formally, multiplicity emerges when the function $A(B, \mu, \cdot)$ 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 $A(B, \mu, \cdot)$ is constant.

observed. In this case, the DM simply takes action ρs when signal s is observed, and action 0 if no signal is observed, i.e. $a_\emptyset(B,p) = 0$ independent of B or p. It follows from a usual 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 $A^{pub}(\mu)$, is given by[14]

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

Acquiring information benefits the adviser by reducing uncertainty, as shown by the first term. More importantly, it enables the adviser 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, $B(\mu) = (1-\rho)\mu < \mu$.[15] 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. This leads to such an adviser exerting more effort towards information acquisition.

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 he 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 (6), is at least as large as the marginal benefit from (sub-optimally) disclosing all signals, which we shall denote $A^{pri}(\mu, a_\emptyset(B(\mu), p^e))$. By mean-variance decomposition again, we have

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

Recall from Proposition 1 the prejudicial effect: for any $p^e > 0$ and $\mu \neq 0$, $a_\emptyset(B(\mu), p^e)\mu < 0$. Hence, for any $p^e > 0$ and $\mu \neq 0$, $A^{pri}(\mu, p^e) > A^{pub}(\mu)$: given that information is private, the DM's rational response to the adviser claiming a lack of information affects the

[14] Alternatively, one can also verify that equation (6) simplifies to equation (9) if the nondisclosure region $S(\cdot) = \emptyset$ and $a_\emptyset(\cdot) = 0$, as is effectively the case under public observation of signal.

[15] Equivalently, the adviser expects action $\rho\mu$ conditional on acquiring information, whereas he knows that action 0 will be taken without information, so the adviser believes that by acquiring information, he can convince the DM to take an action that is closer in expectation to his own prior.

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.[16]

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. (Optimality of Difference of Opinion) *There exists some $\mu_A \neq 0$ such that it is strictly better for the DM to appoint an adviser of type μ_A over a like-minded adviser.*

The optimality of difference of opinion is largely due to the persuasion effect. As the difference of opinion μ is raised slightly, the persuasion motive it generates creates a non-negligible benefit in increased information acquisition, whereas the prejudicial effect (which entails both communication loss and information acquisition gain) is negligible. This can be seen most clearly when the signal is perfectly informative, $\rho = 1$. In this case, $B(\mu) = 0$, so there is full disclosure in the communication stage, analogous to a situation where information is public; hence, appointing an advisor with difference of opinion is clearly desirable.[17] By continuity, there is a set of ρ's near 1 for which the adviser of type μ is better for the DM than the 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, there is some adviser sufficiently near type 0 who is in fact better for the DM than an adviser of type 0.

[16] The proof of Proposition 3 in the Appendix shows that the incentive to avoid prejudice combined with the incentive to persuade leads an adviser with a greater difference of opinion to exert more effort, taking into account the greater strategic suppression of signals by such an adviser.

[17] Formally, the DM's utility from appointing an adviser of type μ when $\rho = 1$ is

$$U_{DM}^{\rho=1}(\mu) := -\sigma_0^2(1 - p(\mu)),$$

which is simply the ex-ante variance in the state multiplied by the probability of not acquiring a signal, because when a signal is acquired, there is no residual uncertainty. Proposition 3 implies that for any $\mu > 0$, $U_{DM}^{\rho=1}(\mu) > U_{DM}^{\rho=1}(0)$.

5 Opinions and Preferences

So far, we have assumed that the DM and the available pool of advisors all have the same fundamental preferences, but differ in opinions. As noted, at a substantive level, this seems reasonable in some circumstances. At a theoretical level, this approach may seem without loss of generality, based on existing models of communication/disclosure games. In such models, it is often irrelevant whether the adviser's bias originates from his preference or his prior: for an adviser with any prior, there exists an adviser with a corresponding preference who will have precisely the same incentive for disclosure/communication, and vice versa.

We show here that this interchangeability of priors and preferences does not extend to a model with endogenous information acquisition. The distinction then opens up new issues. 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 the state of the world will shift the DM's action in the direction of his preferences bias—or a *zealot*—one who believes that his preference will be also be "vindicated by the evidence." In this section, we explore these issues.

To begin, suppose, as is standard in the literature, a player's preferences are indexed by a single bias parameter $b \in \mathbb{R}$, 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), $(b, \mu) \in \mathbb{R}^2$. 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 subgame. For any adviser with opinion bias μ and no preference bias, there exists an 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.[18] 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. Hence, the two types have the same expected utility, absent disclosure. 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 that provides incentives for an opinion-biased adviser does not exist for an adviser biased in preferences alone.*

Publicly observed signal. To see how the two types of biases affect the incentive, it is useful to first consider the case where the adviser's signal (or lack thereof) is publicly observed. This case allows us focus on some of the issues more simply, 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, even though the adviser prefer action $\rho s + B(b, \mu)$; so the adviser has expected utility $-\tilde{\sigma}^2 - (B(b, \mu))^2$. The adviser's expected gain from acquiring information is, therefore,

$$A^{pub}(b, \mu) = \underbrace{\sigma_0^2 - \tilde{\sigma}^2}_{\text{uncertainty reduction}} + \underbrace{\left(2\rho - \rho^2\right) \mu^2}_{\text{persuasion}} + \underbrace{(1 + \rho)\, b\mu}_{\text{reinforcement}}. \tag{11}$$

Suppose first $\mu = 0$, so the adviser is like-minded. In this case, $A^{pub}(b, 0)$ 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 to appointing an adviser who differs only in preference. This stands in stark contrast to the case of difference of opinion, $(0, \mu)$, $\mu \neq 0$, 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

[18] "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.

opinions.

Now suppose $\mu \neq 0$. Then, the persuasion effect appears, 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 manifests itself as the third term 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 offset 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 the adviser's incentive for information acquisition. The next proposition states an incentivizing 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* $B(b',\mu')\mu' \geq 0$.[19] *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 the public information case, a preference bias alone creates incentives for information acquisition when it is private. The adviser exerts effort to avoid the prejudicial inference the DM attaches to nondisclosure. Of course, this incentive benefit is offset by the loss associated with strategic withholding. It turns out that these opposing effects are of the same magnitude locally for $|b| \approx 0$. Hence, a difference of preference is not unambiguously beneficial to DM in the way the difference of opinion is.[20]

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

[20] Indeed, a numerical analysis shows a possibility that the DM's utility falls as the $|b|$ rises from zero but rises above the initial level as $|b|$ becomes sufficiently large. In this case, unlike figure ?? (on the case of opinion bias), DM never prefers an adviser with preference bias unless it is sufficiently large. This difference matters when advisers with large biases (of either type) are difficult to find or (politically) infeasible to select.

More generally, the proposition shows how the two types of biases interact with respect to the incentive for acquisition. The following corollaries record the nature of interaction.

Corollary 1. (Complementarity of opinion and preference) *If* $(b', \mu') > (b, \mu) \geq 0$*, then an adviser with* (b', μ') *choose 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. (Zealot vs. skeptic) *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. (Optimal type) *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. In practice, we suspect there must be an upper bound on μ. Faced with such an upper bound, it may well be optimal for the DM to choose an expert with an interim bias $B(b, \mu) > 0$, as was the case when advisers are differentiated by opinions alone.

6 Discussion

In this section, we discuss some other issues that can be raised in our framework. For simplicity, we return to the baseline setting where advisers are only distinguished by their opinions, sharing the fundamental preferences of the DM.

6.1 Delegation

An important issue in various organizations is the choice between delegation and communication.[21] One prominent view is that delegation of decision-making authority to an agent increases his incentives to become informed about the decision problem if he can extract a greater fraction of the surplus from information gathering, but is costly to the DM insofar as the agent's actions when given authority do not maximize the DM's interests. Aghion and Tirole (1997) label this the "initiative" versus "loss of control" tradeoff.

Our model generates a complementary insight: aside from any loss of control, delegation can also lead to to reduced initiative because it eliminates concerns the agent may have with a DM's decision in the absence of information. The intuition is most transparent when considering the limiting case of perfectly-informative signals, i.e. $\sigma_1^2 = 0$ or $\rho = 1$. Then, $B(\mu) = 0$ and there is no interim bias when a signal is observed, regardless of the adviser's prior. As previously noted, communication then perfectly reveals the adviser's information (including the event that he doesn't receive a signal), regardless of any difference of opinion. What happens if we now permit the DM to delegate decision-making authority to the chosen adviser as an alternative to communication? That is, the DM chooses the adviser and commits to either a regime of communication or full delegation prior to the adviser acquiring information.

First consider the issue of loss of control. If the chosen adviser is like-minded, delegation and communication are equivalent, because the DM and the adviser would take the same decision regardless of the outcome of the adviser's investigation. On the other hand, if the adviser has a difference of opinion, the decisions under delegation and communication would coincide when a signal is received, but not when no signal is received. Hence, there

[21] Our interest here concerns situations where there is an information acquisition problem, but Dessein (2002) shows that the question is subtle even when information is exogenous. Also, we only consider the two extremes: full delegation and no delegation, ignoring constrained delegation. A standard justification is that of incomplete contracting. Szalay (2005) analyzes a related mechanism design problem with commitment, but our setting differs because of verifiable information and because differences of opinion induce interim bias.

is a cost of loss of control if and only if the chosen adviser is not like-minded. It is also clear that, holding effort fixed, the cost to the DM is strictly increasing in the difference of opinion the adviser has.

Now consider the issue of incentives to acquire information. Under the communication regime, since information is fully revealed, and preferences agree when a signal is observed, the incentives to acquire information are driven by the disagreement between adviser and DM in the event the investigation reveals no signal. This is strictly increasing in the difference of opinion between the DM and the adviser, i.e. more extreme advisers (relative to the DM) have greater incentives to acquire information. In contrast, under delegation, the adviser's incentives to acquire information do not vary at all with his prior. This is because there is no penalty for not acquiring information, since the adviser will simply take his ex-ante preferred decision in such an event. Therefore, under delegation, there is no benefit to appointing an adviser with a difference of opinion.

By our earlier results, the discussion above can be extended to the case of partially-informative signals, $\rho < 1$, summarized as follows.

Proposition 6. *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 appropriate adviser with a difference of opinion.*

Proof. The first statement follows from the preceding discussion. The second is a corollary of Proposition 4 and the observation that delegation and communication have identical consequences when the adviser is like-minded. ∎

The point can also be seen through the decomposition of incentives we discussed after Proposition 3. When the authority is delegated, the adviser no longer has the incentive to acquire information to persuade the DM even when his opinion is different, since he views his opinion to be right and he has the decision-making authority. In contrast, when the DM retains authority, the advisor has an incentive to persuade the DM *if* he has a difference of opinion, because he believes that information will bring the DM's decision closer to his own preferred decision. It is this persuasion incentive that is eliminated by delegation.[22]

[22] Recall that when signals are not perfectly revealing of the state, so that there is non-zero interim bias given a difference of opinion, there is also a third component of incentives: avoiding prejudice. This too is eliminated by delegation.

6.2 Heterogeneity in Confidence

Suppose that rather than being differentiated by prior mean, all advisers have the same prior mean as the DM ($\mu = 0$), but they differ in confidence in their ability, represented by beliefs about their signal precision. It is convenient to map signal precision into the weight the posterior mean place on signal versus prior mean, as before. That is, if an adviser believes his signal has variance σ_1^2, we represent him via $\rho(\sigma_1^2) := \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}$.[23] The DM has $\rho_{DM} \in (0,1)$, and she can choose an adviser with any $\rho_A \in [0,1]$. Interpreting ρ_{DM} as a baseline, an adviser with $\rho_A > \rho_{DM}$ is overconfident of his ability and an adviser with $\rho_A < \rho_{DM}$ is underconfident.[24]

The preferred action for an adviser of type ρ_A who observes signal s is $\rho_A s$, whereas the DM would take action $\rho_{DM} s$ if the signal is disclosed. Hence, any adviser with $\rho_A \neq \rho_{DM}$ has a conflict of interest in the disclosure sub-game. Surprisingly however, so long as $\rho_A \geq \rho_{DM}$, there is an equilibrium in the disclosure sub-game that fully reveals the outcome of the adviser's investigation, for any effort choice. To see this, notice that if the DM believes the adviser never withholds his signal, he optimally plays $a_\emptyset = 0$. It is then optimal for the adviser to disclose any signal he acquires because $|\rho_A s| \geq |\rho_{DM} s|$ for all s (with strict inequality when $s \neq 0$ and $\rho_A > \rho_{DM}$).[25]

We can show that the incentives to acquire information are strictly higher the more overconfident an adviser is. The intuition is that he believes the value of a signal is larger, and hence perceives a larger marginal benefit of effort, even after accounting for the fact that the DM's decision does not respond to a signal as much as he would like. Therefore, among overconfident advisers, the DM strictly prefers one with $\rho_A = 1$. In the proof of the following result, we establish that an underconfident adviser is never optimal either.

Proposition 7. *If advisers are distinguished only by confidence, ρ_A, the DM uniquely prefers to appoint a maximally overconfident adviser, i.e. one with $\rho_A = 1$ who believes that his signal is perfectly informative.*

To conclude our discussion of confidence, suppose that advisers can differ in both their

[23]This makes it clear that the analysis can also be applied to differences in beliefs about the variance of the prior distribution of state of the world.

[24]Admati and Pfleiderer (2004) and Kawamura (2007) analyze cheap talk games with an over/underconfident Sender. Information acquisition is not endogenous in either paper.

[25]If there is any other equilibrium in the disclosure sub-game, we assume the fully revealing one is played. Aside from being intuitively focal and simple, it is the only disclosure equilibrium that is "renegotiation proof" at the point where effort has been exerted but investigation outcome not yet observed (regardless of how much effort has/is believed to have been exerted).

opinion (i.e, prior mean) and perception of ability. When $\rho_A = 1$, the adviser's prior mean does not affect his posterior when he receives a signal. On the other hand, when he does not receive a signal, his preferred decision is his prior mean. Once again, there is a fully revealing equilibrium in the disclosure sub-game (regardless of the adviser's prior mean or effort), meaning that when no information is disclosed, the DM takes action $0 = \mu_{DM}$. So among advisers with $\rho_A = 1$, those with larger $|\mu_A|$ have even larger incentives to acquire information because of the penalty from the nondisclosure action when no signal is received. Hence, the DM can exploit a combination of difference of opinion and overconfidence.

7 Related Literature

Formally, our model builds upon a Sender-Receiver costless signaling game of verifiable information, first studied by Grossman (1981) and Milgrom (1981). The "unraveling" phenomenon noted by these authors does not arise here because the Sender (adviser) may not possess information, as pointed out by Shin (1994, 1998), among others. This literature typically assumes that independent of his information, the Sender's utility is strictly monotone in the Receiver's decision. In contrast, at the communication stage of our model, the interim preferences—owing to difference of opinion—take the non-monotone form of "finite bias" used in the cheap talk (soft or unverifiable information) literature following Crawford and Sobel (1982).

The idea that differences of opinion induce interim conflicts, in turn entailing communication loss, is not surprising. In most models of strategic communication with exogenous information, if the DM were permitted to choose the type of adviser she communicates with, she would choose an adviser who shares her interim preferences (Suen, 2004, explictly discusses this point).[26] Our point of departure is to endogenize the acquisition of information. While previous authors have incorporated this factor (e.g. Matthews and Postlewaite, 1985; Shavell, 1994), there is little work studying the tradeoff between information acquisition and transmission, and how this is affected by the degree of conflict, which is the insight developed here.

A few exceptions are authors who have recently developed similar themes independently. Dur and Swank (2005) find it optimal for a DM to sometimes choose an adviser of a different preference type from herself, even though communication with such an adviser is

[26]This is also true in models with multiple advisers, such as Krishna and Morgan (2001) and Li and Suen (2004).

less effective than with an adviser of the same type. In a related vein, Gerardi and Yariv (2007a) show, in a jury model with binary decisions and public information, that appointing a juror with a different preference (in terms of the threshold of reasonable doubt) from the DM can be beneficial. Despite the resemblance of these findings to ours, the underlying rationales are distinct, formally due to the binary decision space in the aforementioned papers. The main idea of Dur and Swank (2005) is the benefit of having an adviser who is more open-minded or less extreme than the DM, *rather than one with difference of opinion from DM*—the focus of the current work. Indeed, if the DM in their model is ex-ante unbiased between the two available decision alternatives, it is optimal for her to appoint an adviser of the same type; in our continuous model, this is never the case. Likewise, the incentive effect of a preference-biased juror in Gerardi and Yariv (2007a) stems from the binary decision space: facing the unfavorable default status quo, a biased advisor can only gain from the new information which will either reinforce the status quo or, with luck, shift the DM's decision to the favored alternative.[27] With non-binary decisions, however, new information can lead to a decision that is less favorable than the status quo. In fact, without a restriction on the decision space, new information does not shift the decision on average when the conflict is one of preferences alone. Consequently, as was shown in Section 5, a preference conflict alone with public information cannot motivate information acquisition when the decision is continuous. As such, the persuasion and prejudicial effects we find that create incentives for information acquisition have no counterparts in these papers.

Van den Steen (2004) does find a persuasion incentive stemming from difference of opinion.[28] His model differs from ours, however, in a number of ways; for example he assumes both binary information and decisions. Substantively, unlike in our paper, he does not study the prejudicial effect and its implications for information acquisition when information is privately observed by the adviser. There is, therefore, no analog of our finding that some difference of opinion is optimal for the DM even accounting for the cost of strategic disclosure. The above papers and ours are best seen complementary, developing related but distinct insights, and focusing on different issues.[29]

[27] A related point is made by Hori (2006) 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 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).

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

[29] For instance, Van den Steen (2004) also discusses coordination between individual action choices, and

The optimality of having an adviser whose interim preferences are different from the DM's is also reminiscent of Dewatripont and Tirole (1999), but the frameworks and forces at work are quite different. They 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.

Finally, the current paper is consistent with the emerging theme that ex-post suboptimal mechanisms are sometimes ex-ante optimal because they can provide greater incentives for agents to acquire information. Examples in settings related to ours are Li (2001) and Szalay (2005).[30] Although we are interested in environments where contracting is essentially infeasible, our results have some of this flavor. If the DM chooses a like-minded adviser, the decision she takes will be optimal given the adviser's information, because information is fully revealed. In choosing an adviser of a different opinion, the DM's decision is suboptimal in some cases given the adviser's information, because of the prejudicial effect when the adviser does not acquire a signal; this leads to greater information acquisition from the adviser. It is important to note, however, that ex-post suboptimality is not the only reason that it is ex-ante optimal for the DM to appoint an adviser with a difference of opinion: even if signals are public when acquired—in which case the DM always takes an optimal decision given the adviser's information—an adviser who has a larger difference of opinion with the DM has a greater incentive to acquire information, because of the persuasion incentive.

8 Concluding Remarks

This paper has introduced a novel model to study the costs and benefits of difference of opinion from an incentive point of view. Our main findings are threefold. First, because difference of opinion leads to interim conflict of interest even when fundamental preferences agree, it leads to a loss in information transmission via strategic disclosure. This leads to a prejudicial effect against an adviser who has a difference of opinion unless he provides

Gerardi and Yariv (2007a) consider multiple jurors and sequential consultation. Meanwhile, we consider private information and the distinction/interaction between opinion and preferences. We should also note the less directly related work of Cai (2004), who develops a model with an endogenously chosen number of advisers who strategically acquire and transmit information. In his environment advisers are assumed to be ex-ante unbiased, instead only discovering their biases if they choose to acquire information about the decision problem.

[30] Examples in other environments include Bergemann and Valimaki (2002), Gerardi and Yariv (2007b), and Gershkov and Szentes (2004).

information. Second, difference of opinion increases the incentives for an adviser to exert costly effort towards information acquisition. This is for two reasons: a motive to avoid the prejudicial effect just noted if he does not provide information and a belief that he can systematically shift the decision maker's action in the direction of his opinion by acquiring information (the persuasion motive). Third, in trading off these costs and benefits, the decision marker finds it optimal to consult an adviser who has at least some difference of opinion with her.

We have discussed various other issues, including the relationship between opinions and preferences, their interaction, and questions about delegation and confidence. We conclude by discussing the significance of some of our modeling choices, and some avenues for future research.

The nature of effort. We have assumed that effort is success-enhancing in the sense of increasing the probability of observing an informative signal about the state. Alternatively, one could assume that there is a fixed probability with which the adviser observes a signal, but higher effort increases the precision of the observed signal. We believe our insights would extend to this setting. The incentivizing effect of difference of opinion and non-optimality of a like-minded expert certainly hold when the signal is publicly observed, because in this case there is no disclosure loss, and an adviser with a more extreme opinion has more to gain by producing a more precise signal.

Social Welfare. Our focus has been entirely on the DM's optimal choice of adviser. One could, instead, discuss the issue of social welfare in our model. It is clear how to evaluate welfare under full information about the state, since all individuals have the same fundamental preferences. More difficult, however, is the question of what the "right" prior is for a social planner (SP), and also how cost of effort from the adviser should be incorporated. We allow for various possibilities under a single umbrella, by considering a social planner who has his own prior mean, μ_{SP}, and the following social welfare function:

$$u_{SP}(a, \omega, e) = -k(a - \omega)^2 - c(e),$$

where $k \geq 1$ is the weight placed on correct decision-making relative to effort cost. One can think of k as reflecting the number of people affected by the decision problem at hand. As $k \to \infty$, the effort cost is essentially ignored by the SP. This may be justified for example if the decision is about a new medicinal drug and the unknown state is the possible side effects the drug may have (cf. Persico, 2004). On the other hand, if the decision affects only a small number of people—perhaps just the DM and adviser—then the effort cost plays a significant role in the social calculus.

We can ask what kind of an adviser the SP would choose if the adviser were communicating directly with the SP, and how this compares with the choice of adviser by the DM who does not account for the cost of effort. First consider $k = 1$. Then, it is straightforward to see that if $\mu_{SP} = 0$, so that society and DM share the same prior, the socially optimal adviser is a like-minded one, and the DM's choice is inefficient. The logic is simple: a like-minded adviser reveals information perfectly, and chooses effort to solve the first best when $k = 1$. However, this obviously relies upon the adviser fully internalizing the tradeoff between effort and good decisions. If $k > 1$, the optimal choice of adviser would be one with a different opinion, but generally not as extreme as the DM would choose. We expect that as k moves from 1 to ∞, the socially optimal choice of adviser ranges from a like-minded one to the DM's choice. Finally, note that if $\mu_{SP} \neq 0$, so that the planner has a different prior from the DM, the DM's choice of adviser will not generally coincide with the planner's choice.

Monetary payments. We have completely abstracted from monetary contracting altogether. In many applications, this seems reasonable, for example if a leader is choosing amongst current members of her organization to investigate some issue, or the wage for an outside adviser is exogenously fixed as far as the DM is concerned, for bureaucratic or other reasons, and richer contracts are not feasible. A first step towards relaxing this would be require the DM to promise a wage that will induce the adviser's ex ante participation. While the full extension is beyond the scope of current work, we conjecture that our main insights carry over to this setting. Intuitively, the DM in this case maximizes some weighted average of her payoff and the adviser's payoff, with the weight on the adviser's payoff corresponding to the shadow value on the individual rationality constraint. This is similar to the treatment of "social welfare" above, with a planner who shares the DM's prior.

Extremism and bounded action space. The set of feasible decisions in our model is $\mathbb{R}$. If, instead, the decision space were restricted, to, say, a compact interval symmetric around 0, we would have an additional effect: a more extreme adviser (in terms of opinion or preference) would place less direct value on acquiring information than a less extreme adviser (cf. Dur and Swank, 2005). Intuitively, this is because with a restricted decision space, it is unlikely that information will "change the mind" of an extreme adviser. This would introduce another element to the calculus of difference of opinion that may be worthy of further examination. It seems likely, however, that while this would generally make the optimal adviser less extreme that in our current setting (since less extreme advisers would have greater direct interest in acquiring information), our central insights would continue to hold.

The nature of information. We have treated information as hard or verifiable. Aside from the technical tractability this offers, the assumption seems appropriate in numerous relevant applications. On the other hand, there are some situations where information may be soft or unverifiable. Although a full analysis is left to future research, we can make some observations. It is important to distinguish between two possibilities here.

First, as in Austen-Smith (1994), suppose that the adviser can prove that he has acquired a signal, but statements about what the signal is are cheap talk. Our main insights appear to extend to this setting, subject to the caveat of multiple equilibria in the communication stage owing to the usual difficulties with cheap talk. In particular, when signals are perfectly revealing of the state, there is an equilibrium in the communication game where the adviser (regardless of his opinion) reveals whether he is informed or not, and if informed, perfectly reveals his signal. Loosely, this is the "most informative" communication equilibrium. Given its selection, a greater difference of opinion leads to greater information acquisition, because of the persuasion motive, while there is no information loss from communication. By continuity, some difference of opinion will be optimal when signals are close to perfect, even though there is a loss of information from strategic communication in this case.

On the other hand, if the adviser cannot even prove that he has acquired a signal, so that all statements are cheap talk, the incentivizing effects of differences of opinion may be significantly mitigated. Intuitively, the persuasion motive from difference of opinion creates incentives because the adviser believes that by exerting more effort, he creates the ability to systematically shift the the DM's action in his desired direction. Under complete cheap talk, effort does not affect the set of claims the adviser can make, so this component of incentives appears tenuous in such cases.

A Appendix: Proofs

Proof of Proposition 1: By assumption, $p < 1$. Without loss, we assume $B \geq 0$. A symmetric argument will 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

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

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

$$\overline{s}(B,p) = \frac{p}{p\int_{\overline{s}(B,p)-\frac{2B}{\rho}}^{\overline{s}(B,p)} \gamma(s;0)\,ds + 1 - p} \int_{\overline{s}(B,p)-\frac{2B}{\rho}}^{\overline{s}(B,p)} s\gamma(s;0)\,ds. \tag{12}$$

We will show that there is a unique solution to (12) in two steps below; this implies that 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 (12) when $B > 0$. This latter is accomplished by showing that there is a unique solution to

$$\Upsilon(\overline{s}; B,p) := -p\int_{\overline{s}-\frac{2B}{\rho}}^{\overline{s}} (\overline{s}-s)\gamma(s;0)\,ds - \overline{s}(1-p) = 0. \tag{13}$$

Without loss of generality, we can restrict attention to $\overline{s} < 0$, because there is no solution to (13) 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 in (13) is positive and bounded above by $4(\frac{B}{\rho})^2$, and $p < 1$).

To prove uniqueness, observe

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

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 (13). ‖

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

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

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,

$$\frac{\partial \overline{s}\left(B, p\right)}{\partial B} = -\frac{\frac{\partial \Upsilon\left(\overline{s}(B,p); B,p\right)}{\partial B}}{\frac{\partial \Upsilon\left(\overline{s}(B,p); B,p\right)}{\partial \overline{s}}} < (=)0,$$

(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, $\frac{\partial \overline{s}(B,p)}{\partial p} = -\frac{\frac{\partial \Upsilon\left(\overline{s}(B,p); B,p\right)}{\partial p}}{\frac{\partial \Upsilon\left(\overline{s}(B,p); B,p\right)}{\partial \overline{s}}}$. 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 $p=0$, and is strictly decreasing in B for $p>0$.*

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:

$$U_1(p;p^e,B,\mu) = p\Big[\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_1(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right] + \mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_1(a_\emptyset(B,p^e),\omega)|s,\mu]\mid\mu\right]\Big] \\ +(1-p)\mathbb{E}_\omega[u_1(a_\emptyset(B,p^e),\omega)|\mu] - c(p).$$

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:

$$\begin{aligned}
& U_1(p;p^e,B,\mu) \\
= \;& p\Big[\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_1(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right] + \mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_1(a_\emptyset(B;p^e),\omega)|s,\mu]\mid\mu\right]\Big] \\
& \quad +(1-p)\mathbb{E}_\omega[u_1(a_\emptyset(B;p^e),\omega)|\mu] - c(p) \\
= \;& p\left(\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_1(\alpha_0(s),\omega)\mid s,\mu]\mid\mu\right] + \mathbb{E}_{s\in S(B,p^e)}\left[\mathbb{E}_\omega[u_1(a_\emptyset(B;p^e),\omega)|s,\mu]\mid\mu\right)\right. \\
& \quad +(1-p)\mathbb{E}_s[\mathbb{E}_\omega[u_1(a_\emptyset(B;p^e),\omega)|s,\mu]] - c(p) \\
= \;& p\left(\mathbb{E}_{s\notin S(B,p^e)}\left[\mathbb{E}_\omega[u_1(\alpha_0(s),\omega) - u_1(a_\emptyset(B;p^e),\omega)\mid s,\mu]\mid\mu\right]\right) \\
& \quad +\mathbb{E}_s\left[\mathbb{E}_\omega[u_1(a_\emptyset(B;p^e),\omega)|s,\mu]\mid\mu\right] - c(p) \\
= \;& p\left(\int_{s\notin S(B,p^e)}\left[\left(a_\emptyset\left(B,p^e\right)-\rho s-B\right)^2-B^2\right]\gamma\left(s;\mu\right)ds\right) \\
& \quad -\int\left(a_\emptyset\left(B,p^e\right)-\left(\rho s+B\right)\right)^2\gamma\left(s;\mu\right)ds-\tilde{\sigma}^2-c(p),
\end{aligned}$$

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

Proof of Lemma 2: By Proposition 1, there is full disclosure when $p^e=0$, hence we evaluate $A(B,\mu,0)=\sigma_0^2+2\rho B>0$ from (6). For any (B,μ), $A(B,\mu,\cdot)$ is a bounded mapping. Therefore, by the Inada conditions, we have $c'(0) = 0 < A(B, \mu, 0)$ and $c'(p) > A(B, \mu, 1)$ 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

$$\frac{\partial U_1(\tilde{p}; p, B, \mu)}{\partial p} = A(B, \mu, p) - c'(\tilde{p}) \gtreqless 0 \text{ if } \tilde{p} \lesseqgtr p,$$

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 prior μ. We can then 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.[31] This will mean 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:

$$\mathcal{A}(\mu, p(\mu)) = c'(p(\mu)), \tag{14}$$

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

where $\mathcal{A}(\mu, p) := A(B(\mu), \mu, p)$. Observe first

$$\begin{aligned}
\mathcal{A}(\mu, p) &= \int_{s \notin S(B(\mu),p)} \rho(s - \overline{s}(B(\mu), p)) \left(2(1-\rho)\mu + \rho(s - \overline{s}(B(\mu), p))\right) \gamma(s; \mu) ds \\
&= 2\mu\tilde{\sigma}^4 \int_{s \notin S(B(\mu),p)} (s - \overline{s}(B(\mu), p))\gamma(s; \mu) ds \\
&\qquad + \rho^2 \int_{s \notin S(B(\mu),p)} (s - \overline{s}(B(\mu), p))^2 \gamma(s; \mu) ds. \qquad (15)
\end{aligned}$$

A straightforward, but tedious, calculation yields

$$\mathcal{A}_{11}(0, p) > \mathcal{A}_1(0, p) = \mathcal{A}_2(0, p) = 0, \qquad (16)$$

where $\mathcal{A}_i$ denotes a partial derivative of $\mathcal{A}$ with respect to i-th variable, for $i = 1, 2$, and $\mathcal{A}_{ij}$ denotes a second order derivative with respect to i and j-th variables for $i, j = 1, 2$.

Totally differentiating (14) with respect to μ evaluating it at $\mu = 0$, we obtain

$$p'(0) = \frac{\mathcal{A}_1(0, p(0))}{c''(p(0)) - \mathcal{A}_2(0, p(0))}.$$

It then follows from (16) that $p'(0) = 0$. Now, totally differentiate (14) twice, and evaluate the outcome using $p'(0) = 0$. We then obtain

$$\begin{aligned}
p''(0) &= \frac{\mathcal{A}_{11}(0, p(0))}{c''(p(0)) - \mathcal{A}_2(0, p(0))} - \frac{(p'(0))^2 c'''(p(0))}{c''(p(0))} + \frac{p'(0)}{c''(p(0))} \left[2\mathcal{A}_{12}(0, p(0)) + \mathcal{A}_{22}(0, p(0))\right] \\
&= \frac{\mathcal{A}_{11}(0)}{c''(p(0))} > 0. \parallel
\end{aligned}$$

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,

$$w(0, p) - v(0, p) = \int_{-\infty}^{\infty} \left(a_{\emptyset}(0, p) - \frac{s\sigma_0^2}{\sigma^2}\right)^2 \gamma(s; 0) ds > 0,$$

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

$$v_1(0, p) = v_2(0, p) = v_{11}(0, p) = w_1(0, p) = w_2(0, p) = w_{11}(0, p) = 0. \qquad (17)$$

It then follows from (17) 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

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

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 advisor, 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 with (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 Maxima), there must exist $p'' > p(\mu)$ 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 $p' \leq p(\mu)$ given DM's belief that the adviser will choose $p(B, \mu)$. The disclosure subgame 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

$$A(B', \mu', p(B, \mu)) = c'(p') \leq c'(p(B, \mu)) = A(B, \mu, p(B, \mu)). \tag{18}$$

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

The proof follows several steps.

Step 1. *The following inequality holds.*

$$\begin{aligned} A(B', \mu', p(B, \mu)) &\geq \Pi(B', \mu') \\ &:= \int_{s \notin S(B)} \left[(a_\emptyset(B', p(B, \mu)) - \rho s - B')^2 - B'^2 \right] \gamma(s; \mu') ds. \end{aligned} \tag{19}$$

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

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

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

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

which implies the desired inequality by the definition of π. $\|$

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

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

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

We then obtain the desired inequality:

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

The first inequality follows from the fact that $(B',\mu') > (B,\mu)$ and that $B \geq 0$. That $B' \geq B \geq 0$ implies $0 \geq \overline{s} = \overline{s}(B) \geq \overline{s}(B') = \overline{s}'$ (Step 3 in the proof of Proposition 1), which in turn implies that $B' - \rho\overline{s}' \geq B - \rho\overline{s} \geq 0$. Next, $\mathbb{E}[s \mid s \notin S(B),\mu'] \geq \mathbb{E}[s \mid s \notin S(B),\mu]$, since the Normal density $\gamma(\cdot;\mu')$ dominates in likelihood ratio the Normal density $\gamma(\cdot;\mu)$. We also have $\mathbb{E}[s \mid s \notin S(B),\mu'] \geq 0$, since $\mu' \geq 0$ (which follows from the fact that $|\mu'| \geq |\mu|$ and that $\mu' \geq \mu$) and since $S(B) \subset \mathbb{R}_-$. Next, $\mu' \geq \mu$ implies that $\Pr\{s \notin S(B) \mid \mu'\} \geq \Pr\{s \notin S(B) \mid \mu\}$. Combining all these facts imply the first inequality in weak form. The inequality is strict, however, since $B' > B$ or $\mu' > \mu$, which means one of the inequalities established above must be strict.

The second inequality 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 since $|\mu'| \geq |\mu|$, 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 $s \in S(B) \subset \mathbb{R}_-$), and since $\Pr\{s \in S(B) \mid \mu'\} \leq \Pr\{s \in S(B) \mid \mu\}$. The string of inequalities thus proves the claim. ‖

Combining Step 1 and Step 2, we have

$$A(B',\mu',p(B,\mu)) > \Pi(B,\mu). \tag{20}$$

By definition, it also follows that

$$\Pi(B,\mu) = A(B,\mu,p(B,\mu)). \tag{21}$$

Combining (20) and (21) yields

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

which contradicts (18). 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 $(B',\mu') := (-B,\mu)$. ▮

Proof of Proposition 7: First consider $\rho_A \geq \frac{\rho_{DM}}{2}$. It is straightforward to verify that if $\rho_A \geq \frac{\rho_{DM}}{2}$, there is a full disclosure equilibrium in the disclosure sub-game, independent of effort, p. Given full disclosure, the gain in utility for the adviser from observing a signal s

over not observing it is

$$\int_{-\infty}^{\infty} (\rho_{DM} s - \omega)^2 \gamma(\omega|s,\rho_A) d\omega + \int_{-\infty}^{\infty} (-\omega)^2 \gamma(\omega|s,\rho_A),$$

where $\gamma(\cdot|s,\rho_A)$ denotes the posterior distribution over the state for the adviser with type ρ_A given signal s. The above can be simplified via algebra to

$$\rho_{DM} s^2 (2\rho_A - \rho_{DM}),$$

whose derivative with respect to ρ_A is $2\rho_P s^2$, which is strictly positive at all $s \neq 0$. Therefore, the marginal benefit of acquiring a signal is strictly higher for a more confident adviser, and consequently he exerts more effort. Since we have full disclosure in the disclose sub-game, it follows that the DM strictly prefers an adviser with $\rho_A = 1$ among all $\rho_A \geq \frac{\rho_{DM}}{2}$.

For $\rho_A < \frac{1}{2}\rho_{DM}$, it suffices to prove that there is a unique equilibrium in the disclosure sub-game, independent of effort, in which the adviser never discloses a signal. Consider any nondisclosure action $a_\emptyset \geq 0$ (the argument for $a_\emptyset < 0$ is symmetric to $a_\emptyset > 0$). It is straightforward to verify that the adviser's best response in terms of nondisclosure region is

$$S(a_\emptyset) = \left(-\infty, -\frac{a_\emptyset}{\rho_{DM} - 2\rho_A} \right] \cup \left[\frac{a_\emptyset}{\rho_{DM}}, \infty \right).$$

Note that this is no disclosure if (and only if) $a_\emptyset = 0$. It follows that $a_\emptyset = 0$ is an equilibrium independent of effort, p, because if the adviser is never disclosing, the DM will follow his prior upon nondisclosure. To see that there is no equilibrium with $a_\emptyset > 0$, observe that for any $a_\emptyset \geq 0$, the set $S(a_\emptyset)$ has expectation (with respect to the DM's prior density on signals) no greater than 0, and hence the DM's best response to $S(a_\emptyset)$, $a_N(p, S(a_\emptyset))$, is no larger than 0 for any p. Consequently, there does not exist $a_\emptyset > 0$ such that $a_\emptyset = a_N(p, S(a_\emptyset))$. ▮

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2007): "Learning and Disagreement in an Uncertain World," mimeo, Massachusetts Institute of Technology. 3

Admati, A. R., and P. Pfleiderer (2004): "Broadcasting Opinions with an Overconfident Sender," *International Economic Review*, 45(2), 467–498. 23

Aghion, P., and J. Tirole (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1–29. 6, 21

Aumann, R. J. (1976): "Agreeing to Disagee," *Annals of Statistics*, 4, 1236–1239. 3

Austen-Smith, D. (1994): "Strategic Transmission of Costly Information," *Econometrica*, 62(4), 955–63. 29

Banerjee, A., and R. Somanathan (2001): "A Simple Model Of Voice," *The Quarterly Journal of Economics*, 116(1), 189–227. 3, 7

Bergemann, D., and J. Valimaki (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70(3), 1007–1033. 26

Blackwell, D. (1951): "Comparison of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics*, pp. 93–102. 7

Cai, H. (2004): "Costly Participation and Heterogeneous Preferences in Informational Committees," mimeo, UCLA. 26

Crawford, V., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451. 6, 24

Degroot, M. H. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York. 9

Dessein, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69(4), 811–838. 21

Dewatripont, M., and J. Tirole (1999): "Advocates," *Journal of Political Economy*, 107(1), 1–39. 26

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. 24, 25, 29

ELIAZ, K., AND R. SPIEGLER (2006): "Contracting with Diversely Naive Agents," *Review of Economic Studies*, 71(3), 689–714. 7

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. MIT Press, Cambridge, MA. 8

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

GERARDI, D., AND L. YARIV (2007a): "Costly Expertise," mimeo, Yale University and Caltech. 25, 26

——— (2007b): "Information Acquisition in Committees," forthcoming, *Games and Economic Behavior*. 26

GERSHKOV, A., AND B. SZENTES (2004): "Optimal Voting Scheme under Costly Information Acquisition," mimeo, University of Chicago. 26

GREEN, J. R., AND N. L. STOKEY (2007): "A Two-person Game of Information Transmission," *Journal of Economic Theory*, 127(1), 90–104. 7

GROSSMAN, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law & Economics*, 24(3), 461–483. 6, 24

HARRINGTON, JOSEPH E, J. (1993): "Economic Policy, Economic Performance, and Elections," *American Economic Review*, 83(1), 27–42. 7

HORI, K. (2006): "The Bright Side of Private Benefits," mimeo, Hitotsubashi University. 25

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

KAWAMURA, K. (2007): "Confidence and Competence in Expertise," mimeo, University of Oxford (Nuffield College). 23

KRISHNA, V., AND J. MORGAN (2001): "A Model of Expertise," *Quarterly Journal of Economics*, 116(2), 747–775. 24

Li, H. (2001): "A Theory of Conservatism," *Journal of Political Economy*, 109(3), 617–636. 26

Li, H., and W. Suen (2004): "Delegating Decisions to Experts," *Journal of Political Economy*, 112(S1), S311–S335. 24

Matthews, S., and A. Postlewaite (1985): "Quality Testing and Disclosure," *RAND Journal of Economics*, 16(3), 328–340. 24

Milgrom, P., and J. Roberts (1986): "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 17(1), 18–32. 8

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391. 6, 8, 24

Morris, S. (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 11, 227–253. 7

Persico, N. (2004): "Committee Design with Endogenous Information," *Review of Economic Studies*, 71, 165–191. 28

Shavell, S. (1994): "Acquisition and Disclosure of Information Prior to Sale," *RAND Journal of Economics*, 25(1), 20–36. 24

Shin, H. S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64(1), 253–264. 24

——— (1998): "Adversarial and Inquisitorial Procedures in Arbitration," *RAND Journal of Economics*, 29(2), 378–405. 8, 24

Spector, D. (2000): "Rational Debate and One-Dimensional Conflict," *Quarterly Journal of Economics*, pp. 181–200. 7

Suen, W. (2004): "The Self-Perpetuation of Biased Beliefs," *Economic Journal*, 114(495), 377–396. 24

Szalay, D. (2005): "The Economics of Clear Advice and Extreme Options," *Review of Economic Studies*, 72, 1173–1198. 21, 26

Tversky, A., and D. Kahneman (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124–1131. 3

Van den Steen, E. (2004): "The Costs and Benefits of Homogeneity, with an Application to Culture Clash," mimeo, MIT Sloan. 25

Van den Steen, E. (2005): "Organizational Beliefs and Managerial Vision," *Journal of Law, Economics and Organization*, 21(1), 256–283. 7

Yildiz, M. (2003): "Bargaining without a Common Prior—An Immediate Agreement Theorem," *Econometrica*, 71(3), 793–811. 7