2008

When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests

Kristof Madarasz
Ming Li

When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests ⋆

Ming Li

1455 Boul. de Maisonneuve O., Department of Economics, Concordia University, Montreal, QC, Canada H3G 1M8

Kristóf Madarász

549 Evans Hall, Department of Economics, University of California at Berkeley, Berkeley, CA 94720

Abstract

We study the quality of advice that an informed and biased expert gives to an uninformed decision maker. We compare two scenarios: mandatory disclosure of the bias and non-disclosure, where information about the bias can only be revealed through cheap-talk. We find that in many scenarios non-disclosure allows for higher welfare for both parties. Hiding the bias allows for more precise communication for the more biased type and, if different types are biased in different directions, may allow for the same for the less biased type. We identify contexts where equilibrium revelation allows but mandatory disclosure prevents meaningful communication. JEL Codes: C72, D83.

Key words: cheap-talk, conflicts of interest, disclosure

1 Introduction

In a variety of contexts informed experts advise clients about what actions to take. Stock analysts advise investors on how to allocate their portfolios. Policy experts advise politicians on what policies to adopt. Medical researchers advise the government on whether to approve a new drug. However, experts are often prone to not provide accurate advice because they have conflicting interests. They may have incentives to distort information so as to induce decisions that are favorable to them. A stock analyst might have taken a short or long position on the stock she recommends. A policy expert might have a conservative or liberal bias or be beholden to special interests. A medical researcher may wish to promote drugs from a company funding her research. Conflicts of interest are at the heart of many recent corporate scandals and

⋆ This paper is combined from our two independent working papers: Li (2005) and Madarász (2005). Li is grateful to Larry Samuelson for numerous insightful discussions and constant encouragement, and thanks Yeon-Koo Che, Jing Li, Tima Mylovanov, Lucía Quesada, seminar participants at Concordia, McGill, Wisconsin, and the Canadian Economic Theory Conference, and especially Bill Sandholm, for helpful comments. Li acknowledges support from Concordia University Startup Research Fund. Madarasz thanks Attila Ambrus, Vincent Crawford, Matthew Rabin, Tomasz Sadzik, and seminar participants at UC Berkeley and the Hungarian Academy of Sciences for great suggestions, and especially Botond Koszegi for encouragement and unbiased advice. Both of us are grateful to the editor and two anonymous referees for very constructive suggestions. All remaining errors are our own.

Email addresses: `mingli@alcor.concordia.ca` (Ming Li),
`madarasz@econ.berkeley.edu` (Kristóf Madarász).

are often thought as major impediments to ideal information transmission in markets.

A commonly advocated partial solution to the conflicts of interest problem is mandatory disclosure. It is argued that mandatory disclosure eliminates the need for the decision maker to second-guess the recommendations of the expert and such transparency allows for decisions that are more aligned with the interest of the decision-maker. [1] In contrast to this common presumption, we find that for a wide range of environments, both the expert and the decision maker benefit from keeping the expert's conflict of interest private. Therefore, disclosure often not only does not alleviate the conflict-of-interest problem, but exacerbates it.

Our work extends the seminal cheap-talk model introduced by Crawford and Sobel [7] (CS henceforth) by allowing the decision maker to be uncertain about the expert's bias. In their model, an expert who observes some payoff-relevant information advises an uninformed decision maker through costless reports. CS show that when the expert's conflict of interest is common knowledge, the expert cannot fully reveal the state to the decision maker, but may transmit information about the state through noisy communication. In our model, the

[1] For example, Stiglitz [26] argues that "conflicts of interest will never be fully eliminated, either in the public or private sector. But by sensitizing ourselves to their presence, by increasing required disclosure – as the old saying goes, sunshine is the strongest antiseptic – by becoming aware of the incentives that are in place that can exacerbate these conflicts of interest, and by imposing regulations that limit their scope, we can do much to mitigate their consequences, both in the public and the private sector."

expert's bias is her private information.[2] The corresponding uncertainty that the decision maker faces may concern both the direction and the degree of the expert's bias. The situation where a policy expert might be conservative or liberal constitutes an example of the former; the situation where a stock analyst might be more self-serving or less constitutes one of the latter.

To gain some intuition for the results of the paper, consider a world with five possible states, decisions, and reports: "extremely low," "moderately low," "average," "moderately high," or "extremely high." Each state occurs with equal probability. The expert has either a right or left bias with equal probability. A right-biased expert would like the decision maker's action to be one notch above the true state (for example, action "extremely high" in state "moderately high"), and a left-biased expert would like it to be one notch below. Assume also that the decision maker and the expert's loss in utility from suboptimal actions is symmetric around his or her most preferred action. As a concrete example, one could think of a medical researcher (expert) assessing the severity of the side effects of a newly developed drug and giving recommendations to doctors or the government (decision maker). If the medical researcher receives funding from the drug maker, then she might have an incentive to understate its side effects (left bias); if she receives funding from the drug maker's competitor, then she might have an incentive to overstate them (right bias).

As demonstrated by CS, when the expert's bias is commonly known, effective communication between the expert and the decision maker is hindered by mis-

[2] In this paper, we use the word "bias" and the phrase "conflict of interest" interchangeably.

aligned preferences. Only noisy communication is possible. Take, for example, a right-biased expert. At the state "moderately low," she is indifferent between actions "extremely low" and "extremely high." Therefore, she cannot credibly distinguish between the three highest states. Hence, these three states are pooled together into one message. But given the decision maker takes an action equal to the expected state, the expert would also want to pool "moderately low" with the high message. The only equilibrium with any information transmission then involves the expert pooling the four highest states together and reporting the "extremely low" state as a separate message. Much information is lost here.

When there is uncertainty about the expert's bias, however, the decision maker can reason as follows. When he receives the message "moderately high," it could be because the state is "extremely high," and the expert has the left bias, or because the state is "average," and the expert has the right bias. Thus, the "expected value" of the state *is* moderately high and he finds it optimal to take the action "moderately high." This reasoning works for all three moderate messages.[3] Given the decision maker's actions, the expert finds it optimal to report such messages, since by so doing she receives her most preferred action for the bias-state combinations. This effect enables the expert and the decision maker to achieve better communication. Clearly, the expert is better off under nondisclosure. Because the decision maker has concave preferences, he too is better off.

This example shows that uncertainty about the expert's bias allows the decision maker to be more rather than less trusting of an actual recommendation

[3] However, a right-biased expert still cannot separate the two highest states.

because he does not know whether the expert is exaggerating or understating the truth. The expert can then respond to this fact by giving more precise recommendations which in turn improves the welfare of both parties. To generalize this intuition, we consider a model with a continuum of states and allow for uncertainty both about the direction and the degree of the conflict. We introduce two types of experts: a "high type" whose conflict takes a high value and a "low type" whose conflict takes a low one.

We show that there always exist conflict-hiding equilibria, where each report is issued by both types. However, each type issues the same report for a different set of states: the high type for a lower set and the low type for a higher set. In certain contexts, there can also be equilibria where for a limited set of states, the conflict is revealed through cheap-talk: we call such equilibria partially conflict-revealing.

We show that it is always true that the high type can communicate with less noise in a conflict-hiding equilibrium than under disclosure. Because the decision maker discounts her advice less than if the conflict were known, less information is lost under nondisclosure. Given concave preferences, this implies higher welfare for both players. If uncertainty concerns only the direction of the conflict then the same is true for the low type. This implies that for all concave preferences conflict-hiding equilibria allow for higher welfare than disclosure for both players. If uncertainty concerns only the degree of the conflict but its direction is known, the low-type expert's advice is discounted more in a conflict-hiding equilibrium than under disclosure. Thus, the low type can communicate with less noise and become better off when her bias is disclosed. In this case, the overall welfare implications of disclosure might depend on the exact shape of the players' preferences.

In general, consider a case where uncertainty concerns the degree and possibly also the direction of the conflict. We show that if players exhibit a sufficiently increasing distaste for inaccurate advice, as reflected by the concavity of the players preferences, then their dominant concern is to decrease the amount of noise when communication is least accurate. Since this occurs when the conflict of interest is known to be high and nondisclosure reduces the amount of noise in the high type's messages, we show that nondisclosure is welfare-improving.

All these results indicate that in a large set of environments, not requiring disclosure could be a better policy. The situation where disclosure is clearly a better choice happens if the only conflict-hiding equilibrium is a babbling one yet it is possible for the low type to transmit information when her bias is disclosed. This can only happen if there is no information transmission in the disclosure equilibrium when the expert is of the high type. Even in this case, however, disclosure is not necessarily better than nondisclosure. In certain scenarios, the low-type expert is able to reveal the value of her conflict through some equilibrium messages. We show that such partial conflict revelation can allow meaningful communication in environments where neither conflict-hiding equilibria nor disclosure equilibria do. In particular, this is the case when the two possible values of conflicts are of large magnitudes and opposite signs.

The rest of the paper is organized as follows. In Section 2, we review literature related to our work. In Section 3, we develop a simple model of cheap-talk with uncertainty about biases. In Section 4, we characterize conflict-hiding equilibria. In Section 5, we present a welfare analysis, and show that in many scenarios nondisclosure dominates disclosure. In Section 6, we discuss the possibility of conflict revelation in nondisclosure equilibrium which improves welfare. Finally, in Section 7, we summarize the results and propose directions for future research.[4]

2 Related literature

The paper most closely related to ours is that by Morgan and Stocken [21]. In their model, a stock analyst could be "bad," who prefers investors to overvalue stocks, or "good," who prefers investors to value them correctly. They show that the presence of bad experts prevents good experts from revealing good news but not from revealing bad news and induces good experts to issue favorable reports more frequently.

Our model differs from that of Morgan and Stocken [21] in two key ways. First, we consider how the players' dislike of inaccurate advice determines the welfare effects of uncertainty about the conflict. Second, through extensive comparative static analysis, we compare welfare under disclosure and nondisclosure to address the optimality of these two regimes. These departures from their model allow us to identify the mechanism through which uncertainty about incentives might benefit both the decision maker and the expert, which in turn allows us to relate to a wide range of important policy considerations.

Papers on cheap-talk with uncertain biases also include those by Dimitrakas and Sarafidis [10] and Morris [22]. Similarly to [21], they consider bias distributions skewed in one direction – unbiased and right-biased. In contrast, we allow both the magnitude and the direction of the expert's bias to be un-

[4] In addition to the proofs in the Appendix, we provide an online supplement for proofs we omitted from the paper.

certain. Dimitrakas and Sarafidis [10] characterize cheap-talk equilibria with uncertain biases, when the bias value is allowed to be distributed on a continuum. [5] Morris [22] employs a discrete state space, allows two possible values of bias, but assumes that the expert only imperfectly observes the state. He shows how an expert's concern for reputation as an unbiased expert may prevent him from revealing information.

Cain, Loewenstein, and Moore [5] analyze the effects of disclosure of conflicts of interest on expert advice in an experimental setting. They find that disclosure leads to greater distortions than nondisclosure, i.e., more noise in advice and lower earnings for the decision maker, both of which are consistent with our results. Although they attribute some of these effects to a mix of strategic and psychological factors (e.g., credulity, naïveté, and anchoring), our analysis shows that these effects can arise in a perfectly Bayesian setting.

Sobel [25] and Bénabou and Laroque [2] study reputation concerns of experts when there are honest advisors and strategic ones. [6] Other authors focus on uncertainty about another dimension – competence of experts or accuracy of experts' information. Austen-Smith [1], Ottaviani and Sørensen [24], and Moscarini [23] are a few examples.

Farrell and Gibbons [14] study the effects of the presence of different *audiences* on cheap-talk, under a setup based on a discrete state space. They find

[5] However, they do not provide welfare comparisons between disclosure and nondisclosure.

[6] Dziuda [12] also studies a model with biased and honest types, but in her model, the expert's information is partially verifiable, in that the expert cannot claim to have evidence she does not possess.

the effect could be subversion, one-sided discipline, or mutual discipline.[7] In contrast, we study the effect of the existence of different types of *speakers*. The effect that one type of expert has on the other is close in spirit to their "mutual discipline" effect.

3 Model

A privately informed expert (E or she) gives advice to an uninformed decision maker (D or he). The decision maker's decision affects both parties' payoffs, which also depend on the value of an underlying state, s. The state s is a random variable uniformly distributed on $[0, 1]$. The expert privately observes the realization of s and sends to the decision maker a costless message m from an arbitrarily large message set M. After receiving the message, the decision maker takes an action $y \in \mathbf{R}$. In state s, the decision maker's most preferred action is equal to s. If the expert has bias β, her most preferred action is $s+\beta$. If $\beta > 0$, we say that the expert has a right bias, while if $\beta < 0$, we say that the expert has a left bias.

The decision maker's and the expert's utility functions are respectively

[7] *Subversion* refers to cases in which the speaker is able to communicate to one audience in private, but the presence of another audience prevents him from such communication in public. *One-Sided Discipline* means that the speaker cannot communicate to one audience in private, but the presence of another audience enables him to effectively communicate with this audience. *Mutual discipline* refers to cases in which the speaker is not able to communicate to either audience in private, but is able to do so in public.

$$U^D(y,s) = U(y-s),$$
$$U^E(y,s,\beta) = \tilde{U}(y-(s+\beta)),$$

where both U and $\tilde{U}$ are strictly decreasing in the absolute value of their sole argument, the difference between the actual action y and the decision maker's (expert's) most preferred action.[8] At times, we refer to this property as *distance aversion*. Note that distance aversion implies *symmetry*, i.e., U and $\tilde{U}$ are even functions. In addition, we assume U and $\tilde{U}$ are strictly concave and twice continuously differentiable functions.[9] We will use V to indicate the expected utility of the decision maker or expert in a strategy profile, with the appropriate subscripts/superscripts and arguments given the contexts. For example, V^D is the expected utility of the decision maker in a strategy profile.

The expert's bias β is her private information and drawn from the following distribution:

$$\beta = \begin{cases} b_h \text{ with probability } p, \\ \\ b_l \text{ with probability } 1-p, \end{cases}$$

where $-b_h \leq b_l < b_h$. This is without loss of generality due to the symmetry of players' preferences. The following two quantities are useful in the characterization of equilibrium:

$$d = b_h - b_l \in [0, 2b_h], \; v = \frac{b_h + b_l}{2} \in [0, b_h].$$

[8] This argument is sometimes called the "outcome" of the decision maker's action in the cheap-talk literature.

[9] The assumption on the expert's utility, $\tilde{U}$, is only needed when we discuss welfare results concerning the expert. For all the results concerning equilibrium characterization and the decision maker's payoff, we need only assume it to be distance-averse.

Thus, d is the difference between the two bias values, and v the average of them.

We consider only pure strategies for the expert. This is without loss of generality in terms of players' expected payoffs, as the set of states in which the expert is indifferent between actions has measure zero. A pure strategy for an expert with bias β can then be characterized by the function $\mu_\beta : [0, 1] \to M$. Let $P(\cdot|m)$ be the belief of the decision maker about the state when he receives the message m. Let $y(m)$ be the action taken by the decision maker if he receives message m.

We focus our attention on *Perfect Bayesian Equilibrium*: (E1.) The decision maker's beliefs, $P(\cdot|m)$, be formed using Bayes' rule for any message m whenever possible;[10] (E2.) The decision maker's actions, $y(m)$, maximize his expected utility given his belief $P(\cdot|m)$ for all m; (E3.) The expert's messages, $\mu_\beta(s)$, maximize her utility for all s among all $m \in M$ given the decision maker's strategy.[11]

To eliminate essentially equivalent equilibria we make two assumptions. First, when an expert is indifferent between two equilibrium actions she induces the lower one if she is of the high type and the higher one if she is of the low type.[12] Second, we assume that if messages m and m' induce the same action

[10] To be precise, we need $P(\cdot|m)$ to be the regular conditional probability defined by the joint distribution of m and s. See Durrett [11] for a detailed discussion.

[11] When characterizing equilibria, we will omit beliefs, as beliefs about unsent messages can be specified the same as any message sent in equilibrium, so that they do not disrupt E2 and E3.

[12] This assumption ensures all equilibrium actions that are potentially optimal for a type of expert will be induced by that type of expert.

then $m = m'$ and adopt the language "the message that induces action y," or simply, "message y."

We call an equilibrium "informative" if it gives the decision maker higher payoffs than the babbling equilibrium, in which the expert's reports do not reveal any information about the state.[13]

To study the welfare effects of mandatory disclosure of conflicts of interest, we compare two information regimes. In the first regime, which we call *disclosure*, there is a prior commitment to disclosing the conflict. In the second regime, which we call *nondisclosure*, there is no such prior commitment and the conflict cannot be verifiably revealed. We say that one regime allows for higher welfare than another if there is an equilibrium which delivers higher welfare for both players in the former than any equilibrium in the latter. We will talk about the decision maker and the expert's ex ante expected payoff, namely, before the realization of both the expert's bias and the underlying state. The ex interim expected payoff of the expert refers to her expected payoff after the value of her bias is realized, but before that of the state does.

Note that our characterization of the equilibrium under disclosure and nondisclosure also characterizes the equilibrium in a third scenario, where the expert could ex interim *voluntarily* disclose the conflict (mixing is not allowed). This

[13] In cheap-talk games, it is not unreasonable to expect to observe informative equilibria when they exist. In fact, there is experimental evidence (Blume et al. [3] and Cai and Wang [4]) to support this claim. A recent paper by Chen, Kartik, and Sobel [6] provides a justification in the form of an equilibrium selection criterion. However, there are also theoretical limitations to selecting such equilibria, as discussed by Farrell [13] and Farrell and Rabin [15].

is because if the strategy of disclosure is observable and one type finds it optimal to disclose, then the decision maker will find out the value of the conflict independently of the other type's decision. Hence, this scenario reduces to the disclosure case. If both types keep the conflict private then there is no change in the decision maker's prior beliefs about the conflict and it reduces to the nondisclosure case. Under this scenario the expert's voluntary decision not to disclose her conflict regardless of its value accords with a stronger welfare criterion. We say that in this case nondisclosure ex interim dominates disclosure for the expert.

Our first lemma describes how an expert's ranking of two actions depends on the underlying state and her bias.

Lemma 1 *Given two actions, y and y' with $y < y'$, an expert of bias β strictly prefers y to y' in state s if and only if the expert's most preferred action, $s+\beta$, is less than $\frac{y+y'}{2}$. When $s+\beta = \frac{y+y'}{2}$, the expert is indifferent between them.*

PROOF. [***Lemma 1, Page 14***] Observe $s + \beta \leq \frac{y+y'}{2}$ is equivalent to $|y - (s+\beta)| \leq |y' - (s+\beta)|$, which is equivalent to $\tilde{U}(y - (s+\beta)) \geq \tilde{U}(y' - (s+\beta))$ as $\tilde{U}$ is decreasing in the absolute value of its sole argument.

The lemma implies that in any equilibrium, for either type of expert, the set of states in which she finds it optimal to induce a particular action is a closed interval, as arbitrary intersections of closed intervals remain closed intervals. In addition, the lemma is true regardless of the state distribution.

In general, there can be only a finite number of actions in equilibrium. [14]

[14] Our result generalizes Lemma 1 of [7] and Lemma 2 of [21].

Lemma 2 *There are only a finite number of actions in equilibrium as long as neither* b_l *nor* $pU'(b_h) + (1-p)U'(b_l)$ *is equal to* 0*. Furthermore, there cannot be an infinite number of actions induced by both types, unless* $pU'(b_h) + (1-p)U'(b_l)$ *is equal to* 0*.*

Unless one of the two expressions in the lemma is equal to zero (note they cannot both be zero except in degenerate cases), actions in equilibrium must be relatively far apart from one another, or equivalently, messages must contain enough noise, to ensure that the expert stick to her equilibrium messages. As a result, there can be only a finite number of equilibrium actions.

It is possible to have an infinite number of actions in equilibrium if one of the expressions in Lemma 2 *is* zero. First, we consider a case where $pU'(b_h)+(1-p)U'(b_l) = 0$, where it is possible to have an infinite number of equilibrium actions that are induced by both types. [15]

Lemma 3 *Let* $b_h = -b_l = b < \frac{1}{2}$ *and* $p = \frac{1}{2}$*. Then, the following strategy profile constitutes an equilibrium:*

1. *An expert of bias* β*'s strategy satisfies*

$$\mu_\beta(s) = \begin{cases} b & \textit{if } \beta = -b, \textit{ and } s \in [0, 2b]; \\ s+\beta & \textit{if } s \in [b-\beta, 1-b-\beta]; \\ 1-b & \textit{if } \beta = b, \textit{ and } s \in [1-2b, 1]. \end{cases}$$

2. *Upon receiving message* m*, the decision maker takes action* $y(m) = m$ *for all* $m \in [b, 1-b]$*.*

[15] The equilibrium construction also appears in [8].

We omit the proof as it is straightforward. Figure 1 provides an illustration of the equilibrium. For any message y in the interval $(b, 1-b)$, both types of expert induce their most preferred action. Also, as the expected misstatement is zero, the decision maker finds it optimal to take action y when he receives this message. Thus, any message $y \in (b, 1-b)$ is precise, in that any variation in the underlying state is reflected in the expert's report. This already suggests that nondisclosure generates higher welfare. We will return to this point in Section 4. Here, we want to stress two points. The first point is regarding comparisons between communication versus delegation. Dessein [9] shows that under uniform distribution of state, quadratic preferences, and common knowledge of bias, full delegation always gives the decision maker higher utility than communication, as long as informative communication is possible. However, the equilibrium in the above lemma demonstrates this need not be the case when the decision maker is uncertain about the expert's bias. Full delegation gives the decision maker a payoff of $-b^2$. But when $p = \frac{1}{2}$, the most informative communication equilibrium gives the decision maker a payoff of $-b^2 + \frac{4}{3}b^3$, which is always higher.[16] The second point is regarding the assertion by CS that the expert's bias is self-defeating in that the expert would prefer to commit to telling the truth if possible. However, under quadratic preferences, in Lemma 3, either type of expert's payoff is $-\frac{8}{3}b^3$. In contrast, if the expert committed to truth-telling, her payoff would be $-b^2$. Therefore, the expert does not want to commit to truth-telling if $b \leq \frac{3}{8}$. In other words, only large biases are self-defeating.

When $b_l = 0$, it is also possible to have an infinite number of actions in

[16] Due to continuity of preferences, for any $b \in (0, 1/2)$, if p is close to $1/2$, the decision maker is better off with communication than with delegation.

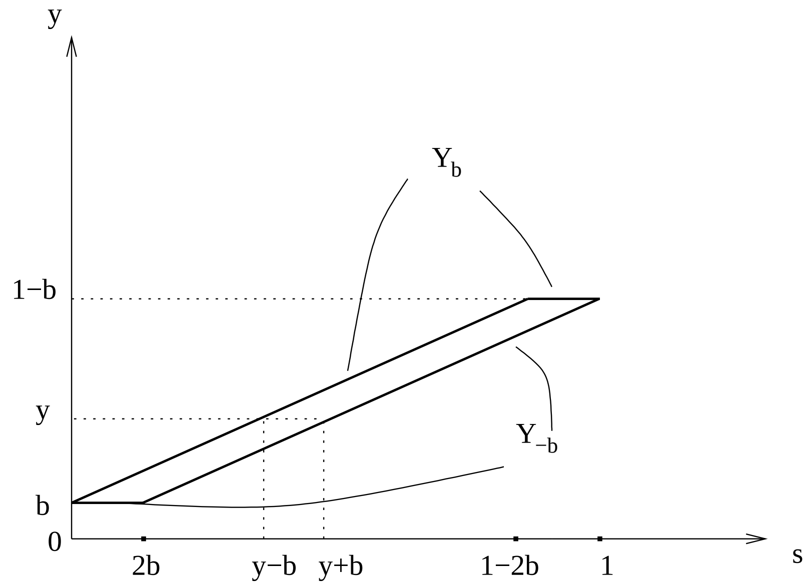

Fig. 1. Equilibrium when $b_h = -b_l$ and $p = \frac{1}{2}$, where Y_β maps from the true state to the action through an expert of bias β and the decision maker.

equilibrium.

Example 4 *Let* $b_h = \frac{1}{4}$*,* $b_l = 0$*, and* $p = \frac{1}{2}$*. Suppose also the decision maker's preferences are quadratic. Then the following strategy profile is an equilibrium and it features an infinite number of actions – the interval* $[0, 0.2156)$*, and two points* $y_1 = 0.2156$*, and* $y_2 = 0.6403$*. The low type induces actions equal to the state in* $[0, 0.2156)$*, action* y_1 *in* $[0.2156, 0.4280)$*, and action* y_2 *in* $[0.4280, 1]$*; the high type induces* y_1 *in* $[0, 0.1780]$ *and* y_2 *in* $(0.1780, 1]$.

Note that in the above example, some actions (namely, those in $[0, 0.2156)$) are induced by only one type of expert (the low type). Therefore, messages inducing such actions reveal the type of the expert. We call this type of equilibrium *partially conflict-revealing equilibrium.* Alternatively, if each equilibrium action is induced by both types, we call the equilibrium *conflict-hiding equilibrium.*[17] In this kind of equilibrium, no equilibrium message reveals the type

[17] In the context of stock recommendations, Morgan and Stocken [21] analyze two

of the expert, hence the term conflict-hiding. The equilibrium in Lemma 3 is a limiting case of these two types of equilibria. In the equilibrium, all actions are induced by both types, but there are actions that are induced by one type in an interval and the other type at only one point. When receiving such a message, the decision maker can infer with probability one what the expert's bias is. We will nevertheless call equilibria of this type conflict-hiding equilibria, as our results on conflict-hiding equilibria do apply to them.

While conflict-hiding equilibrium always exists, the same is not true for partially-conflict revealing ones. In the rest of Section 4 and Section 5, we will focus on conflict-hiding equilibria and return to revelation of conflicts in Section 6.

kinds of equilibria, semiresponsive and categorical ranking. The categorial ranking equilibria correspond to our conflict-hiding equilibria, while the semireponsive equilibria are a kind of partially conflict-revealing equilibria in which certain messages exclusive to the unbiased expert perfectly reveal the underlying states to the decision maker, just as in our Example 4.

4 Conflict-Hiding Equilibria

In our characterization of conflict-hiding equilibrium, we will focus on the case where at least one of the following holds: [18]

$$p \geq \frac{1}{2} \text{ or } b_l \geq 0.$$

Consider an equilibrium with n actions, $y_1, \ldots, y_n$, arranged in ascending order. Lemma 2 asserts that these are the only possible conflict-hiding equilibria, unless $pU'(b_h) + (1-p)U'(b_l)$ is equal to 0. For $i = 1, \ldots, ..., n$, the point at which an expert of bias b_j $(j = h, l)$ is indifferent between y_i and y_{i+1} is

$$a_i^j = \frac{y_i + y_{i+1}}{2} - b_j. \tag{1}$$

In addition, let $a_0^j = 0$ and $a_n^j = 1$. Thus, $\left\{a_i^j\right\}_{i=0}^n$ $(j = h, l)$ forms a partition of the interval $[0, 1]$. In the $i-$th interval for type j, the expert sends the message that induces y_i, where $j = h, l$ and $i = 1, \ldots, n$. For economy of notation, let

$$a_i \equiv a_i^h.$$

Thus, $a_i^l = a_i + d$ when $i = 1, \ldots, n-1$. [19]

[18] In the case where both $p < 1/2$ and $b_l < 0$ hold, we cannot fully characterize the set of conflict-hiding equilibria without additional assumptions. We know that if an $n-$action equilibrium exists, then it is unique. The difficulty lies in showing that the existence of an $(n+1)-$action equilibrium implies that of an $n-$action one. When the decision maker's preferences are quadratic, results similar to that of Theorem 5 can be established. We should note, however, that the difference equations characterizing an $n-$action equilibrium does apply to the case with $p < 1/2$ and $b_l < 0$.

[19] From this relationship, we can immediately observe that $d < 1$ is necessary for

Let $y(a^h, \tilde{a}^h, a^l, \tilde{a}^l)$ be the action that maximizes the expected utility of the decision maker given that the high type sends the corresponding message in $[a^h, \tilde{a}^h]$ and the low type in $[a^l, \tilde{a}^l]$. Thus, the equilibrium actions can be written as

$$\begin{aligned} y_1 &= y(0, a_1, 0, a_1 + d), \\ y_i &= y(a_{i-1}, a_i, a_{i-1} + d, a_i + d), \qquad\qquad i = 2, \ldots, n-1, \\ y_n &= y(a_{n-1}, 1, a_{n-1} + d, 1). \end{aligned} \tag{2}$$

Adding adjacent equations in (2) and using (1), we may obtain a system of difference equations of a_i.[20] The following theorem fully characterizes the conflict-hiding equilibria.

Theorem 5 *Consider the game $\{U, \tilde{U}, p, b_h, b_l\}$ and assume $p \geq \frac{1}{2}$ or $b_l \geq 0$. Unless $p = \frac{1}{2}$ and $b_h = -b_l$, there exists some positive integer $N(p, b_h, b_l)$ such that for each $n = 1, 2, ..., N(p, b_h, b_l)$ there exists a unique solution, $(\{a_i\}_{i=0}^n, \{y_i\}_{i=1}^n)$ to (1) and (2). The solution corresponds to a unique equilibrium with n partition intervals for each type. Furthermore, these are the only conflict-hiding equilibria. When $p = \frac{1}{2}$ and $b_h = -b_l < \frac{1}{2}$, there exists a partition equilibrium with n partition intervals for each positive integer n.*

Let $\{m_1, .., m_n\}$ be a set of distinct messages. The equilibrium with n partition

the existence of informative conflict-hiding equilibrium.

[20] Please refer to Equation (11) in the appendix for the exact form of the difference equations. The set of difference equations satisfies a version of "condition M" used by CS. That is, keeping $a_1, \ldots, a_{i-1}$ unchanged and increasing a_i, in the forward solution (solving the difference equations while ignoring the last equation), all cutoff points from a_{i+1} onward will be shifted to the right. This guarantees the uniqueness of an $n-$action equilibrium.

elements can be characterized by:

(1) *Each type* ($j = h, l$) *of expert partitions* $[0,1]$ *into* n *intervals with cutoff points* $\{a_i^j\}_{i=1}^{n-1}$ *and sends message* m_i *in the* i*-th interval, where* $a_i^h = a_i$ *and* $a_i^l = a_i + d$.

(2) *The decision maker takes action* y_i *upon receiving message* m_i.

In addition, $a_{i+1} - a_i \geq a_i - a_{i-1}$ *for all* $i = 1, \ldots, n-1$*, and* $a_{i+1}^l - a_i^l \geq a_i^l - a_{i-1}^l$ *for all* $i = 1, \ldots, n-1$ *when* $b_l \geq 0$.

Theorem 5 establishes that, for each positive integer up to an upper bound, there exists a single equilibrium with that number of actions. Given our assumption that the expert on average has a right bias, the theorem also implies that higher messages contain more noise than lower ones, similar to the conclusion of CS.

The equilibrium under disclosure is a limit case of the conflict-hiding equilibrium when p goes to 1. To compare equilibria as p changes, we define two measures: balancedness and precision of an equilibrium. We say an equilibrium is more *balanced* the smaller is the increment in noise from a lower message to a higher message; we say an equilibrium is more *precise* if it includes a higher number of distinct actions. The decision maker prefers a more balanced and precise equilibrium since he has concave preferences. We also say that a message is more *precise* or less *noisy* than another if the set of states where it is sent in equilibrium has a lower measure than the other one.

To examine the properties of the conflict-hiding equilibria more closely, let us define x_i by

$$y_i = \frac{a_{i-1} + a_i}{2} + x_i. \tag{3}$$

where the x_i depend on the shape of U. They are decreasing in p, equal to d for $p = 0$, and 0 for $p = 1$.[21] Then, the system of difference equations for $\{a_i\}_{i=0}^n$ can be written

$$\begin{aligned} a_2 - a_1 &= a_1 - 0 + 4b_h - [2y_1 - a_1 + 2x_1], \\ a_{i+1} - a_i &= a_i - a_{i-1} + 4b_h - 2[x_i + x_{i+1}], \qquad\qquad i = 2, \ldots, n-1, \\ 1 - a_{n-1} &= a_{n-1} - a_{n-2} + 4b_h - [2y_n - a_{n-1} - 1 + 2x_n]. \end{aligned} \tag{4}$$

When there is no uncertainty about the conflict, i.e., $b_h = b_l = b$, the above equation reduces to

$$a_{i+1} - a_i = a_i - a_{i-1} + 4b \text{ for } i = 1, 2, ..., n-1. \tag{5}$$

with $a_0 = 0$ and $a_n = 1$. This case also corresponds to the disclosure equilibrium (the equilibrium studied by CS) where the value of the bias b is common knowledge. In such a disclosure equilibrium the increment in noise is $4b$. Note that this quantity is negative if $b < 0$, implying that when the expert has a left bias then higher messages are more precise than lower messages. The increments in noise in conflict-hiding messages is always smaller than $4b_h$ and greater than $4b_l$. In addition, given our assumption that the expert on average has a right bias, higher messages are always less precise. The following two corollaries pin down the comparative statics with respect to the distribution of the conflict.

Corollary 6 *Fixing $n \geq 2$, the equilibrium cutoff points a_i are strictly decreasing in p if an $n-$action equilibrium exists.*

[21] For $i = 2, ..., n-1$, there exists a function x such that $x_i = x(p, a_i - a_{i-1})$ and $x(1/2, \cdot) = d/2$. The latter makes it easy to characterize the conflict-hiding equilibrium when $p = 1/2$.

Corollary 7 *The maximum number of intervals (actions) in an equilibrium, $N(p, b_h, b_l)$, is nonincreasing in p.*

Corollary 6 states that for an equilibrium with a given number of actions, the cutoff points move to the left as the probability of high type increases. Intuitively, a higher p corresponds to the bias value being more right-biased "on average." This causes an increase in the increments in the noise of messages moving from the left to the right. Therefore, the first message has to be more accurate, and the noise in all equilibrium messages becomes less balanced. This intuition is the same as identified by CS. As we will see in Section 5, since the decision maker's preferences are strictly concave, this leads to a decrease in his payoffs. Corollary 7 states that a higher p renders it less likely for there to exist equilibria with larger number of actions. The total amount of noise in all messages is fixed since the set of states is a fixed interval. Since higher p requires increments in noise to increase, it makes it less likely for an $n-$action equilibrium to exist.

In the conflict-hiding equilibrium, the high type and the low type induce the same action in two different intervals: a higher one for the low type and a lower one for the high type. The decision maker knows this fact, but from observing a message he cannot tell the type of the expert. As a result he is more willing to believe the expert than if he knew that she was the high type. Consequently, the high type will be able to communicate with a more balanced noise distribution in an n-action conflict-hiding equilibrium than in an n-action disclosure equilibrium. Furthermore, from Corollary 7, it follows that whenever there is an n-action disclosure equilibrium with the high type there is also an $n-$action conflict-hiding equilibrium. Therefore, nondisclosure allows for more precise and more balanced communication for the high type.

If uncertainty concerns only the degree of the bias then at the same time the decision maker is less willing to believe the expert than if he knew that she was the low type. Yet if there is uncertainty about the direction of bias then the decision maker might be more willing to believe the expert than if he knew she was a low type. Thus, if the low type's bias is of the same direction as that of the high type, then noise is less balanced in a conflict-hiding equilibrium than under disclosure. Also, a disclosure equilibrium with the low type allows for more precise communication as follows from Corollary 7. When there is uncertainty about the direction of the conflict however, it is possible that the low type can communicate with more evenly distributed noise in an n-action conflict-hiding equilibrium than in an n-action disclosure equilibrium.In particular, when $p = 1/2$, $b_l < 0$, and $|b_l| > v$, we can show that conflict-hiding equilibrium can be more precise than a disclosure equilibrium with bias b_l.[22]

5 Welfare

With the help of the comparative static results of the previous section, we now turn to the welfare comparison between the conflict-hiding and disclosure equilibria. We maintain our assumption that either $p \geq 1/2$ or $b_l \geq 0$..

Our first results concerns how the decision maker's expected payoff depends on the bias distribution when we fix the number of equilibrium actions.

Lemma 8 *Given any $n \leq N(p, b_h, b_l)$ and $n \geq 2$, the decision maker's equilibrium expected payoff in the n-action equilibrium, $V_n^D(p, b)$, is strictly decreasing in p.*

[22] Please refer to [20].

This lemma says that the decision maker strictly prefers an equilibrium in which the probability of the high type, p, is smaller, given a fixed number of actions in the equilibrium. A lower p corresponds to a lower weight placed on the expected utility from consulting the high type. This increases the decision-maker's payoff as his expected payoff from consulting the low type is higher. A lower p also leads to a change in the noise in each type of expert's messages due to a rightward shift of cutoff points. When $b_l \geq 0$ this shift causes the noise of the low type's messages to become more balanced. When $b_l < 0$, it makes it less balanced. However, it always causes the noise of the high type's messages to become more balanced. When $p \geq \frac{1}{2}$ this offsets the effect of the noise of the low type's messages becoming less balanced. Thus, given $p \geq \frac{1}{2}$ or $b_l \geq 0$, the decision makes becomes better off as a result of the change in the noise distributions of each type of expert's messages. Intuitively, the combination of the two effects means that the overall noise in communication becomes more balanced, which benefits the decision maker due to his strictly concave preferences. [23]

We now turn to welfare comparisons for the expert. Intuitively, the high type prefers not to have her bias disclosed because she can imitate the low type under nondisclosure, thereby improving her influence on the decision maker. For the low type, the comparison depends on whether her bias is of the same direction as the high type. The following theorem shows that if it is the case, the low type prefers disclosure to nondisclosure.

Theorem 9 *Assume the expert's utility function, $\tilde{U}$, is concave. Then,*

[23] The equilibrium actions also change as p changes. But, they are chosen by the decision maker to maximize his expected payoff. The envelope theorem tells us that such changes have no marginal effect on the decision maker's expected payoff.

(1) *For the high type, the highest expected payoff achievable under nondisclosure is higher than that under disclosure;*

(2) *For the low type, when $b_l \geq 0$, the highest expected payoff achievable under nondisclosure is lower than that under disclosure.*

The intuition of the result is as follows. First, fixing the number of equilibrium actions, the noise across the high type's messages under nondisclosure is more balanced than under disclosure. Even if the decision maker knows the messages are sent by the high type and takes the corresponding optimal action (the average of the two cutoff points, as he would under disclosure), the high type's expected payoff would still be higher than that under disclosure, since her preferences are concave. However, in the conflict-hiding equilibrium under disclosure, the decision maker has to account for the possibility that the message may have been sent by the low type. As a result, he chooses an action that is more favorable to the high type than if he knew the message has been sent by the high type. Combining these two comparisons, the high type is better off under nondisclosure than under disclosure, given the same number of actions in equilibrium. Finally, Corollary 7 states that nondisclosure allows at least as many actions as disclosure for the high type. Therefore, the high type is better off under nondisclosure. The argument is analogous for why the low type prefers disclosure when her bias is of the same direction as the high type, except that all three comparisons in the argument go the opposite direction.

As we have shown above, nondisclosure allows for more balanced and more precise communication between the high type and the decision maker than a disclosure equilibrium, which benefits both the decision maker and the high type. Our next result shows that the same argument applies for the low type if the low type's bias is of the same magnitude but opposite direction of the high type's. That is, $b_h = -b_l = b$. In this case, symmetry of the players' preferences implies that the set of equilibria when $p = p_0$ is simply the mirror image of that when $p = 1 - p_0$.[24] As a result, the welfare comparison of Lemma 8 can be generalized to imply that fixing the number of actions, the decision maker's expected payoff becomes higher as p becomes closer to $\frac{1}{2}$. The following theorem then follows from Lemma 8 and Theorem 9.

Theorem 10 *When $b_h = -b_l$, the highest expected payoff the decision maker can receive is higher under nondisclosure than under disclosure. In addition, the highest expected payoff the expert receives is also higher under nondisclosure than under disclosure, both ex ante and ex interim.*

PROOF. In terms of the decision maker's expected payoff, disclosure is equivalent to $p = 0$ and $p = 1$. According to Lemma 8, the n-action equilibrium under nondisclosure gives the decision maker higher payoff than the $n-$action equilibrium under disclosure. In addition, as we have shown in Corollary 7, the maximum number of actions allowed decreases as p increases from $\frac{1}{2}$ to 1 (or, by symmetry, decreases from $\frac{1}{2}$ to 0). Combining these two statements, we conclude the most informative equilibrium under nondisclosure gives the decision maker higher payoff than disclosure.

The second part follows from Theorem 9. Without loss of generality, we focus on the case $p \geq \frac{1}{2}$. First, note that the high type's and the low type's highest expected payoffs under disclosure are the same, again, due to symmetry. Theorem 9 implies the high type ex interim prefers nondisclosure to disclosure. In

[24] In fact, Theorem 12 in Section 6 implies that in this case all equilibria are conflict-hiding.

a conflict-hiding equilibrium under nondisclosure, it can be shown that the low type's ex interim expected utility is higher than the high type's.[25] Therefore, the low type also ex interim prefers nondisclosure to disclosure. Hence, the expert's ex ante expected payoff under nondisclosure is also higher than that under disclosure.

As we have seen in Lemma 3, when $p = 1/2$ the expert can induce an infinite number of actions. In that case, the conflict-hiding equilibrium is more precise and more balanced than any disclosure equilibrium, implying higher welfare for both players. The above theorem shows that this results extends to all p. Thus, when uncertainty concerns only the direction of the conflict, nondisclosure always allows for higher welfare. For the expert, this is true both before and after she learns her bias.[26]

When the decision maker faces uncertainty about the magnitude of the conflict, the welfare analysis is more ambiguous. We have shown that when the expert is of the high type, not disclosing the conflict always allows for higher expected payoffs for the decision maker and the expert. The case for the low type is more complex. As long as uncertainty about the direction of the conflict is dominant then in the spirit of Theorem 10 both the low type and the decision maker are also better off. When uncertainty concerns only the degree of the conflict then by Lemma 8 and Theorem 9 the decision-maker and the low type can be made better off in a disclosure equilibrium than in any

[25] The proof is omitted here, but is available in the online supplement.

[26] Under quadratic preferences, we may obtain additional welfare results. For example, the decision maker's highest expected payoff is decreasing in the value of b. We would like to refer the interested reader to [19].

conflict-hiding equilibrium. In these scenarios, the welfare comparison depends on whether the improvement in the quality of communication when the conflict is high or the deterioration when the conflict is low is more important. This tradeoff depends on the *concavity* of the players' preferences. Concavity of the preferences implies that everything else equal, players will incur losses from a decrease in the balancedeness of the noise distribution in the expert's advice. The more concave preferences are, the greater is the loss. When preferences are sufficient concave, the most important factor affecting the players' welfare is the loss from the least precise messages. From Corollary 6, it follows that the noise in communication is more balanced in an n-action conflict-hiding equilibrium than in an n-action disclosure equilibrium with the high type. This suggests that nondisclosure may allow for higher welfare than disclosure for both the decision maker and the expert, simply because the least precise message under disclosure is more precise than its counterpart in nondisclosure. Our next theorem formalizes this idea.

Theorem 11 *Fix* (p, b_h, b_l) *such that* $N(1, b_h, b_l) \geq 2$*. Then, there exists sufficiently concave utility functions for the players such that there is a conflict-hiding equilibrium that ex-ante allows for higher welfare than any disclosure equilibrium for both the decision maker and the expert.*

This theorem shows that when there is meaningful communication in a conflict-hiding equilibrium and preferences are sufficiently concave, the players can both achieve higher ex-ante welfare when the conflict remains private to the expert. Thus, the last two theorems above lead to the conclusion that there exists a wide range of scenarios where transparency about conflicts of interest does not improve communication. Thus, in our admittedly stylized framework, it is not optimal to implement a policy that requires mandatory disclosure of the expert's conflicts of interest prior to communication.

It is important to note that unlike Theorem 10, Theorem 11 holds only ex ante, i.e., before the expert learns her bias. In addition, a necessary condition for the theorem to be true is that meaningful communication is possible in a conflict-hiding equilibrium, which is implied by our assumption that there can be meaningful communication between the high type and the decision maker under disclosure. In the next section, we consider cases where there is no meaningful communication when the conflict is hidden yet nondisclosure might still allow for higher welfare than disclosure.

6 Revelation of the Conflict

In the previous sections we have characterized conflict-hiding equilibria and given sufficient conditions for a conflict-hiding equilibrium to Pareto dominate a disclosure equilibrium. One of these sufficient conditions is that there is meaningful communication in a disclosure equilibrium when the conflict is high because this ensures that there is meaningful communication when the conflict is hidden. If this condition fails, then the above welfare comparison will not necessarily hold. In particular this happens in the case where $|b_l| < 1/4$ but $N(b_h, bl, p) = 1$. Here, there is meaningful communication in the disclosure equilibrium with the low type, but there is no meaningful communication either in the conflict-hiding equilibria or in the disclosure equilibrium with the high type. It follows that as long as $p > 0$ the disclosure equilibrium will Pareto dominate the conflict-hiding one for all utility functions. In this section, we show that even in this case it's not true that nondisclosure necessarily leads to lower welfare. The reason is that there might exist equilibria other than the conflict-hiding ones under nondisclosure. To investigate this, we turn to the other possible type of equilibria – the partially conflict-revealing ones.

An equilibrium message can reveal the expert's type only if there exist states for one type such that the induced action is preferred to all other equilibrium actions but there do not exist such states for the other. Therefore, the high (low) type can reveal her type only if her message induces an action that is higher (lower) than any other equilibrium action, otherwise the other type would have an incentive to also send this message. Our next theorem shows that only the low type can be revealed through cheap-talk. Furthermore, the only actions that the low type can induce without the high type wanting to imitate her are sufficiently low actions.

Theorem 12 *For all distributions of the conflict, no equilibrium reveals the conflict of the high type and no equilibrium reveals the conflict of the low type for $s > d$.*

It follows from this theorem that conflict revelation is necessarily partial and that any message revealing the conflict has to contain information about s. While the high type cannot be revealed, there can be several conflict-revealing messages for the low type. In fact when $b_l = 0$, as in Example 4, there can be an infinite number of such messages. In addition, a corollary of the above theorem is that when $b_h = -b_l$ there can only be conflict-hiding equilibria.

The above theorem establishes only necessary but not sufficient conditions for conflict revelation. To see that conflict revelation might not be possible even for the low type at $s < d$ consider the following example:

Example 13 *Suppose players have quadratic preferences and $(p, b_h, b_l) = (\frac{1}{2}, \frac{1}{5}, \frac{1}{6})$.*

Theorem 12 implies that conflict revelation is only possible if $s < d = \frac{1}{30}$ *and if* $b = b_l$*. Since* $b_l = 1/6 > \frac{1}{30}$ *there can be at most one conflict-revealing message. Let's denote this action by* y_1*. It has to be true that* $y_1 < \frac{d}{2} = \frac{1}{60}$*. Since at* $s = 2y_1$ *the low type has to be indifferent between* y_1 *and* y_2*, it follows that* $y_2 = 3y_1 + 2b_l \in [\frac{20}{60}, \frac{23}{60}]$*. Suppose* y_2 *is the only other action induced. This means that the high type induces* y_2 *for all* s*, but then* $y_2 > \frac{1}{2} > \frac{23}{60}$*, a contradiction. Alternatively, suppose that there is a third action* y_3*. Then* $2a_2^l + 2b_l = y_3 + y_2$ *has to be true in equilibrium. The lower bound for* y_3 *occurs when* $y_1 \to 0$*, but then* $y_2 \to \frac{20}{60}$ *and* $a_2^l \to 0.68291$ *and as a result the lower bound on* y_3 *is* 1.3658*, which is impossible since the rationalizable action space is the unit interval.*

This example shows a reason why a partially conflict-revealing equilibrium might not exist. Intuitively, if the two types want to distort the decision maker's action to the same direction and similar extents, then low type can only reveal her type for low values of the state. But since the low type only wants to reveal her type if this gives her higher expected utility, the lowest commonly induced action cannot contain too much noise and hence cannot be far from the highest action induced by a conflict-revealing message. On the other hand, in order for the expert to be credible the increase in noise from a lower action to a higher one has to be sufficiently large and this increment becomes larger as the conflict of the low type increases. If these two forces are in conflict, as the above example suggests, there might not be room for conflict-revelation.

When d is large and there is uncertainty about the *direction* of the conflict, then the same logic is suggestive about a surprising property of partially conflict-revealing equilibria. Since the low type would like to induce lower actions and the high type higher ones than the state, there is more room for actions induced by only the low type. Furthermore, when b_l is negative then for the expert to remain indifferent, higher states have to be communicated with *less* noise. This means that the conflict-revealing message is credible if it is *less* precise than the first conflict-hiding message. When $b_l < 0 < b_h$ and $N(1, b_h, b_l) = N(0, b_h, b_l) = 1$ then partially conflict-revealing equilibria might be the only way to transmit information. The following example illustrates this possibility.

Example 14 *Suppose $b_l = -1/4$, $b_h = 1/3$, and $p = 1/2$ and assume the decision maker has quadratic preferences. Then, the following outcome is an equilibrium. There are two actions $y_1 = 0.38763$ and $y_2 = 0.66289$. The high type always induces action y_2; the low type induces action y_1 when $s \in [0, 2y_1]$ and action y_2 when $s \in (2y_1, 1]$.*

Note that in this example the expert can never be informative if the conflict is common knowledge. It is the combination of hiding the conflict for certain values of s and revealing it for others that greatly improves communication. Simple calculation shows that both players are better off than in the unique babbling equilibrium under disclosure. This phenomenon is not limited to the above parameter values and we conjecture that it is true in a much wider range of contexts. In fact, in the limiting case of partially conflict-revealing equilibrium when $b_l = -b_h$ and $p = 1/2$ there is always an informative nondisclosure equilibrium with the structure described in Lemma 3 in Section 3 as long as $b < 1/2$. In contrast, informative equilibria are only possible in the disclosure case if $b < 1/4$.

7 Conclusion

In this paper, we considered a model of strategic information transmission in which the decision maker is uncertain about the conflict of interest of the expert. We found that in a wide range of scenarios, it is not beneficial to the decision maker to have the conflict of interest mandatorily disclosed. We identified scenarios in which the expert is better off under nondisclosure, before or after she learns her conflict of interest. Our result shows that mandatory disclosure of conflicting interests might hurt rather than help parties to reach their desired outcomes because it often leads to less efficient communication. If the expert is not forced to verifiably disclose her bias but is allowed to hide or reveal her bias through equilibrium cheap-talk, then she can often be more precise and more credible. For all distributions of the conflict we consider, if there is meaningful communication under mandatory disclosure when the conflict is high and the players are sufficiently averse to inaccurate advice, then there is always a conflict-hiding equilibrium that allows for higher welfare than any equilibrium under mandatory disclosure. A surprising result that we did not fully explore in this paper is that, when there is uncertainty about the direction of the conflict, equilibrium revelation of the conflict may trigger meaningful communication even when mandatory disclosure prevents it. Future research could provide a more general characterization of such partially conflict-revealing equilibria.

As indicated by our discussion following Lemma 3, the equilibrium characterization and welfare analysis might be connected to variety of applications of the CS model such as legislative procedures (Gilligan and Krehbiel [16] and Krishna and Morgan [18]), organizational structure (Dessein [9]), and political lobbying (Grossman and Helpman [17]). Furthermore, although we consider a simple static framework our results might matter for the understanding of the quality of information transmission in settings that involve repeated interactions or multiple experts. Further research is needed to deepen our understanding of important issues such as the welfare effect of reputation concerns and the dynamics of conflict revelation.

Appendix: Detailed Derivations and Proofs

The following discussion (preceding the proof of Lemma 2) contains some definitions and conditions that are useful for the characterization of equilibrium. The online supplement has further details.

Let

$$y(l,p,d) \equiv \text{argmax}_y V(y,l,p,d) \equiv p \int_0^l U(y-s)\, ds + (1-p) \int_d^{l+d} U(y-s)\, ds. \quad (6)$$

We use $l > 0$ to denote the length of the two intervals in which the message is sent. When $l = 0$, the maximand becomes $pU(y) + (1-p)U(y-d)$.[27]

Recall that $y(a^h, \tilde{a}^h, a^l, \tilde{a}^l)$ is the action that maximizes the expected utility of the decision maker given that the high type sends the corresponding message in $[a^h, \tilde{a}^h]$ and the low type in $[a^l, \tilde{a}^l]$. By the distance aversion of U, we have

$$y(a^h, \tilde{a}^h, a^h + d, \tilde{a}^h + d) = y(\tilde{a}^h - a^h, p, d) + a. \quad (7)$$

Now we find the solution to Problem (6). First, since U is a strictly concave C^2 function,we have $\frac{\partial^2 V}{\partial y^2} < 0$. Therefore, the optimal solution $y(l,p,d)$ is the unique y that satisfies

$$0 = \frac{\partial V}{\partial y} = p \int_0^l U'(y-s)\, ds + (1-p) \int_d^{l+d} U'(y-s)\, ds, \quad (8)$$

[27] We have slightly abused notations with the use of functions y and V, in that their arguments change according to the optimization problem. However, we believe the meaning is clear from the context. In the rest of the Appendix, occurrences of the function y will be as defined in (6), unless otherwise noted.

if a solution to it exists. The following lemma affirms its existence and describes the range of $y(l,p,d)$.

Lemma 15 *Assume $d > 0$ and $p \in (0,1)$. There exists a unique value of y, $y(l,p,d)$, that solves (8). When $p > \frac{1}{2}$, $y(l,p,d) \in (\frac{l}{2}, \frac{l}{2} + \frac{d}{2})$. When $p = \frac{1}{2}$, $y(l,p,d) = \frac{l}{2} + \frac{d}{2}$. When $p < \frac{1}{2}$, $y(l,p,d) \in (\frac{l}{2} + \frac{d}{2}, \frac{l}{2} + d)$.*

PROOF. [***Lemma 15, Page 37***] The proof is by inspecting (8) and is omitted here. But, it is available in the online supplement.

PROOF. [***Lemma 2, Page 15***] The proof utilizes similar ideas to those in [21]. It can be found in the online supplement.

The following (preceding the proof of Theorem 5) further characterizes the conflict hiding equilibrium for the general setup.

Using (1) and (2), we obtain

$$
\begin{aligned}
2(a_1 + b_h) &= a_1 + y(a_2 - a_1, p, d) + y_1(a_1, p, d), \\
2(a_i + b_h) &= a_{i-1} + a_i + y(a_{i+1} - a_i, p, d) + y(a_i - a_{i-1}, p, d), \\
& \qquad i = 2, \ldots, n-2, \\
2(a_{n-1} + b_h) &= a_{n-2} + y(a_{n-1} - a_{n-2}, p, d) + y_n(a_{n-1}, p, d),
\end{aligned} \tag{9}
$$

where $y_1(a_1, p, d)$ is the action y that maximizes[28]

$$
\begin{aligned}
V(y, 0, a_1, p, d) \equiv\, & pE(U(y-s)|\ s \text{ uniform between 0 and } a_1) \\
& +(1-p)E(U(y-s)|\ s \text{ uniform between 0 and } a_1 + d),
\end{aligned}
$$

[28] The value of $V(y, 0, a_1, p, d)$ is not written in integration form because we allow a_1 to be negative. Thus, comparisons between the various bounds depend on the value of a_1. The case is similar for $V(y, a_{n-1}, 1, p, d)$ below.

and $y_n(a_{n-1}, p, d)$ is the y that maximizes

$$\begin{aligned} V(y, a_{n-1}, 1, p, d) \equiv\, & pE(U(y-s)|\ s \text{ uniform between } a_{n-1} \text{ and } 1) \\ & +(1-p)E(U(y-s)|\ s \text{ uniform between } a_{n-1}+d \text{ and } 1). \end{aligned}$$

Let us define the function $\delta(\cdot)$ as

$$\delta(a, p, b_h, d) \equiv a + b_h - y_1(a, p, d).$$

Observe that when $a = 0$ or $-d$, $\delta(0, p, b_h, d) = v$. Also, since $y_1(a, p, d) \in [\frac{a}{2}, \frac{a+d}{2}]$, we have

$$\delta(a, p, b_h, d) \in [\frac{a}{2} + v, \frac{a}{2} + b_h]. \tag{10}$$

Observe that $y_n(a_{n-1}, p, d) - 1 = y_1(a_{n-1} - 1, p, d)$. Therefore,

$$(a_{n-1} + b_h) - y_n(a_{n-1}, p, d) = \delta(a_{n-1} - 1, p, d).$$

With the introduction of the function δ, we may rewrite (9) as

$$\begin{aligned} \delta(a_1, p, b_h, d) &= y(a_2 - a_1, p, d) - b_h \\ (a_i - a_{i-1}) - [y(a_i - a_{i-1}, p, d) - b_h] &= y(a_{i+1} - a_i, p, d) - b_h \\ & \quad i = 2, \ldots, n-2, \\ (a_{n-1} - a_{n-2}) - [y(a_{n-1} - a_{n-2}, p, d) - b_h] &= -\delta(a_{n-1} - 1, p, d). \end{aligned} \tag{11}$$

The following lemma concerns properties of the functions δ and y.

Lemma 16 *The functions δ (with arguments a, p, b_h, and d) and y (with arguments l, p, and d) have the following properties:*

(1) δ is continuously differentiable; for $a \geq 0$ or $a \leq -d$, $\frac{\partial \delta}{\partial a} \geq 0$ and $= 0$ only when $a = 0$ or $-d$; when $a \geq 0$ or $a \leq -d$, $\frac{\partial \delta}{\partial p} \geq 0$ and $= 0$ only when $a = 0$ or $a = -d$.

(2) y is continuously differentiable; $0 < \frac{\partial y}{\partial l} < 1$*;* $\frac{\partial y}{\partial p} < 0$*;* $\frac{\partial y}{\partial d} \geq 0$ *and* $= 0$ *only when* $l = 0$*.*

PROOF. [***Lemma 16, Page 38***] The proof is by applying the implicit function theorem to the first order conditions that determine $y(\cdot)$ and $\delta(\cdot)$ and is omitted here. But, it is available in the online supplement.

PROOF. [***Theorem 5, Page 20***] First, we claim when $p \geq \frac{1}{2}$ or $b_l \geq 0$,

$$\begin{aligned} y(0, p, d) - b_h &\leq v, & (12)\\ l - [y(l, p, d) - b_h] &\geq y(l, p, d) - b_h, & (13) \end{aligned}$$

The proof of these claims follows from Lemma 15. First,

$$\begin{aligned} y(0, p, d) - b_h &\leq \frac{d}{2} - b_h = -v \leq v \quad \text{when } p \geq \frac{1}{2},\\ y(0, p, d) - b_h &\leq d - b_h = -b_l \leq v \quad \text{when } b_l \geq 0, \end{aligned}$$

which gives us (12). Second,

$$\begin{aligned} y(l, p, d) - b_h &\leq \frac{l}{2} + \frac{d}{2} - b_h = \frac{l}{2} - v \quad \text{when } p \geq \frac{1}{2},\\ y(l, p, d) - b_h &\leq \frac{l}{2} + d - b_h = \frac{l}{2} - b_l \quad \text{when } b_l \geq 0, \end{aligned}$$

which implies (13) in both cases. Together with (11-i) and the fact $\frac{\partial y}{\partial l} > 0$ from Lemma 16, it implies for $n \geq 4$,

$$a_{i+1} - a_i \geq a_i - a_{i-1}, \qquad (14)$$

where $i = 2, \ldots, n-2$.

By (12) and (13) we may specify an arbitrary $a_1 \geq 0$, and solve (11) forward up to (11-(n-1)), and be assured that a_i is increasing with its index. Note that in order for an n-action equilibrium to exist, there must exist an $a_1 \geq 0$ such that the forward solution satisfies (11-(n-1)). We claim that in order for an $n-$action equilibrium to exist, it is necessary and sufficient that when $a_1 = 0$ the forward solution satisfies

(i) $a_{n-1} \leq 1 - d$;

(ii) the LHS of (11-(n-1)) is less than or equal to the RHS of (11-(n-1)).

The following chain of argument shows that the LHS of (11-(n-1)) is increasing in $a_1 \geq 0$.

$$\begin{aligned}
a_1 \uparrow ((11\text{-}1), \delta \text{ is increasing in } a \geq 0) &\Rightarrow y(a_2 - a_1, p, d) \uparrow \left(\frac{\partial y}{\partial l} > 0\right) \\
\Rightarrow a_2 - a_1 \uparrow \left(1 - \frac{\partial y}{\partial l} > 0\right) &\Rightarrow (a_2 - a_1) - (y(a_2 - a_1, p, d) - b_h) \uparrow ((11\text{-}2)) \\
\Rightarrow y(a_3 - a_2, p, d) \uparrow \left(\frac{\partial y}{\partial l} > 0\right) &\Rightarrow \ldots \Rightarrow a_{n-1} - a_{n-2} \uparrow \left(1 - \frac{\partial y}{\partial l} > 0\right) \\
&\Rightarrow (a_{n-1} - a_{n-2}) - (y(a_{n-1} - a_{n-2}, p, d) - b_h) \uparrow .
\end{aligned}$$

In the above process, a statement in the parentheses preceding each "$\Rightarrow$" are needed to justify the statement that follows the "$\Rightarrow$". Note the above argument also shows a_{n-1} is increasing in a_1 as we solve (11) forward since

$$a_{n-1} = a_1 + (a_2 - a_1) + \ldots + (a_{n-1} - a_{n-2}).$$

By Lemma 16, $\frac{\partial \delta}{\partial a} \geq 0$ when $a \leq -d$ or $a \geq 0$, and $\frac{\partial \delta}{\partial a} = 0$ only when $a = -d$ or $a = 0$. Thus, the RHS of (11-$(n-1)$) is decreasing in a_1 when we solve (11) forward, as long as the resulting a_{n-1} is strictly less than $1 - d$.

Note we have shown in the forward solution, the LHS of (11-(n-1)) is increasing in a_1, a_{n-1} is increasing in a_1, and the RHS of (11-$(n-1)$) is decreasing in a_1. Also, they vary continuously with a_1 by the continuity of all the functions involved in the difference equations.

If $n \geq 3$, for $a_1 = 0$,

$$\text{LHS of (11-1)} = \delta(0, p, b_h, d) = v = y(a_2 - a_1, p, d) - b_h = \text{RHS of (11-1)}.$$

Now, we solve (11) forward. By (13), we have for $n \geq 3$ and $a_1 = 0$, LHS of (11-$(n-1)$) $\geq$ v. Since the LHS of (11-$(n-1)$) is increasing in a_1, it is also true for $a_1 \geq 0$.

Since a_{n-1} is continuously and strictly increasing in a_1, and a_{n-1} is greater than or equal to $1-d$ when $a_1 = 1-d$, condition (i) is equivalent to there existing $\bar{a}_1 \leq 1-d$ such that the forward solution starting at $a_1 = \bar{a}_1$ gives $a_{n-1} = 1-d$. This would imply the RHS of (11-$(n-1)$) is equal to $-v$ ($\delta(-d, \cdot) = v$), and hence

$$\text{LHS of (11-}(n-1)) \geq \text{RHS of (11-}(n-1)). \tag{15}$$

Given the above inequality, and the continuity of both sides of (11-$(n-1)$) in a_1, in order for there to exist an a_1 to equate both sides of (11-$(n-1)$), it is necessary and sufficient that condition (ii) holds. Also, the strict monotonicity of both sides of (11-$(n-1)$) in a_1 ensures the solution is unique.

If $n = 2$, there is only one equation in (11). Conditions (i) and (ii) entail

$$\begin{aligned} 0 &\leq 1-d \\ \delta(0, p, b_h, d) = v &\leq -\delta(0-1, p, b_h, d), \end{aligned}$$

where the first condition holds if $d \leq 1$ and the second condition ensures the existence of a unique a_1 that satisfies the lone equation.

Since fixing a_1, the LHS of (11-(n-1)) is larger for larger n, and the RHS of (11-(n-1)) is smaller for larger n. Therefore, if conditions (i) and (ii) hold for n, they also hold for all $n' < n$. That is, the existence of an n-action equilibrium implies the existence of a unique n'-action equilibrium for each $n' < n$.

When $b_h = -b_l < \frac{1}{2}$ and $p = \frac{1}{2}$, both (i) and (ii) are trivially satisfied, as for any n, at $a_1 = 0$, the forward solution of (11) yields $a_1 = \ldots = a_{n-1} = 0$.

Finally, we claim in equilibrium, $a_2 - a_1 \geq a_1$ and $1 - a_{n-1} \geq a_{n-1} - a_{n-2}$; furthermore, when $b_l \geq 0$, $a_2 - a_1 \geq a_1 + d$ and $1 - (a_{n-1} + d) \geq a_{n-1} - a_{n-2}$. These statements, combined with (14), constitute the last statement in the Theorem.[29] Here, we only show $1-(a_{n-1}+d) \geq a_{n-1} - a_{n-2}$ when $b_l \geq 0$, but the other statements can be proved analogously. First, $-\delta(a_{n-1} - 1, p, b_h, d) \leq -\frac{a_{n-1}-1}{2} - v$ by (10). Second, $(a_{n-1} - a_{n-2}) - [y(a_{n-1} - a_{n-2}, p, d) - b_h] \geq \frac{a_{n-1}-a_{n-2}}{2} + b_l$. Last, $b_l \geq 0$ implies $d \leq 2v$. These three facts, together with (11-(n-1)), imply $1 - a_{n-1} - d \geq a_{n-1} - a_{n-2}$.

PROOF. [***Corollary 6, Page 22***] Suppose $p > p'$. Let a_i and a_i' be the corresponding equilibrium cutoff points. We want to prove $a_i < a_i'$ for all $i = 1, ..., n-1$.

First we show $a_1 < a_1'$. The proof is by contradiction. Suppose $a_1 \geq a_1'$. In the chain of argument below, we will make use of Lemma 16 without explicitly referring to it. We have

[29] The corresponding statements in the case $n = 2$, though slightly different, can be shown similarly.

$$\begin{aligned}
a_1 \geq a_1' \geq 0, p > p' \Rightarrow \delta(a_1, p, b_h, d) \geq \delta(a_1', p', b_h, d) \,((11\text{-}1)) \\
\Rightarrow y(a_2 - a_1, p, d) > y(a_2' - a_1', p', d) \Rightarrow y(a_2 - a_1, p, d) > y(a_2' - a_1', p, d) \\
\Rightarrow a_2 - a_1 > a_2' - a_1' \Rightarrow (a_2 - a_1) - [y(a_2 - a_1, p, d) - b_h] > (a_2' - a_1') - [y(a_2' - a_1', p, d) - b_h] \\
\Rightarrow (a_2 - a_1) - [y(a_2 - a_1, p, d) - b_h] > (a_2' - a_1') - [y(a_2' - a_1', p', d) - b_h] \,((11\text{-}2)) \\
\Rightarrow y(a_3 - a_2, p, d) > y(a_3' - a_2', p', d) \Rightarrow \ldots \\
\Rightarrow \text{LHS of } (11\text{-}(n-1)) \text{ of } p > \text{LHS of } (11\text{-}(n-1)) \text{ of } p'.
\end{aligned}$$

In addition, the above argument implies $a_{n-1} = a_1 + (a_2 - a_1) + \ldots + (a_{n-1} - a_{n-2}) > a_1' + (a_2' - a_1') + \ldots + (a_{n-1}' - a_{n-2}') = a_{n-1}'$. In equilibrium, a_{n-1} and a_{n-1}' must be both not more than $1 - d$. Thus, by Part 1 of Lemma 16, we conclude

$$\text{RHS of } (11\text{-}(n-1)) \text{ of } p = -\delta(a_{n-1} - 1, p, d) < -\delta(a_{n-1}' - 1, p', d) = \text{RHS of } (11\text{-}(n-1)) \text{ of } p'.$$

The above two statements directly contradict (11-$(n-1)$). Thus, $a_1 < a_1'$.

That $a_i < a_i'$ can be proved by induction and contradiction. Let j be the smallest index i such that $a_i \geq a_i'$. Then,

$$a_j - a_{j-1} > a_j' - a_{j-1}'.$$

With a similar chain of argument to the one above, we arrive at a contradiction. Therefore, $a_i < a_i'$ for all $i = 1, \ldots, n-1$.

Note that the above proof applies to the case $n \geq 3$. In the case $n = 2$, (11) has only one equation, and the proof is straightforward there.

PROOF. [***Corollary 7, Page 23***] We want to show that if $p > p'$, and there exists an $n-$interval equilibrium for p, then there exists an $n-$interval equilibrium for p'.

In the proof of Theorem 5, we have shown that an n-action equilibrium exists if and only if when $a_1 = 0$ the forward solution satisfies

(i) $a_{n-1} \leq 1 - d$;

(ii) the LHS of (11-(n-1)) is less than or equal to the RHS of (11-(n-1)).

Let us fix $a_1 = a_1' = 0$. When $n \geq 3$, using the same argument as that in the proof of Corollary 6, we have

$$\begin{aligned} a_{n-1} &> a'_{n-1}, \\ \text{LHS of (11-}(n-1)\text{) of } p &> \text{LHS of (11-}(n-1)\text{) of } p', \\ \text{RHS of (11-}(n-1)\text{) of } p &< \text{RHS of (11-}(n-1)\text{) of } p'. \end{aligned}$$

Since there exists an $n-$interval equilibrium for p, Conditions (i) and (ii) hold for p. The inequalities above then imply that they also hold for p'. The case $n = 2$ is straightforward.

PROOF. [***Lemma 8, Page 24***] Note

$$V_n^D(p, b) \equiv \sum_{i=1}^{n} p \int_{a_{i-1}^h}^{a_i^h} U(y_i - s)ds + (1-p) \int_{a_{i-1}^l}^{a_i^l} U(y_i - s)\, ds.$$

By the Envelope Theorem, we may ignore the dependence of V_n^D on p through y_i as the y_i satisfy the relevant first order conditions. Thus,

$$\frac{\partial V_N^D(p, b)}{\partial p} = A_1 + A_2, \tag{16}$$

where

$$A_1 = \sum_{i=1}^{n-1} p[U(y_i - a_i^h) - U(y_{i+1} - a_i^h)]\frac{\partial a_i^h}{\partial p} + (1-p)[U(y_i - a_i^l) - U(y_{i+1} - a_i^l)]\frac{\partial a_i^l}{\partial p}$$
$$A_2 = \sum_{i=1}^{n} \int_{a_{i-1}^h}^{a_i^h} U(y_i - s)ds - \int_{a_{i-1}^l}^{a_i^l} U(y_i - s)\, ds.$$

The expression A_1 can be viewed as the change in the decision maker's expected payoff due to shifts in cutoff points. The expression A_2 is the difference between the decision maker's expected payoff from consulting a high type and that from consulting a low type. We claim:

(1) $A_1 \leq 0$ with equality if and only if $p = 1/2$ and $b_h = -b_l$;

(2) $A_2 \leq 0$ with equality if and only if $b_h = -b_l$.

In the online supplement, we provide detailed proofs of these two statements. Here, we provide intuition for the signs of each term. Note that we are fixing the number of equilibrium actions/messages. Hence, the welfare implications are due to changes in the balancedness of messages only.

First, consider the term A_1. As p increases from $1/2$ to 1, all cutoff points are shifted to the left, which makes a right-biased expert's messages less balanced (including the high type and the low type when $b_l \geq 0$) but a left-biased expert's messages more so (the low type when $b_l < 0$). When $b_l \geq 0$, the total effect is to make both types' messages less balanced. When $b_l < 0$ and $p \geq 1/2$, the balancedness-enhancing effect for the high type dominates the balancedness-reducing effect for the low type. Thus, the total effect is also to reduce the balancedness of all messages. Due to the strict concavity of the decision maker's preferences, he becomes worse off.

Now, consider the term A_2. It reflects the fact that as p increases, it becomes more likely that the decision maker is consulting the high type. Since the high type's messages are less balanced, the decision receives a lower expected payoff from consulting the high type. Therefore, he becomes worse off as more weight is put on the high type.

Given the above two statements about A_1 and A_2, we have

$$\frac{\partial V_N^D(p,b)}{\partial p} = A_1 + A_2 \leq 0,$$

when $p \geq \frac{1}{2}$ and $b_l \geq 0$, with equality only when $p = \frac{1}{2}$ and $b_h = -b_l$.

PROOF. [***Theorem 9, Page 25***] By Corollary 6, the cutoff points under nondisclosure, $\{a_i\}_{i=1}^{n-1}$, are strictly decreasing in p. Let us use $\{a_i^{dh}\}_{i=1}^{n-1}$ ($\{a_i^{dl}\}_{i=1}^{n-1}$) to denote the cutoff points for the high (low) type under disclosure. Then, a_i is equal to a_i^{dh} when $p = 1$, and $a_i + d = a_i^{dl}$ when $p = 0$. Therefore, the corollary implies that for all $i = 1, ..., n-1$, (i) for $p \geq \frac{1}{2}$, $a_i^{dh} \leq a_i$; (ii) when $b_l \geq 0$, for all $p \in [0,1]$, $a_i + d \leq a_i^{dl}$.

The following lemma is crucial to the proof of this theorem.

Lemma 17 *Let $\{l_i\}_{i=1}^n$ and $\{l_i'\}_{i=1}^n$ be two sets of nonnegative numbers satisfying:*

(1) $\sum_{i=1}^n l_i = \sum_{i=1}^n l_i' = 1$;

(2) $l_{i+1} \geq l_i$ and $l_{i+1}' \geq l_i'$ for all $i = 1, ..., n-1$;

(3) $l_1 \geq l_1'$, $l_n \leq l_n'$ and $l_{i+1} - l_i \leq l_{i+1}' - l_i'$ for all $i = 1, ..., n-1$;

Let $f : \mathbb{R} \to \mathbb{R}$ be a concave function. Then,

$$\sum_{i=1}^n \int_0^{l_i} f(\frac{l_i}{2} - s)ds \geq \sum_{i=1}^n \int_0^{l_i'} f(\frac{l_i'}{2} - s)ds.$$

If f is strictly concave and $l_i \neq l_i'$ for some i, then the inequality is strict.

The intuition of the lemma is as follows. Let there be two size-n partitions of the interval $[0,1]$, $\{l_i\}$ and $\{l_i'\}$. The conditions state that $\{l_i\}$ is more balanced than $\{l_i'\}$. If an agent's payoff is measured by a concave function of the deviation of the state from the mean of each interval, as is the case for the payoffs of the decision maker and the expert in our model, then the agent has higher expected utility from the more balanced partition, $\{l_i\}$.

Now, let $l_i = a_i - a_{i-1}$ and $l_i' = a_i^{dh} - a_{i-1}^{dh}$. First, they clearly satisfy Condition 1 of Lemma 17. To see they satisfy Condition 2, we may refer to (5) (as a_i^{dh} satisfies (5) with $b = b_h$) and Theorem 5. Finally, they satisfy Condition 3 by Corollary 6 and (4). Let $f(x) = \tilde{U}(x - b)$. Since $\tilde{U}$ is concave, we may apply Lemma 17 and conclude

$$\sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} \tilde{U}(\frac{a_i + a_{i-1}}{2} - (s+b))\, ds = \sum_{i=1}^{n} \int_{0}^{l_i} \tilde{U}((\frac{l_i}{2} - s) - b)\, ds$$
$$\geq \sum_{i=1}^{n} \int_{0}^{l_i'} \tilde{U}((\frac{l_i'}{2} - s) - b)\, ds = \sum_{i=1}^{n} \int_{a_{i-1}^{dh}}^{a_i^{dh}} \tilde{U}(\frac{a_i^{dh} + a_{i-1}^{dh}}{2} - (s+b))\, ds.$$

In other words, for the partition under nondisclosure, the high type would have been better off than under disclosure, even if the decision maker takes action $\frac{a_i+a_{i-1}}{2}$ after receiving m_i. Observe that the action $\frac{a_i+a_{i-1}}{2} + b_h$ would have maximized the expert's expected payoff. Furthermore, $y_i \in (\frac{a_i+a_{i-1}}{2}, \frac{a_i+a_{i-1}}{2} + b_h)$ for $p \in (0,1)$. Thus, the high type's expected payoff in equilibrium under nondisclosure is even higher. Our discussion so far is based on a fixed n. In addition, $N(p, b_h, b_l) \geq N(1, b_h, b_l)$ by Corollary 7. Hence, the high type's highest payoff achievable under nondisclosure is strictly higher than under disclosure.

We now turn to the case $b_l \geq 0$ and show that the low type prefers to have her bias disclosed. Let $l_i = a_i^{dl} - a_{i-1}^{dl}$ and $l_i' = a_i^l - a_{i-1}^l$ (note that $a_0^l = 0$, $a_i^l = a_i + d$ for $i = 1, \ldots, n-1$, and $a_n^l = 1$). Following similar procedures to the one for the high type, we can show they satisfy the three conditions of Lemma 17. Finally, Corollary 7 implies $N(p, b_h, b_l) \leq N(0, b_h, b_l)$ when $b_l \geq 0$. Thus, we conclude the low type's highest payoff achievable under nondisclosure is strictly lower than that under disclosure.

PROOF. [***Theorem 11, Page 29***]

Consider a game $\Gamma = \{U, \tilde{U}, p, b_h, b_l\}$ and fix a conflict-hiding equilibrium with $N = N(1, b_h, b_l)$ actions. From Corollary 7 it follows that such an equilibrium always exists since $N(p, b_h, b_l) \geq N(1, b_h, b_l)$. Let $Y_j(\cdot)$ be the equilibrium mapping from states into actions as the result of a type-j expert's strategy in the nondisclosure conflict-hiding equilibrium. Let $Y_{dj}(\cdot)$ (d for disclosure) be its counterpart in the disclosure equilibrium.

Part 1 In this part, we show that there exists a sufficiently concave utility function $TU : \mathbb{R} \to \mathbb{R}$ and $z^* > 0$ such that (i) $|Y_j(s) - s| \leq z^*$ for all $s \in [0, 1]$ and $j = h, l$; (ii) but, there exist j and a set of non-zero measure $s \in [0, 1]$ such that $|Y_{dj}(s) - s| > z^*$. Similarly, for the expert we show that given TU there exists a $z^{**} > 0$ such that (i) $|Y_j(s) - s - b_j| \leq z^{**}$ for all $s \in [0, 1]$ and $j = h, l$; (ii) but, there exist j and a set of non-zero measure s such that $|Y_{dj}(s) - s - b_j| > z^{**}$.

Take the conflict-hiding equilibrium and consider the case when $\beta = b_h$. First note that $0 < |a_{i-1} - y_{i-1}| < y_i - a_{i-1}$ for all $i > 1$ since $a_{i-1} = (y_{i-1} + y_i)/2 - b_h$.

Claim 18 *$y_{i+1} - a_i > y_i - a_{i-1} > 0$ and $\max_s |Y_h(s) - s| \leq y_N - a_{N-1}$.*

Let us substitute $y_{i+1} = 2a_i + 2b_h - y_i$ into $y_{i+1} - a_i > y_i - a_{i-1}$ so that we get $y_i < \frac{a_{i-1}+a_i}{2} + b_h$. This inequality holds because if $b_l > 0$ then $b_h - d = b_l > 0$ and if $p \geq 1/2$ then $b_h - d/2 = (b_h + b_l)/2 > 0$. Note that for $1 < i \leq N$, $|a_{i-1} - y_{i-1}| < y_i - a_{i-1}$ since $2a_{i-1} + 2b_h = y_i + y_{i-1}$ where $b_h > 0$. From the fact that $y_{i+1} - a_i$ is increasing in i and that $|a_{i-1} - y_{i-1}| < |y_i - a_{i-1}|$ it follows that the maximum is either $y_N - a_{N-1}$ or $1 - y_N$. Since $y_N > (1 + a_{N-1})/2$ the maximum equals $y_N - a_{N-1}$.

Claim 19 $\max_s |Y_l(s) - s| \leq y_N - a_{N-1}$.

Note that $0 < y_i - a^l_{i-1} < a^l_i - y_i$ for all $i < N$ since $y_i < \frac{a^l_{i-1}+a^l_i}{2}$. Equation (**??**) in the proof of Lemma 8 implies $a^l_i - y_i > a^l_{i-1} - y_{i-1} > 0$ for all $i = 2, \ldots, N-1$. Given that $a^l_i - y_i > a^l_{i-1} - y_{i-1}$ and that $y_i - a^l_{i-1} < a^l_i - y_i$ it follows that the maximum is achieved either for $a^l_{N-1} - y_{N-1}$ or for $1 - y_N$. Note that $a^l_{N-1} - y_{N-1} = a_{N-1} + d - y_{N-1} = y_N - a_{N-1} - 2b_h + d = y_N - a_{N-1} - (b_h + b_l)$ $< y_N - a_{N-1}$. Given the proof of Claim 18, it follows that $\max_s |Y_l(s) - s| <$ $y_N - a_{N-1}$.

Claim 20 *There exists a sufficiently concave utility function TU such that* $y_N - a_{N-1} < (1 - a^{dh}_{N-1})/2$.

Consider a concave transformation of U denoted by TU. Assume that TU satisfies our assumptions from Section 3 and $dTU/dx > dU/dx > 0$ for $x < 0$ and $dTU/dx < dU/dx < 0$ for $x > 0$. Now, we demonstrate that there is a sufficiently concave transformation, TU, which will move y_N sufficiently close to a_{N-1} such that $y_N - a_{N-1} < (1 - a^{dh}_{N-1})/2$ is satisfied. Consider the following condition that determines y_N.

$$\frac{\partial V^D}{\partial y} = p \int_0^l U'(y-s)ds + (1-p) \int_d^l U'(y-s)ds = 0 \tag{17}$$

and consider a concave transformation TU. Then given this transformed utility function the equilibrium condition is satisfied if

$$p\int_0^l TU'(y-s)ds + (1-p)\int_d^l TU'(y-s)ds = 0. \tag{18}$$

Let y^* satisfy the equilibrium condition with U and consider the following two quantities:

$$L = p\int_0^{y^*} U'(y^*-s)ds+(1-p)\int_d^{y^*} U'(y^*-s)ds-p\int_0^{y^*} TU'(y^*-s)ds-(1-p)\int_d^{y^*} TU'(y^*-s)ds$$

and

$$R = \int_{y^*}^l U'(y^*-s)ds - \int_{y^*}^l TU'(y^*-s)ds$$

Since $|dTU/dx| > |dU/dx|$ it follows that $L < 0 < R$. Furthermore, given that TU is a concave transformation of U and $l/2 < y_N < (l+d)/2$, it follows that there is a function TU such that $|L| > |R|$, which implies $L + R > 0$. Using Equation (17), we know the left hand side of Equation (18) is equal to $-(L + R)$, hence negative. The strict concavity of TU implies that left hand side of Equation (18) is decreasing in y, hence y_N under TU has to be lower than y^* since Equation (18) has to be satisfied in equilibrium. Given the assumption that U and TU are continuously differentiable, it is easy to see that for every $\varepsilon > 0$ there exists TU such that $\mid y_N - a_{N-1} \mid < (1+a_{N-1})/2+\varepsilon$. Since $|y_N - a_{N-1}| < \max|Y_{dh}(s) - s| = \left|1 - a_{N-1}^{dh}\right|/2$ this proves the inequality for the decision maker.

Note that in the game $\Gamma = \{TU, \tilde{U}, p, b_h, b_l\}$ the number of conflict-hiding equilibria might be different than in $\Gamma = \{U, \tilde{U}, p, b_h, b_l\}$ but $N(p, b_h, b_l) \geq N(1, b_h, b_l)$ is true in both cases. Also note that the disclosure equilibrium strategies in $\Gamma = \{TU, \tilde{U}, p, b_h, b_l\}$ and $\Gamma = \{U, \tilde{U}, p, b_h, b_l\}$ are the same. These equilibria are fully determined by (p, b_h, b_l) as long as the player's preferences satisfy our assumptions from Section 3.

Claim 21 $\max_{j,s}|Y_j(s)-s-b_j| \leq \frac{1-a_{N-1}}{2}+b_h < \frac{1-a^{dh}_{N-1}}{2}+b_h$.

It follows from CS that in the disclosure equilibrium the maximum of $|Y_{dj}(s)-s-b_j|$ is achieved for $s=1$ and $j=h$ and is equal to $\frac{1-a^{dh}_{N-1}}{2}+b_h$.

To obtain an upper bound on $\max_{j,s}|Y_j(s)-s-b_j|$ consider first that $\beta = b_h$. In this case $|Y_j(s)-s-b_h|$ is maximal for either $s=a_{N-1}$ and equals $a_{N-1}-y_{N-1}+b_h>0$ or for $s=1$ and equals $1-y_N+b_h$. This follows from the fact $|Y_j(s)-s-b_h|$ is always maximal at a cut-off point and that $|y_i-a_i|$ is increasing in i. An upper bound on $\max\{a_{N-1}-y_{N-1}+b_h, 1-y_N+b_h\}$ is obtained when $y_N \to \frac{1+a_{N-1}}{2}$ and equals $\frac{1-a_{N-1}}{2}+b_h$. Combining this inequality with $a^{dh}_{N-1}<a_{N-1}$ from Corollary 6 it follows that $\max_s|Y_j(s)-s-b_h| < \max_{j,s}|y_d(s)-s-b_j|$ as long as $p<1$.

Consider now the case when $\beta=b_l$ and $b_l>0$. Since $\left|y_i-a^l_i-b_l\right| = \left|y_{i+1}-a^l_i-b_l\right|$ for $1<i<N$ and $\left|y_i-a^l_i\right|$ is increasing in i, the maximum is achieved for $a^l_{N-1}-y_{N-1}+b_l = a_{N-1}-y_{N-1}+b_h$. Finally, consider $\beta=b_l$ and $b_l<0$. Here the maximum is achieved either for $\left|y_N-a^l_{N-1}-b_l\right| = a_{N-1}-y_{N-1}+b_h$ or for $|1-y_N-b_l| < \frac{1-a_{N-1}}{2}-d/2-b_l = \frac{1-a_{N-1}}{2}-(b_h+b_l)/2 < \frac{1-a_{N-1}}{2}+b_h$.

Part 2. Let V^I $(I=D,E)$ denote the expected utility of the decision maker (expert) in the N-action nondisclosure equilibrium and let V^I_{dl} and V^I_{dh} $(I = D,E)$ denote their counterparts in the $N-$action disclosure equilibria with the low type and the high type respectively. If $V^I > pV^I_{dl}+(1-p)V^I_{dh}$ for a player, then we are done. So, assume that $V^I < pV^I_{dl}+(1-p)V^I_{dh}$. Consider TU such that the conditions that $y_N-a_{N-1} > (1-a^{dh}_{N-1})/2$ is already satisfied. As the proof of Part 1 shows such TU always exists. Since TU

is continuous, it is bounded on $[0, 1]$ and there exists $H \in (0, \infty)$ such that $pV_{dh}^D + (1-p)V_{dl}^D < V^D + H$. Now consider a function F which is continuously differentiable, concave around zero, and distance-averse. Let's define MU to be $MU(x) = TU(x)$ if $x \leq z^*$and $MU(x) = F(TU(x))$ if $x \geq z^*$, where F is continuously differentiable symmetric around zero and $MU(x)$ satisfies our assumptions from Section 3. Since $H < \infty$, and TU is distance-averse, we can choose F such that

$$\int_{\Phi^*} TU(x)Q(x \mid Y_{dh})dx - \int_{\Phi^*} MU(x)Q(x \mid Y_{dh})dx > H,$$

where $\Phi^* = \{x \geq z^*\}$ and with a slight abuse of notion for every x in Φ^*, $Q(x \mid Y_{dh}$ is the likelihood that the outcome of the game equals x given the mapping Y_{dh} from state to action of the N-action disclosure equilibrium. This proves the theorem for the decision maker. The same type of argument shows that there is a function G such that if $M\tilde{U}(x) = \tilde{U}(x)$ if $x \leq z^{**}$ and $M\tilde{U}(x) = G(\tilde{U}(x))$ if $x \geq z^{**}$ then $V^E > [(1-p)V_{dl}^E + pV_{dh}^E]$. Clearly given the properties of the integral, if the welfare result holds for MU or $M\tilde{U}(x)$ then given our discussion above on how y depends on the concavity of the decision maker's preferences, it holds for any concave transformation of these functions as long as the these new functions satisfy our assumptions on the utility functions of the players from Section 3.

PROOF. [***Theorem 12, Page 31***] To show that the high type cannot be revealed, suppose there is an action $y \in [0, 1]$ that is induced by the high type only. Then either $y < b_l$ (when $b_l > 0$) or $y > 1 + b_l$ (when $b_l < 0$). Otherwise, the low type would find it optimal to induce the action in some states. Let Y^* be the set of equilibrium actions.

The case $y < b_l$ is possible only if $b_l > 0$. In equilibrium there could be at most one action less than or equal to b_l. If there were two such actions, then by Lemma 1, both types would strictly prefer the higher action to the lower one in all states. Therefore, y is an isolated point in Y^*. Let y' be the lowest equilibrium action other than y, which is well-defined since Y^* is a closed set.[30] By Lemma 1, the set of states in which the high type induces y is then $[0, \frac{y+y'}{2} - b_h]$, which has to be nonempty. But the low type wants to induce y in $[0, \frac{y+y'}{2} - b_l]$, again nonempty since $b_l \leq b_h$, contradicting y being induced only by the high type.

Now, suppose $y > 1 + b_l$. This is possible only if $b_l < 0$. In equilibrium, the high type can induce at most one action higher than or equal to $1 + b_l$. The reason is as follows. In order for the decision maker to optimally choose to take an action higher than $1 + b_l$, the inducing message must be sometimes sent in states higher than $1 + b_l$. But in all states in the interval $[1 + b_l, 1]$, the high type prefers the highest such action to any other action. Note that it also implies y is the highest equilibrium action. Let y' be the second highest action in equilibrium. Again, it is well defined since Y^* is closed and y is not a limit point of Y^* (see the proof of Lemma 2). Let a be the cutoff point for the high type between y and y'. By Lemma 1, $a = \frac{y+y'}{2} - b_h$. In addition, $y = \frac{a+1}{2}$. Also, a cannot be lower than $1 - d$, since the low type would then want to induce y in $[a + d, 1]$. Since $y = \frac{a+1}{2}$ and $y' < y$, we have

$$a < \frac{a+1}{2} - b_h,$$
$$\text{or } a < 1 - 2b_h.$$

[30] The proof of this fact can be found in the proof of Lemma 2 in the online supplement.

This contradicts $a \geq 1 - d \geq 1 - 2b_h$. To conclude, there exist no equilibria that reveal the high type's bias.

Now, we show the low type can send conflict-revealing messages only for $s \leq d$. Let y be the maximal action that is induced only by the low type.[31] Then, $y \leq b_h$. Let y' be the minimal action induced by both types. Then, $y' \geq y$. Otherwise, in all states the high type would prefer y to y' as $\frac{y+y'}{2} - b_h < 0$, a contradiction. By Lemma 1, for the low type, the cutoff point between y and y' is $\frac{y+y'}{2} - b_l$. Unless it is less than $d = b_h - b_l$, we would have $\frac{y+y'}{2} - b_h \geq 0$, which makes it optimal for the high type to induce y in $[0, \frac{y+y'}{2} - b_h]$, contradicting y being induced only by the low type. Thus, the low type can only be revealed in states $s < d$.

[31] Take the supremum if no maximum exists. Similarly for the y' below.

References

[1] D. Austen-Smith, Information transmission in debate, Amer. J. Polit. Sci. 34 (1) (1990) 124–152.

[2] R. Bénabou, G. Laroque, Using privileged information to manipulate markets: Insiders, gurus, and credibility, Quart. J. Econ. 107 (3) (1992) 921–958.

[3] A. Blume, D. V. DeJong, Y.-G. Kim, G. B. Sprinkle, Experimental evidence on the evolution of meaning of messages in sender-receiver games, Amer. Econ. Rev. 88 (5) (1998) 1323–1340.

[4] H. Cai, J. T.-Y. Wang, Overcommunication in strategic information transmission games, Games Econ. Behav. 56 (1) (2006) 7–36.

[5] D. M. Cain, G. Loewenstein, D. A. Moore, The dirt on coming clean: Perverse effects of disclosing conflicts of interest, J. Legal Stud. 34 (1) (2005) 1–25.

[6] Y. Chen, N. Kartik, J. Sobel, On the robustness of informative cheap talk, mimeo (2007).

[7] V. Crawford, J. Sobel, Strategic information transmission, Econometrica 50 (6) (1982) 1431–1452.

[8] T. de Garidel-Thoron, M. Ottaviani, The economics of advice, mimeo (2000).

[9] W. Dessein, Authority and communication in organizations, Rev. Econ. Stud. 69 (2002) 811–838.

[10] V. Dimitrakas, Y. Sarafidis, Advice from an expert with unknown motives, mimeo, (2005) INSEAD.

[11] R. Durrett, Probability: Theory and Examples, 2nd ed., Duxbury Press, 1996.

[12] W. Dziuda, Strategic argumentation, mimeo, Princeton University.

[13] J. Farrell, Meaning and credibility in cheap-talk games, Games Econ. Behav. 5 (1993) 514–31.

[14] J. Farrell, R. Gibbons, Cheap talk with two audiences, Amer. Econ. Rev. 79 (5) (1989) 1211–1224.

[15] J. Farrell, M. Rabin, Cheap talk, J. Econ. Perspect. 10 (3) (1996) 103–118.

[16] T. Gilligan, K. Krehbiel, Asymmetric information and legislative rules with a heterogeneous committee, Amer. J. Polit. Sci. 33 (2) (1989) 459–490.

[17] G. M. Grossman, E. Helpman, Special Intereste Politics, MIT Press: Cambridge, 2001.

[18] V. Krishna, J. Morgan, Asymmetric information and legislative rules: Some amendments, Amer. Polit. Sci. Rev. 95 (2) (2001) 435–452.

[19] M. Li, To disclose or not to disclose? cheap talk with uncertain biases, mimeo, (2005), available at http://alcor.concordia.ca/˜mingli/research/disclosure.pdf.

[20] K. Madarász, Disclosing conflicts of interest, mimeo, (2005).

[21] J. Morgan, P. Stocken, An analysis of stock recommendations, RAND J. Econ. 34 (1) (2003) 183–203.

[22] S. Morris, Political correctness, J. Polit. Economy 109 (2001) 231–265.

[23] G. Moscarini, Competence implies credibility, Amer. Econ. Rev. 97 (1) (2007) 37–63.

[24] M. Ottaviani, P. N. Sørensen, Professional advice, J. Econ. Theory 126 (1) (2006) 120–142.

[25] J. Sobel, A theory of credibility, Rev. Econ. Stud. 52 (4) (1985) 557–573.

[26] J. Stiglitz, Corporate corruption: The conflict of interest driving US financial scandals and being replicated on a global scale, Guardian, July 4, 2002.