Mandatory disclosure of conflicts of interest: Good news or bad news?*

Ming Li† and Ting Liu‡

August 20, 2022

Preliminary and incomplete. Do not quote.

*We would like to thank (virtual) conference and seminar audiences at Asian Econometric Society Meeting 2021, Annual Meeting of the Austrian Economic Association 2021, China Econometric Society Meeting 2021, Communication and Persuasion Workshop Japan, CUHK-HKU-HKUST Economic Theory Workshop, GAMES 2020, Midwest Economic Theory Meeting–East Lansing, Society for Economic Design Conference 2022-Padua, Stony Brook International Game Theory Conference 2022, Düsseldorf Institute for Competition Economics (DICE), Stony Brook University, and especially Sandro Brusco, Chia-Hui Chen, Laura Doval, Hanming Fang, Junichiro Ishida, Kohei Kawamura, Yunan Li, Shintaro Miura, Arijit Mukherjee, Alexander Rasch, Hitoshi Sadakane, Marco Schwarz, Chris Wallace, Huanxing Yang, and Yishu Zeng for their very helpful comments and suggestions. All errors are our own. Li gratefully acknowledges financial support from SSHRC Grant 435-2020-1169, "The economics of strategic communication and persuasion: Theory, applications, and experiments."

†Concordia University, CIRANO, and CIREQ, 1455 Boulevard de Maisonneuve Ouest, Department of Economics, Concordia University, Montréal, Québec, Canada H3G 1M8. Email: ming.li@concordia.ca.

‡Economics Department and College of Business, Stony Brook University. Email: ting.liu@stonybrook.edu.

Abstract

We investigate the welfare effect of disclosure of conflict of interest when an expert advises a decision maker. In a model with verifiable information and uncertainty about the expert's conflict of interest and the informedness of the expert, we show that disclosure of the expert's bias is counterproductive when the magnitude of the expert's bias is not too large and the likelihood of the expert being informed is low. Moreover, the harm of disclosing the expert's conflict of interest is more significant when there is a larger uncertainty about the nature of the expert's conflict of interests.

Keywords: information transmission, conflict of interest, bias, disclosure, verifiable information, transparency

JEL codes: D72, D82.

1 Introduction

Consider a situation where a homeowner wants to install solar panels and has two options: leasing or purchasing solar panels. One option may dominate the other depending on the productivity of solar panels, shade around her house, and the terms of government subsidies. The homeowner may not have all the relevant information for decision-making and consults a salesperson. However, the salesperson may bias toward one of the options based on the structure of his commission, which is usually unknown to clients.[1] How much information will the salesperson reveal, given his bias? Will the homeowner be better off if the salesperson discloses conflicts of interest? Decision-makers have similar questions when they seek advice from financial advisors, medical professionals, or management consultants.

It is often argued that experts should disclose conflicts of interest to enhance transparency and facilitate more informed decisions by clients. Based on this principle, Physician Payments Sunshine Act (PPSA) requires physicians to disclose financial interests in manufacturers of drugs, devices, biologicals, and medical supplies; SEC issues guidance requiring that investment advisors disclose financial conflicts of interest when advising clients. In this paper, we formally study this problem in a theoretical model of information transmission, where a biased expert advises a decision-maker who is uncertain about the expert's bias. We investigate whether and when mandatory disclosure of the expert's bias can improve the decision maker's welfare.

We adopt Crawford and Sobel's (1982) setup where the state is drawn from an interval, and the expert and the decision-maker have partially aligned preferences that are represented by quadratic loss functions. However, we assume that information is verifiable, in that the expert can withhold information but cannot lie about it.

[1]For expositional convenience, we use "he" to refer to the advisor/expert and "she" to the client/decision-maker.

So, we have a "hard-information" model in the style of Milgrom (1981) and Grossman (1981). The hard-information assumption applies in expert-client relationships where lying is detectable, is a violation of the professional code of conduct, and is subject to severe penalties and punishment. There are two types of uncertainties about the characteristics of the expert–his informedness and his bias. The expert's characteristics are his private information. When the state of the world is realized, the informed expert perfectly observes it, and the uninformed expert learns nothing. Then, the expert sends a verifiable message to the decision-maker, who then takes an action to maximize her utility. The assumption that the expert has different degrees of informedness is motivated by the fact that experts differ in their expertise. For example, some salespersons are fully aware of all the relevant information for solar installation, whereas others may not know the details of government subsidies.

It is illustrative to consider the case where the expert's biases are in the opposite directions with the same magnitude. If the decision maker faces uncertainty about the expert's bias, the expert will withhold information close to the expected state and disclose information in extreme states. Specifically, the upward biased expert withholds information in an interval connected to the left of the interval in which the downward biased expert withholds information. The point where the two intervals connect is close to the expected state. If the expert is silent, the decision maker infers that there are three possible events: (i) the expert is uninformed about the state, (ii) the expert is informed and has an upward bias, and the state of the world is low, and (iii) the expert is informed and has a downward bias, and the state is high. The decision maker will take into account the probabilities of the three events and take an action at the conditional expected state, which is the point where the two intervals connect.

Now, consider mandatory disclosure of conflicts of interest. The expert's bias is known to the decision maker in this case. Suppose the expert has an upward bias and does not reveal any information. The decision maker will be more suspicious that the state is low than the case when she is unsure about the expert's bias. As a result, staying silent will lead to a lower action than the case with uncertainty on the expert's bias. The case with downward bias is symmetric. Unraveling does not arise in equilibrium because the expert may not know the state. The intervals in which the upward biased and downward biased experts withhold information are disjoint and close to the two extreme states. Hence, the expert will withhold more extreme information compared with no mandatory disclosure.

We identify a novel tradeoff that affects the decision-maker's welfare if the expert discloses conflicts of interest. On the one hand, disclosing conflicts of interest may induce the informed expert to reveal more information than without disclosure. This is because the expert cannot pool with others with different biases and hence faces more pressure of unraveling. This force increases the decision maker's payoff.

On the other hand, disclosure reduces the decision maker's expected payoff if the expert is uninformed about the state of the world. If the decision maker knew that the expert is uninformed, her optimal action is equal to the expected state of the world. Suppose fthe decision-maker does not know the nature of the expert's conflicts of interest. Then, she tends to interpret the expert's "silence" in a more neutral way and takes an action close to the expected state of the world. If the decision-maker knows that the expert has an upward bias, she is more suspicious that the state is low when the expert stays silent than if the expert's bias is unknown. As a result, when the upward-biased expert stays silent, the decision maker will take an action further away from the expected state toward the low end. Hence, under disclosure, the decision-maker will bear a larger loss from distortion in action if the expert is uninformed about the state.

We find that mandatory disclosure of conflicts of interest hurts the decision-maker if the expert's bias is small and the likelihood that he is informed about the state is low. The decision maker benefits from disclosure, otherwise. Moreover, the decision maker suffers a larger loss from disclosure of conflicts of interest if she faces a larger uncertainty about the expert's bias. We also find that the expert prefers not to disclose his bias. Hence, he will not voluntarily disclose conflicts of interest.

Our results are robust against the distribution of the expert's bias. Disclosure of conflicts of interest may hurt the decision maker even when the expert biases toward the same direction, provided that the magnitudes of the biases are small. The tradeoff holds qualitatively true given the decision maker's quadratic loss function. Under disclosure, the decision maker will take different actions when different types of experts withhold information. By contrast, she will take the same action when the expert with unknown biases withholds information. The quadratic loss function implies that the expected distortion in action is higher under disclosure than without disclosure.

Academic researchers have come to different conclusions on this subject, both theoretically (Li and Madarasz, 2008) and experimentally (Cain et al., 2005; Chung and Harbaugh, 2019; Ismayilov and Potters, 2013).[2] Dye (1985) (see also Jung and Kwon 1988) considers a model in which the expert with a known bias may not be informed. He shows that information will not unravel in equilibrium. We focus on the case in which the decision maker faces uncertainty about the expert's bias and evaluate the welfare consequence of the mandatory disclosure of conflicts of interest.

In previous work, Seidmann and Winter (1997) generalize Milgrom's (1981) "unravelling" result to the case where there is not a straightforward delineation between "good news" and "bad news."[3] In particular, they show that in the uniform-quadratic case of Crawford and Sobel (1982), full revelation is a plausible equilibrium if the expert is able to make any true statement in the form of "the state of the world is

[2]Loewenstein et al. (2014) provide a review of the relevant literature, with a special focus on behavioural models and experimental evidence.

[3]See Hagenbach et al. (2014) for a further generalization of their finding.

in set S," and only such statements.

Li and Madarasz (2008) consider information transmission in a setting where the decision maker does not know the *direction* of the bias of the expert, and show that in the "uniform-quadratic" setup if the expert can only use *cheap-talk* messages, both the expert and the decision maker are strictly better off from not having the bias of the expert disclosed, as long as the magnitude of the expert's bias is not too big. Although our results share some similarity with Li and Madarasz (2008), the driving forces for the results are very different. In our model, the expert is not always informed and mandatory disclosure policy is harmful to the client when the expert is uninformed. In fact, in our setting, mandatory disclosure of interests benefits the client if the expert is always informed. By contrast, Li and Madarasz (2008) assume that expert is always perfectly informed.

Our theoretical framework is related to that of Shavell (1989) and Bhattacharya and Mukherjee (2013), who like us consider verifiable information and "reveal or not reveal" communication strategies. However, they do not consider uncertainty about the expert's bias. Wolinsky (2003) studies a model of communication with verifiable information and some uncertainty about the expert's preferences, but the form of uncertainty is different and he does not consider welfare consequences of disclosure of conflicts of interest. In concurrent work, Mezzetti (2020) considers a hard-information communication game in which the expert is fully informed and has access to a wider set of available messages. Unlike ours, his focus is not on the welfare comparison between disclosure of conflicts of interest and nondisclosure.

2 Model

There are an expert and a decision maker, both of whom have quadratic loss preferences. The expert's (sender) payoff is denoted by

$$u(\theta, y, \beta) = -\left(y - (\theta + \beta)\right)^2,$$

and the decision maker's (receiver) payoff is denoted by

$$v(\theta, y) = -(\theta - y)^2,$$

where $\theta \in \Theta$ is the state of the world with distribution function F, y is the decision maker's action, and β is the random bias of the expert. We assume that $\Theta = [0, 1]$ and the density function f is unimodal and symmetric around the mean $\mu \equiv E(\theta)$. Many distributions satisfy these properties, including uniform distribution, triangle distribution, (truncated) normal distribution, and (truncated) Cauchy distribution. Given our assumption about the symmetry of the distribution, $\mu = 1/2$. Furthermore, we assume that $\beta \in \{-b, b\}$, with $0 < b \leq \mu/2$, and $\beta = b$ with probability $p \in (0, 1)$. So the size of the expert's bias is commonly known, but its direction is the expert's private information. This assumption allows a clean analysis that demonstrates the basic intuition. In Section 5, we consider general distributions of β.

Regardless of the expert's bias β, the expert is perfectly informed about θ with probability $q \in (0, 1)$ and is uninformed with the complementary probability. The probabilities q and p are common knowledge. Whether or not the expert is informed about the state of the world and the direction of his bias is the expert's private information.

The expert sends the decision maker a verifiable message. If the expert is informed, he can either report the state of the world or report nothing. Hence, the message space is $M = \Theta \cup \{\emptyset\}$, where $\emptyset$ means "report nothing." The uninformed type can only report message $\emptyset$. The informed type learns the state of the world, and his strategy is $m_I^{\beta}(\theta) : \{-b, b\} \times \Theta \to M$. The decision maker's strategy is a mapping from the message space to the action space $Y : M \to [0, 1]$.

The game unfolds as follows:

- Stage 1. Nature draws the state of the world and the expert's type, which indicates whether he is informed and the direction of his bias. The expert privately learns his type. The informed expert also privately learns the state of the world.

- Stage 2. The expert sends a verifiable message to the decision maker.

- Stage 3. The decision maker receives the message and then takes an action.

In what follows, we characterize the perfect Bayesian equilibria when the expert does not disclose his bias and when he is required to do so. Then, we analyze the welfare implications of the disclosure policy.

3 No Disclosure

We first analyze the equilibrium when the expert does not disclose the direction of his bias.

Because the expert cannot lie about the state of the world, in any perfect Bayesian equilibrium, the decision maker either receives a message perfectly revealing the state of the world or the message $\emptyset$ revealing no information about the state. Given the decision maker's preferences, upon receiving the perfectly revealing message, the decision maker takes an action which matches the message and obtains the maximum payoff 0. Given the decision maker's action, the expert's payoff is

$$u(\theta, y, \beta) = -(\beta)^2 .$$

If the decision maker receives the message $\emptyset$, she forms an expectation about the state of the world based on the prior about the expert's type and his equilibrium strategy. The decision maker then takes the action denoted by $y_\emptyset \in [0,1]$ to maximize her expected payoff.

Now, consider the informed expert's decision about which message to send. As shown above, the expert obtains $u(\theta,\theta,\beta) = -(\beta)^2$ by sending the truth revealing message and $u(\theta, y_\emptyset, \beta) = -(y_\emptyset - (\theta+\beta))^2$ by sending the no information message $\emptyset$. Because $u(\theta,\theta,\beta)$ is constant in θ and $u(\theta, y_\emptyset, \beta)$ is concave in θ, the expert prefers to withhold information if and only if $\theta \in [\max\{y_\emptyset - \beta - |\beta|, 0\}, \min\{y_\emptyset - \beta + |\beta|, 1\}]$.[4] Hence, the expert will send the message $\emptyset$ for $\theta \in [\underline{y}(y_\emptyset), y_\emptyset]$ if he has a positive bias and for $\theta \in [y_\emptyset, \overline{y}(y_\emptyset)]$ if he has a negative bias, where

$$\begin{aligned} \underline{y}(y_\emptyset) &\equiv \max\{y_\emptyset - 2b, 0\}; \\ \overline{y}(y_\emptyset) &\equiv \min\{y_\emptyset + 2b, 1\}. \end{aligned}$$

To see the intuition, recall that the positive biased expert wants to induce an action higher than the state of the world. If $\theta > y_\emptyset$, withholding information will result in an action lower than the state of the world. As a result, it is optimal for the the expert to truthfully report the state. The expert is indifferent between whether or not to withhold information at $\theta = y_\emptyset$ because the decision maker will take the same action. For $\theta < y_\emptyset$, the expert has incentives to withhold information. Because the expert has a moderate bias, i.e. $b \leq \mu/2$, he is better off inducing the action $y_\emptyset$ when the state is not too far away from it. When state is significantly lower than $y_\emptyset$, the expert prefers to truthfully report it because withholding information will lead to a larger loss due to the big distortion in the decision maker's action. If the the expert has an extreme bias, he will wish to induce $y_\emptyset$ for all states lower than it. We

[4]The expert is indifferent between whether or not to report the state of the world when $\theta \in \{y_\emptyset - \beta - |\beta|, y_\emptyset - \beta + |\beta|\}$, and we assume that the expert withholds the information when he is indifferent.

ignore the case with extreme bias because it does not add any new insights and is less interesting. The intuition for the negative biased expert is analogous.

Upon receiving the message $\emptyset$, the decision maker updates her expected state of the world to

$$E(\theta|\emptyset) = Pr(U|\emptyset)E(\theta|U,\emptyset) + Pr(I|\emptyset)E(\theta|I,\emptyset), \tag{1}$$

where U (I) means that the expert is uninformed (informed) and

$$Pr(U|\emptyset) \equiv \frac{1-q}{1-q+qPr(\emptyset|I)}$$

is the probability that the expert is uninformed conditional on message $\emptyset$, $E(\theta|U,\emptyset)$ is the conditional expectation of θ when the expert is uninformed and sends $\emptyset$, and $E(\theta|I,\emptyset)$ is the conditional expectation of θ when the expert is informed and sends message $\emptyset$.

Given that the decision maker has the quadratic loss function, her optimal action upon receiving message $\emptyset$ equals her expected state of the world. Hence, the decision maker's equilibrium action $y^*_\emptyset$ solves

$$y_\emptyset = E(\theta|\emptyset). \tag{2}$$

Substitute $E(\theta|\emptyset)$ and rearrange terms, (2) can be rewritten as

$$(1-q)(\mu - y_\emptyset) = qPr(\emptyset|I)\left(y_\emptyset - E(\theta|I,\emptyset)\right). \tag{3}$$

Upon receiving message $\emptyset$, the decision maker infers that the expected state is μ if the expert is uninformed and is $E(\theta|I,\emptyset)$ if the expert is informed and choose to withhold information. Equation (3) states that the decision maker's optimal action should equate the expected loss when the expert is uninformed and when he is informed.

Substituting

$$Pr(\emptyset|I) = p\int_{\underline{y}(y_\emptyset)}^{y_\emptyset} dF(\theta) + (1-p)\int_{y_\emptyset}^{\overline{y}(y_\emptyset)} dF(\theta)$$

and

$$E(\theta|I,\emptyset) = \frac{p\int_{\underline{y}(y_\emptyset)}^{y_\emptyset}\theta dF(\theta) + (1-p)\int_{y_\emptyset}^{\overline{y}(y_\emptyset)}\theta dF(\theta)}{p\int_{\underline{y}(y_\emptyset)}^{y_\emptyset}dF(\theta) + (1-p)\int_{y_\emptyset}^{\overline{y}(y_\emptyset)}dF(\theta)}$$

into the right hand side of (3), we have

$$(1-q)(\mu - y_\emptyset) = q\left(p\int_{\underline{y}(y_\emptyset)}^{y_\emptyset}(y_\emptyset - \theta)dF(\theta) - (1-p)\int_{y_\emptyset}^{\overline{y}(y_\emptyset)}(\theta - y_\emptyset)dF(\theta)\right). \quad (4)$$

Note that left-hand side of (4) decreases in $y_\emptyset$ whereas the right-hand side of it increases in $y_\emptyset$. Moreover, the left-hand side is greater than the right-hand side at $y_\emptyset = 0$ and is smaller than the right-hand side at $y_\emptyset = 1$. Hence, there exists an action $y_\emptyset$ that satisfies (4). The assumption that $f(\theta)$ is symmetric and unimodal ensures the uniqueness of the equilibrium action. In the following two propositions, we characterize the decision maker's equilibrium action for a given prior probability of the direction of the expert's bias.

Proposition 1. *Suppose $p = 1/2$. In the unique equilibrium, when receiving the null message, the decision maker takes action $y_\emptyset^* = \mu$. The informed expert withholds information for $\theta \in [\mu - 2b, \mu]$ if he has a positive bias and for $\theta \in [\mu, \mu + 2b]$ if he has a negative bias.*

Proof of Proposition 1. We have characterized the expert's strategy given the decision maker's action $y_\emptyset$. It suffices to show that $y_\emptyset^* = \mu$ is the unique solution for equation (4).

First, we show that $y_\emptyset = \mu$ is a solution for (4). The left-hand side of (4) is zero at $y_\emptyset = \mu$. Changing variables, the right-hand side of (4) is simplified to

$$\begin{aligned} & q\left(-p\int_{y_\emptyset - \underline{y}(y_\emptyset)}^{0} tf(y_\emptyset - t)dt - (1-p)\int_0^{\overline{y}(y_\emptyset) - y_\emptyset} tf(y_\emptyset + t)dt\right) \\ = \; & q\left(p\int_0^{y_\emptyset - \underline{y}(y_\emptyset)} tf(y_\emptyset - t)dt - (1-p)\int_0^{\overline{y}(y_\emptyset) - y_\emptyset} tf(y_\emptyset + t)dt\right). \quad (5) \end{aligned}$$

Given the assumption $b < \mu/2$, $y_\emptyset = \mu$ implies that $\underline{y}(y_\emptyset) \equiv max\{\mu - 2b, 0\} = \mu - 2b$ and $\bar{y}(y_\emptyset) \equiv min\{\mu + 2b, 1\} = \mu + 2b$. Substituting $y_\emptyset = \mu$, $\underline{y}(y_\emptyset)$, $\bar{y}(y_\emptyset)$, and $p = \frac{1}{2}$, (5) becomes

$$\begin{aligned} & \frac{q}{2}\left(\int_0^{2b} tf(\mu - t)dt - \int_0^{2b} tf(\mu + t)dt\right) \\ = & \frac{q}{2}\left(\int_0^{2b} tf(\mu + t)dt - \int_0^{2b} tf(\mu + t)dt\right) \\ = & 0, \end{aligned}$$

where the second equality follows from the symmetry of $f(\theta)$.

Next, we show $y_\emptyset = \mu$ is the unique solution. Suppose $y_\emptyset < \mu$. Then, the left-hand side of (4) is positive. The right-hand side of (4) is

$$\begin{aligned} & \frac{q}{2}\left(\int_0^{y_\emptyset - \underline{y}(y_\emptyset)} tf(y_\emptyset - t)dt - \int_0^{\bar{y}(y_\emptyset) - y_\emptyset} tf(y_\emptyset + t)dt\right) \\ \leq & \frac{q}{2}\left(\int_0^{2b} tf(y_\emptyset - t)dt - \int_0^{2b} tf(y_\emptyset + t)dt\right) \\ < & \frac{q}{2}\left(\int_0^{2b} tf(y_\emptyset + t)dt - \int_0^{2b} tf(y_\emptyset + t)dt\right) \\ = & 0. \end{aligned}$$

The first inequality follows because $\underline{y}(y_\emptyset) \geq y_\emptyset - 2b$ and $\bar{y}(y_\emptyset) = y_\emptyset + 2b$ given $y_\emptyset < \mu$. The second inequality holds because $f(y_\emptyset - t) < f(y_\emptyset + t)$. To see this, recall $f(\theta)$ is symmetric and unimodal. It follows that $f(y_\emptyset - t) < f(y_\emptyset + t)$ if $y_\emptyset + t \leq \mu$, and $f(y_\emptyset - t) < f(\mu - t) = f(\mu + t) < f(y_\emptyset + t)$ if $y_\emptyset + t > \mu$.

The above argument shows that the right-hand side of (4) is strictly less than the left-hand side of (4) at $y_\emptyset < \mu$. This contradicts the claim that $y_\emptyset$ is a solution for (4).

Now, suppose $y_\emptyset > \mu$. The left-hand side of (4) is negative . The right-hand side of (4) is

$$
\begin{aligned}
& \frac{q}{2}\left(\int_0^{y_\emptyset-\underline{y}(y_\emptyset)} tf(y_\emptyset-t)dt-\int_0^{\bar{y}(y_\emptyset)-y_\emptyset} tf(y_\emptyset+t)dt\right) \\
\geq\; & \frac{q}{2}\left(\int_0^{2b} tf(y_\emptyset-t)dt-\int_0^{2b} tf(y_\emptyset+t)dt\right) \\
>\; & \frac{q}{2}\left(\int_0^{2b} tf(y_\emptyset-t)dt-\int_0^{2b} tf(y_\emptyset-t)dt\right) \\
=\; & 0.
\end{aligned}
$$

The first inequality follows because $\underline{y}(y_\emptyset) = y_\emptyset - 2b$ given $y_\emptyset > \mu$ and $\bar{y}(y_\emptyset) \leq y_\emptyset + 2b$. The second inequality holds because $f(y_\emptyset+t) < f(y_\emptyset-t)$ if $y_\emptyset - t \geq \mu$ and $f(y_\emptyset+t) < f(\mu+t) = f(\mu-t) < f(y_\emptyset-t)$ if $y_\emptyset - t < \mu$. It follows that the right-hand side of (4) is strictly greater the left-hand side of (4) at $y_\emptyset > \mu$. This contradicts the claim that $y_\emptyset$ is a solution for (4). $\square$

First, consider the expert's strategy. The uninformed expert has to send the message $\emptyset$, irrespective of the direction of his bias. The informed expert chooses between withholding information and truthfully reporting it. Recall that for a given action $y_\emptyset$, the positive biased expert withholds information for state in $[\underline{y}(y_\emptyset), y_\emptyset]$ whereas the negative biased expert withholds information for state in $[y_\emptyset, \bar{y}(y_\emptyset)]$. Given the decision maker's action $y_\emptyset = \mu$ and under the assumption $b \leq \mu/2$, $\underline{y}(\mu) = \mu - 2b$ and $\bar{y}(\mu) = \mu + 2b$. Hence, the range in which the expert with different biases withholds information is symmetric around μ.

Next, consider the decision maker's strategy upon receiving the message $\emptyset$. Based on the expert's strategy, the decision maker infers that the expert sends the message $\emptyset$ in the following three events: (i) the expert is uninformed, (ii) the expert is informed and has the positive bias, and the state is in $[\mu-2b, \mu]$, and (iii) the expert is informed and has the negative bias, and the state is in $[\mu, \mu+2b]$. The decision maker will take an action equals to the expected state taking into account the probabilities of the three events. If $p = 1/2$, the expert has the positive and the negative biases with equal probabilities. Moreover, the range for withholding information is symmetric around μ when the expert has different biases. As a result, the expected state conditional on the expert being informed is μ and coincides with the expected state conditional on the expert being uninformed. It follows that $y_\emptyset = \mu$ is the decision maker's optimal action.

Note that when $p = \frac{1}{2}$, the decision maker's optimal action upon message $\emptyset$ is independent of the size of the expert's bias b and the probability that the expert is informed q. As b increases, the expert becomes more biased, and hence the range in which he withholds information expands.

The next proposition summarizes the equilibrium if the expert is more likely to have a positive (negative) bias.

Proposition 2. *In any equilibrium, the decision maker's action* $0 < y^*_\emptyset < \mu$ *if* $1/2 < p < 1$ *and* $\mu < y^*_\emptyset < 1$ *if* $0 < p < 1/2$. *The informed expert withholds information for* $\theta \in [\underline{y}(y^*_\emptyset), y^*_\emptyset]$ *if he has a positive bias and for* $\theta \in [y^*_\emptyset, \bar{y}(y^*_\emptyset)]$ *if he has a negative bias.*

Proof of Proposition 2. Consider $p > 1/2$. Suppose $y_\emptyset \geq \mu$. We first show that $f(y_\emptyset - t) > f(y_\emptyset + t)$ for all $t > 0$. If $y_\emptyset - t \geq \mu$, then $f(y_\emptyset - t) > f(y_\emptyset + t)$ because symmetry of $f(\theta)$ implies that $f(\theta)$ is monotonically decreasing for $\theta \geq \mu$. If $y_\emptyset - t < \mu$, then

$$f(y_\emptyset - t) = f(2\mu + t - y_\emptyset) > f(y_\emptyset + t).$$

The equality follows from the symmetry of $f(\theta)$ and the inequality follows from $\mu < 2\mu + t - y_\emptyset < y_\emptyset + t$.

If $y_\emptyset \geq \mu$, the left-hand side of (4) is nonpositive. Following (5) in the proof for

Proposition 1, the right-hand side of (4) can be written as

$$\begin{aligned}
& q\left(p\int_0^{y_\emptyset-\underline{y}(y_\emptyset)} tf(y_\emptyset-t)dt-(1-p)\int_0^{\bar{y}(y_\emptyset)-y_\emptyset} tf(y_\emptyset+t)dt\right)\\
\geq\ & q\left(p\int_0^{2b} tf(y_\emptyset-t)dt-(1-p)\int_0^{2b} tf(y_\emptyset+t)dt\right)\\
>\ & q\left(p\int_0^{2b} tf(y_\emptyset-t)dt-(1-p)\int_0^{2b} tf(y_\emptyset-t)dt\right)\\
>\ & q(1-p)\left(\int_0^{2b} tf(y_\emptyset-t)dt-\int_0^{2b} tf(y_\emptyset-t)dt\right)\\
=\ & 0.
\end{aligned}$$

The first inequality holds because $\underline{y}(y_\emptyset)=y_\emptyset-2b$ given $y_\emptyset\geq\mu$ and $\bar{y}(y_\emptyset)\leq y_\emptyset+2b$. The second inequality follows from $f(y_\emptyset+t)<f(y_\emptyset-t)$ and the third inequality follows from $p>\frac{1}{2}$. Hence, if $y_\emptyset\geq\mu$, the left-hand side of (4) is strictly less than the right-hand side of (4). This contradicts the claim that $y_\emptyset$ is the solution for (4).

Next, we show that there exists a $0<y_\emptyset<\mu$ which solves (4). At $y_\emptyset=0$, the left-hand side of (4) is positive and the right-hand side of (4) is negative because $y_\emptyset=\underline{y}(y_\emptyset)=0$. At $y_\emptyset=\mu$, the left-hand side of (4) is zero. The right-hand side of (4) is positive because

$$\begin{aligned}
& q\left(p\int_0^{\mu-\underline{y}(y_\emptyset)} tf(\mu-t)dt-(1-p)\int_0^{\bar{y}(y_\emptyset)-\mu} tf(\mu+t)dt\right)\\
=\ & q\left(p\int_0^{2b} tf(\mu-t)dt-(1-p)\int_0^{2b} tf(\mu+t)dt\right)\\
>\ & q(1-p)\left(\int_0^{2b} tf(\mu-t)dt-\int_0^{2b} tf(\mu+t)dt\right)\\
=\ & 0.
\end{aligned}$$

The first equality holds because $\underline{y}(\mu)=\mu-2b$ and $\bar{y}(\mu)=\mu+2b$ and the last equality follows from the symmetry of $f(\theta)$. Because both the left-hand side and the right-hand side of (4) are continuous functions of $y_\emptyset$, there exists a $y_\emptyset\in(0,\mu)$ that solves (4).

The case of $p < \frac{1}{2}$ is symmetric to the case of $p > \frac{1}{2}$. The analysis is analogous and is skipped to avoid repetition. □

Compared with the case with symmetric probability of bias, the decision maker's action upon message $\emptyset$ will be away from μ in direction opposite to the direction toward which the expert is more likely to be bias. Intuitively, if the expert is more likely to have a positive bias, whenever he withholds information, the decision maker will infer that the state of the world is more likely to be lower than the average state.

4 Disclosure

In this subsection, we analyze the equilibrium when disclosure of bias is mandatory. The decision maker knows the direction of the expert's bias but does not know whether or not the expert is informed. Similar to the case of no disclosure, in any perfect Bayesian equilibrium, the decision maker either receives a perfectly revealing message or the message $\emptyset$. Let $y_\emptyset^b$ denote the decision maker's action when the expert with the positive bias sends the message $\emptyset$ and $y_\emptyset^{-b}$ denote the action when the expert with the negative bias sends $\emptyset$.

Suppose that the expert has the positive bias. Following the same argument in the case of no disclosure, the decision maker's action $y_\emptyset^b$ satisfies

$$(1-q)(\mu - y_\emptyset) = q\left(\int_{\underline{y}(y_\emptyset)}^{y_\emptyset} (y_\emptyset - \theta)dF(\theta)\right). \tag{6}$$

Condition (6) is obtained by substituting $p = 1$ into (4). Similarly, if the expert has the downward bias, the decision maker's action $y_\emptyset^{-b}$ satisfies

$$(1-q)(\mu - y_\emptyset) = -q\left(\int_{y_\emptyset}^{\overline{y}(y_\emptyset)} (\theta - y_\emptyset)dF(\theta)\right), \tag{7}$$

which is obtained by substituting $p = 0$ into (4).

Proposition 3. *If the expert has the positive bias, $y^b_\emptyset$ is unique and $0 < y^b_\emptyset < \mu$; If the expert has the negative bias, $y^{-b}_\emptyset$ is unique and $\mu < y^{-b}_\emptyset < 1$. The informed expert withholds information for $\theta \in [\underline{y}(y^b_\emptyset), y^b_\emptyset]$ if he has the positive bias and for $\theta \in [y^{-b}_\emptyset, \overline{y}(y^{-b}_\emptyset)]$ if he has the negative bias. The action $y^b_\emptyset$ decreases in b and q whereas $y^{-b}_\emptyset$ increases in b and q.*

Proof of Proposition 3. Consider the expert with the positive bias. We first show that $0 < y^b_\emptyset < \mu$. Suppose $y^b_\emptyset = 0$. Then, the left-hand-side of (6) is positive but the right-hand-side of (6) is zero because $y^b_\emptyset = \underline{y}(y^b_\emptyset) = 0$. This contradicts the claim that $y^b_\emptyset$ is the solution for (6). Next, suppose $y^b_\emptyset \geq \mu$. Then, the left-hand-side is nonpositive. The right-hand-side of (6) is positive. Again, we have a contradiction.

Next, we show that $y^b_\emptyset$ is unique. The Left-hand-side of (6) strictly decreases in $y_\emptyset$. Let $R(y_\emptyset)$ denote the right-hand-side of (6). We show that $R(y_\emptyset)$ is increasing in $y_\emptyset$ for $y_\emptyset \leq \mu$. Suppose $y^b_\emptyset \leq 2b$, so $\underline{y}(y^b_\emptyset) = 0$. Then,

$$R'(y^b_\emptyset) = qF(y^b_\emptyset) > 0.$$

Suppose $y^b_\emptyset > 2b$. Then, $\underline{y}(y^b_\emptyset) = y^b_\emptyset - 2b$, and

$$\begin{aligned} R'(y_\emptyset) &= 2bq\left(\frac{F(y^b_\emptyset t) - F(y^b_\emptyset - 2b)}{2b} - f(y^b_\emptyset - 2b)\right) \\ &= 2bq\left(f(\hat{\theta}) - f(y^b_\emptyset - 2b)\right) \\ &> 0, \end{aligned} \tag{8}$$

where $\hat{\theta} \in (y^b_\emptyset - 2b, y^b_\emptyset)$. The second equality follows from the mean value theorem, and the last inequality follows from $y^b_\emptyset - 2b \leq \hat{\theta} < \mu$ and the assumption that $f(\theta)$ is symmetric and unimodal. Given that the left-hand-side of (6) is decreasing in $y_\emptyset$ whereas the right-hand-side of it is increasing in (6), $y^b_\emptyset$ is unique.

Now, we show that $y_\emptyset$ decreases in q and b. Applying the implicit function theorem to (6) with respect to q, we have

$$
\frac{\partial y^b_\emptyset}{\partial q} = \begin{cases} -\frac{\int_0^{y^b_\emptyset}(y^b_\emptyset-\theta)dF(\theta)+\mu-y^b_\emptyset}{qF(y^b_\emptyset)+1-q} < 0 & \text{if } y^b_\emptyset \leq 2b; \\ \\ -\frac{\int_{y^b_\emptyset-2b}^{y^b_\emptyset}(y^b_\emptyset-\theta)dF(\theta)+\mu-y^b_\emptyset}{2bq\left(\frac{F(y^b_\emptyset)-F(y^b_\emptyset-2b)}{2b}-f(y^b_\emptyset-2b)\right)+1-q} < 0 & \text{if } y^b_\emptyset > 2b. \end{cases}
$$

If $y^b_\emptyset \leq 2b$, $\underline{y}(y^b_\emptyset) = 0$ and $\frac{\partial y^b_\emptyset}{\partial q} < 0$ given $y^b_\emptyset < \mu$. If $y^b_\emptyset > 2b$, $\underline{y}(y^b_\emptyset) = y^b_\emptyset - 2b$. In this case, $\frac{\partial y^b_\emptyset}{\partial q} < 0$ because $2bq\left(\frac{F(y_\emptyset)-F(y_\emptyset-2b)}{2b} - f(y_\emptyset - 2b)\right) > 0$ by (8).

Similarly, applying the implicit function theorem to (6) with respect to b, we have

$$
\frac{\partial y^b_\emptyset}{\partial q} = \begin{cases} 0 & \text{if } y^b_\emptyset \leq 2b; \\ \\ -\frac{4qbf(y^b_\emptyset-2b)}{2bq\left(\frac{F(y^b_\emptyset)-F(y^b_\emptyset-2b)}{2b}-f(y^b_\emptyset-2b)\right)+1-q} < 0 & \text{if } y^b_\emptyset > 2b. \end{cases}
$$

The analysis of downward biased expert is analogous and is skipped to avoid repetition. □

When the expert's conflict of interest is known, information about the state does not unravel because the expert may not be informed, as shown by Dye (1985). A notable difference between our setup and Dye's is that we consider the case in which the expert's interest is partially aligned with the decision maker whereas Dye assumes that the expert has an extreme bias. As a result, the expert in our model may withhold information in an interior range of states whereas in Dye, the expert withholds information when the state is below (above) a cutoff.

If the decision maker knows that the expert has a positive bias, she will be more suspicious that the state is very low when the expert withholds information than the case when she is unsure about the nature of the expert's bias. As a result, the decision maker will take an action to the left of the average state μ when the expert withholds information. Moreover, the action $y^b_\emptyset$ will be lower if the magnitude of the expert's bias is larger or the likelihood of the expert being informed is higher because the expert has stronger incentives to hide the low states. The case when the expert has a negative bias is symmetric.

It is useful to categorize the cases under disclosure. Note that given the symmetry assumption, the disclosure equilibria in the cases of the positive and negative biased experts are mirror images of each other. Thus, it is necessary only to consider the case with a positive bias. We call the type of equilibrium when $y_\emptyset^b > 2b$ an interior-cutoff equilibrium, which corresponds to the expert hiding information in an interval that is interior to $[0, 1]$ and the other type the Type B equilibrium, which corresponds to the left cutoff point being 0.

5 Welfare

In this section, we evaluate the welfare implications of the policy that requires the expert to disclose his conflicts of interest. To begin, we compare the decision maker's expected payoff when the expert does not disclose the nature of his bias and when he discloses it. We then compare the expert's expected payoff in these two regimes to investigate whether the expert will voluntarily disclose conflicts of interest. Because there is no closed form solution for the consumer's action $y_\emptyset^b$ and $y_\emptyset^{-b}$ under Disclosure, it is difficult to evaluate the welfare for a general distribution of the state of the world, F. Thus, in our welfare analysis, we focus on the case when the state of the world, θ, follows the uniform distribution.

Define

$$\hat{b}(q) \equiv \left(\frac{\sqrt{1-q}}{1+\sqrt{1-q}}\right)^{\frac{2}{3}} \left(\frac{3+\sqrt{1-q}}{32(1+\sqrt{1-q})}\right)^{\frac{1}{3}}.$$

Proposition 4. *Suppose that θ is uniformly distributed on $[0, 1]$. If $p = \frac{1}{2}$, disclosing the expert's bias reduces the decision maker's welfare if $b \leq \hat{b}(q)$ and increases the decision maker's welfare if $b > \hat{b}(q)$.*

Proof of Proposition 4. The proof has three steps.

Step 1 calculates the decision maker's expected payoff under No Disclosure. Proposition 1, shows that $y_{\emptyset}^{*} = \frac{1}{2}$ and the positive biased expert withholds information for $\theta \in [\frac{1}{2} - 2b, \frac{1}{2}]$ and the negative biased expert withholds information for $\theta \in [\frac{1}{2}, \frac{1}{2} + 2b]$. The decision maker's expected payoff is therefore

$$\begin{aligned} E(v) &= -(1-q)\int_0^1 (\frac{1}{2}-\theta)^2 d\theta - \frac{q}{2}\left(\int_{\frac{1}{2}-2b}^{\frac{1}{2}+2b} (\frac{1}{2}-\theta)^2 d\theta\right) \\ &= -\frac{(1-q)}{12} - \frac{8qb^3}{3}. \end{aligned} \tag{9}$$

Step 2 solves the decision maker's expected payoff under Disclosure. The decision maker's action $y_{\emptyset}^{b}$ following message $\emptyset$ sent by the upward biased expert is determined by equations (6), and her action $y_{\emptyset}^{-b}$ following message $\emptyset$ sent by the downward biased expert is determined by (7).

Suppose b is small so that $\underline{y}(y_{\emptyset}) \equiv \max\{y_{\emptyset} - 2b, 0\} = y_{\emptyset} - 2b$ and $\overline{y}(y_{\emptyset}) \equiv \min\{y_{\emptyset} + 2b, 1\} = y_{\emptyset} + 2b$. The unique solutions for (6) and (7) are $y_{\emptyset}^{b} = \frac{1}{2} - \frac{2qb^2}{1-q}$ and $y_{\emptyset}^{-b} = \frac{1}{2} + \frac{2qb^2}{1-q}$, respectively. The conditions $\underline{y}(y_{\emptyset}) = y_{\emptyset} - 2b$ and $\overline{y}(y_{\emptyset}) = y_{\emptyset} + 2b$ hold if and only if $b \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$. Hence, if $b \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$, the decision maker's expected payoff is

$$\begin{aligned} E^d(v) &= \frac{1}{2}\left(-(1-q)\int_0^1 (y_{\emptyset}^{b}-\theta)^2 d\theta - q\int_{y_{\emptyset}^{b}-2b}^{y_{\emptyset}^{b}} (y_{\emptyset}^{b}-\theta)^2 d\theta\right) + \\ &\quad + \frac{1}{2}\left(-(1-q)\int_0^1 (y_{\emptyset}^{-b}-\theta)^2 d\theta - q\int_{y_{\emptyset}^{-b}}^{y_{\emptyset}^{-b}+2b} (y_{\emptyset}^{-b}-\theta)^2 d\theta\right) \\ &= -(1-q)\int_0^1 (y_{\emptyset}^{b}-\theta)^2 d\theta - q\int_0^{2b} t^2 dt \\ &= -(1-q)\left(\int_0^1 (y_{\emptyset}^{b}-\frac{1}{2})^2 d\theta + \int_0^1 (\frac{1}{2}-\theta)^2 d\theta\right) - \frac{8qb^3}{3} \\ &= -\frac{4q^2b^4}{1-q} - \frac{1-q}{12} - \frac{8qb^3}{3}, \end{aligned} \tag{10}$$

where the second equality follows from $y_\emptyset^{-b} = 1 - y_\emptyset^b$ and $t \equiv y_\emptyset^b - \theta$.

Suppose $\underline{y}(y_\emptyset) = 0$ and $\overline{y}(y_\emptyset) = 1$. The unique solutions for (6) and (7) are $y_\emptyset^b = \frac{\sqrt{1-q}}{1+\sqrt{1-q}}$ and $y_\emptyset^{-b} = \frac{1}{1+\sqrt{1-q}}$, respectively. The conditions $\underline{y}(y_\emptyset) = 0$ and $\overline{y}(y_\emptyset) = 1$ are satisfied if and only if $b \geq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$. The decision maker's expected payoff is

$$\begin{aligned}
E^d(v) &= \frac{1}{2}\left(-(1-q)\int_0^1 (y_\emptyset^b - \theta)^2 d\theta - q\int_0^{y_\emptyset^b} (y_\emptyset^b - \theta)^2 d\theta\right) + \\
&\quad +\frac{1}{2}\left(-(1-q)\int_0^1 (y_\emptyset^{-b} - \theta)^2 d\theta - q\int_{y_\emptyset^{-b}}^1 (y_\emptyset^{-b} - \theta)^2 d\theta\right) \\
&= -(1-q)\int_0^1 (y_\emptyset^b - \theta)^2 d\theta - \frac{q(y_\emptyset^b)^3}{3} \\
&= -(1-q)\int_0^1 (y_\emptyset^b - \frac{1}{2})^2 d\theta - \frac{(1-q)}{12} - \frac{q(y_\emptyset^b)^3}{3}
\end{aligned}$$

Step 3 compares the decision maker's payoff under No Disclosure and under Disclosure.

If $b \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$, the difference between the decision maker's utility without and with mandatory disclosure is

$$E(v) - E^d(v) = \frac{4q^2b^4}{1-q} > 0. \tag{11}$$

If $b > \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$, the decision maker's payoff difference is

$$\begin{aligned}
E(v) - E^d(v) &= \left[-\frac{8qb^3}{3} - \frac{1-q}{12}\right] - \left[-(1-q)\int_0^1 (y_\emptyset^b - \frac{1}{2})^2 d\theta - \frac{(1-q)}{12} - \frac{q(y_\emptyset^b)^3}{3}\right] \\
&= \frac{8q}{3}\left[\frac{(y_\emptyset^b)^3}{8} + \frac{3(1-q)(y_\emptyset^b - \frac{1}{2})^2}{8q} - b^3\right].
\end{aligned} \tag{12}$$

. Consequently, $E(v) > E^d(v)$ if and only if

$$\begin{aligned}
0 &< \frac{(y_\emptyset^b)^3}{8} + \frac{3(1-q)(y_\emptyset^b - \frac{1}{2})^2}{8q} - b^3 \Rightarrow \\
b^3 &< \frac{(1-q)(\sqrt{1-q}+3)}{32(1+\sqrt{1-q})^3} \Rightarrow \\
b &< \hat{b}(q) \equiv \left(\frac{\sqrt{1-q}}{1+\sqrt{1-q}}\right)^{\frac{2}{3}}\left(\frac{3+\sqrt{1-q}}{32(1+\sqrt{1-q})}\right)^{\frac{1}{3}}.
\end{aligned} \tag{13}$$

.

Lastly, we show $\frac{\sqrt{1-q}}{2(1+\sqrt{1-q})} < \hat{b}(q) < \frac{1}{4}$. To show $\hat{b}(q) < \frac{1}{4}$, let $a \equiv 1 + \sqrt{1-q}$. Take the derivative $\frac{d\hat{b}(q)}{dq}$

$$\frac{d\hat{b}(q)}{dq} = \frac{d\hat{b}}{da}\frac{da}{dq} = -\frac{(\frac{a-1}{a})^{2/3}(\frac{2+a}{32})^{1/3}}{2(a-1)} < 0,$$

where the inequality follows because $a > 1$. Since $\hat{b}(0) = \frac{1}{4}$, $\hat{b}(q) < \frac{1}{4}, \forall q \in (0, 1)$.

Next, we shows $\hat{b}(q) > \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$. Take the derivative of $E(v)$ for $b < \frac{1}{4}$,

$$\frac{\partial E(v)}{\partial b} = -8qb^2 < 0.$$

So, $E(v)$ is decreasing in b for $b < \frac{1}{4}$. If $b \geq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$, $E^d(v)$ is constant in b. We have shown that $E(v) > E^d(v)$ at $b = \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$ and $E(v) = E^d(v)$ at $b = \hat{b}(q)$. Because $E(v)$ is monotonically decreasing in b, $\frac{\sqrt{1-q}}{2(1+\sqrt{1-q})} < \hat{b}(q)$. □

Figure 1 illustrates Proposition 4. For a fixed q, disclosing the expert's bias reduces the decision maker's utility if and only if the size of the expert's bias is lower than a cutoff. Two opposing forces affects the decision maker's payoff when the expert discloses the nature of his bias. On the one hand, the expert might reveal more information than without disclosure. This is because the expert cannot pool with others with the opposite bias and hence faces more pressure of unraveling. This force increases the decision maker's payoff. On the other hand, disclosure reduces the decision maker's expected payoff if the expert is uninformed. Recall that under no disclosure, the decision maker takes action $y^*_\emptyset = \mu$ if the expert stays silent. When the expert is uninformed about the state of the world, $y^*_\emptyset = \mu$ maximizes the decision maker's expected payoff. In contrast, under disclosure, when the expert stays silent, the decision maker will take an action lower than μ if the expert has a positive bias and higher than μ if he has a negative bias. The distortion in action from μ hurts the decision maker when the expert is uninformed.

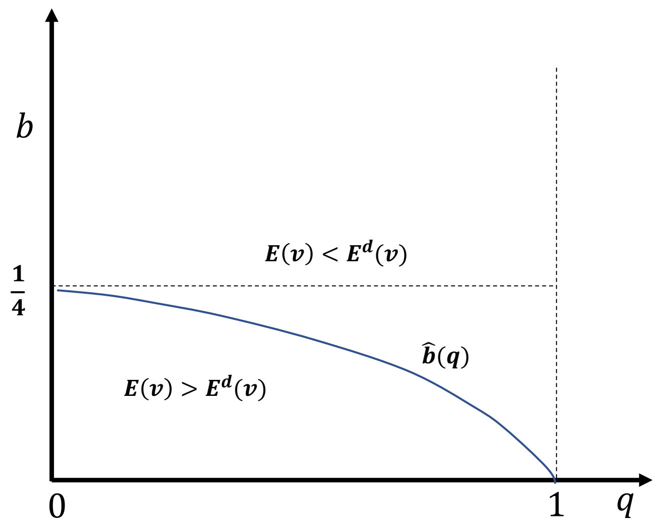

Figure 1: Comparison between disclosure and no disclosure for the case of $p = 1/2$ and uniform distribution of θ. The function $\hat{b}$ gives the points where the decision maker is indifferent between disclosure and nondisclosure.

The intuition is best illustrated in Figures 2 and 3. These two figures demonstrate the expert and the decision maker's strategies under no disclosure regime and disclosure regime for small and large bias, respectively. In each regime, the diagram on the top refers to expert with the positive bias, and the diagram at the bottom refers to expert with the negative bias.

Figure 2 shows that for small biases, the expert withholds the same amount of information under both regimes. The informed expert withholds information in an interval with length $2b$ in both regimes. Conditional on the expert being informed, the decision maker's expected utility only depends on the length of the interval in which the expert withholds information under the uniform distribution. Hence, disclosure does not yield any informational gain for the decision maker. However, under disclosure, the decision maker's action upon message $\emptyset$ is distorted away from the unconditional mean 0.5. As a result, the decision maker's expected payoff condi-

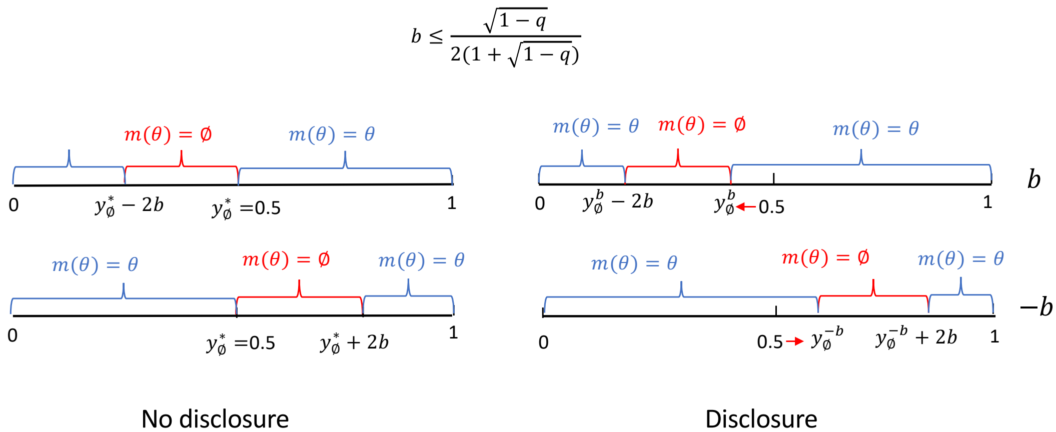

Figure 2: Equilibrium for small bias.

tional on the expert being uninformed is lower under disclosure than no disclosure. In this case, disclosure of conflicts of interest reduces the decision maker's welfare. The decision maker's loss from disclosure is $4q^2b^4/(1-q)$, which increases in b and q. As is shown in Proposition 3, when the expert's bias is known, the message $\emptyset$ will lead to an action further away from 0.5 if the expert has a larger bias or if he is more likely to be informed. It follows that the decision maker's loss from the distortion in action is larger.

Figure 3 illustrates the equilibrium when the expert's bias is large. As b increases, the expert withholds more information in both regimes as the interval in which the expert conceals information expands. However, the interval for concealing information expands faster under no disclosure than under disclosure when b is sufficiently large. In Figure 3, the length of the interval in which the expert withholds information is $2b$ under no disclosure and is less than $2b$ under disclosure. To see the latter, take the expert with the positive bias as an example. As b increases, $y_\emptyset^b$ keeps moving

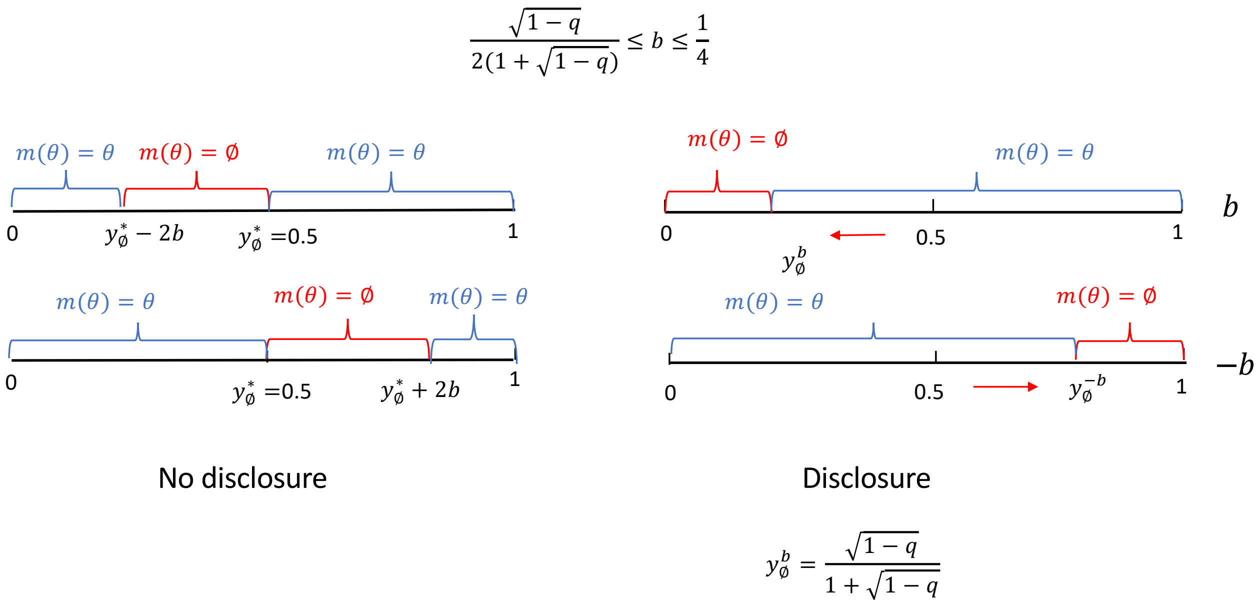

Figure 3: Equilibrium for large bias.

to the left until the lower bound of the interval $y^b_\emptyset - 2b$ reaches the boundary. After that, $y^b_\emptyset$ stays constant as b continues to increase. So, the interval for concealing information under disclosure is shorter than $2b$. Intuitively, if the expert has a large known bias and stays silent, the decision maker is suspicious that the state of the world is very unfavorable to the expert, and hence will take an action close to the extreme against the experts' interests. This speculation makes it more difficult for the expert to hide information than the case when he can pool with experts with an opposite bias. When the bias is greater than $\hat{b}(q)$, the information gain from disclosure dominates the loss from distortion in action. As a result, the decision maker is better off under disclosure.

Let us turn to the case in which the positive and negative bias values do not necessarily occur with equal probability. Without loss of generality, we focus on the case $p \geq 1/2$ as the complementary case is similar.

Note first that this will not affect the equilibrium characterization or the expert's and the decision maker's payoffs under disclosure.

In contrast, the characterization of the equilibria under nondisclosure will be different. In particular, there are now three cases. In the first case the two types of the expert each withholds information in intervals of length $2b$. For the positive-biased expert, this interval is to the left of the no-revelation action and for the negative-biased Sender, it is to the right of that action. The no-revelation action is less than 1/2, given that the positive-biased expert is more likely. The second case involves a hybrid equilibrium, where the positive-biased expert withholds information in all states that are below the no-revelation action, while the negative-biased expert withholds an interval of length $2b$. In the third case, the positive-biased expert withholds information in all states that are below the no-revelation action and the negative-biased expert withholds information in all states that are above the no-revelation action. Again, the no-revelation action is less than 1/2, given that the positive-biased expert is more likely.

Let $\tilde{b} \equiv \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})} \frac{\sqrt{1-2q(1-p)}-\sqrt{1-q}}{(2p-1)(1-\sqrt{1-q})}$.

Proposition 5. *Suppose $p \in [1/2, 1]$ and θ follows the uniform distribution. The unique equilibrium under nondisclosure has the following features:*

If $b \leq \tilde{b}$, the decision maker takes the action $y_\emptyset = \frac{1}{2} - \frac{2b^2(2p-1)q}{1-q}$ following the message $\emptyset$. The upward biased Expert withholds information if and only if $\theta \in (y_\emptyset - 2b, y_\emptyset)$ and the downward biased Expert withholds information if and only if $\theta \in (y_\emptyset, y_\emptyset + 2b)$.

If $\tilde{b} < b < \frac{1}{4}$, the decision maker takes the action

$$y_\emptyset = \frac{\sqrt{(1-q)^2 + pq(1-q+4q(1-p)b^2)} - (1-q)}{pq}$$

following the message $\emptyset$. The upward biased Expert withholds information if and only if $\theta \in [0, y_\emptyset)$ and the downward biased Expert withholds information if and only if $\theta \in (y_\emptyset, y_\emptyset + 2b)$.

Proof. Suppose that $2b \le y_\emptyset \le 1-2b$. Because $b \le \frac{1}{4}$, the set $[2b, 1-2b]$ is nonempty. Then, $y_\emptyset$ satisfies

$$\begin{aligned}(1-q)(\mu - y_\emptyset) &= q\left(p\int_{y_\emptyset-2b}^{y_\emptyset}(y_\emptyset-\theta)dF(\theta) + (1-p)\int_{y_\emptyset+2b}^{y_\emptyset}(y_\emptyset-\theta)dF(\theta)\right)\\ &= 2b^2q(2p-1).\end{aligned}$$

The unique solution is

$$y_\emptyset = \frac{1}{2} - \frac{q}{1-q}2b^2(2p-1).$$

The condition $y_\emptyset \ge 2b$ is satisfied if and only if

$$4q(2p-1)b^2 + 4b(1-q) - (1-q) \le 0, \tag{14}$$

which holds if

$$\begin{aligned}b &\le \frac{\sqrt{(1-q)^2 + q(1-q)(2p-1)} - (1-q)}{2q(2p-1)}\\ &= \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}\frac{\sqrt{1-2q(1-p)} - \sqrt{1-q}}{(2p-1)(1-\sqrt{1-q})}\\ &\equiv \tilde{b} > 0. \end{aligned} \tag{15}$$

The last inequality holds because $p > \frac{1}{2}$. The equality holds by substituting $q = 1 - (\sqrt{1-q})^2$. Because $y_\emptyset < \frac{1}{2}$, $y_\emptyset \ge 2b$ implies $y_\emptyset < 1 - 2b$.

Now, consider $y_\emptyset < 2b$. The action $y_\emptyset$ satisfies

$$\begin{aligned}(1-q)(\mu - y_\emptyset) &= q\left(p\int_{0}^{y_\emptyset}(y_\emptyset-\theta)dF(\theta) + (1-p)\int_{y_\emptyset}^{y_\emptyset+2b}(y_\emptyset-\theta)dF(\theta)\right) \quad (16)\\ &= q[\frac{py_\emptyset^2}{2} - 2(1-p)b^2].\end{aligned}$$

The above condition yields the unique solution

$$y_\emptyset = \frac{\sqrt{(1-q)^2 + pq(4b^2q(1-p)+1-q)} - (1-q)}{pq}$$

The condition $y_\emptyset \leq 2b$ is satisfied if and only if

$$h(b) \equiv 4q(2p-1)b^2 + 4b(1-q) - (1-q) \geq 0,$$

which holds if

$$\tilde{b} \leq b < \frac{1}{4}$$

To see that $(\tilde{b}, \frac{1}{4})$ is a nonempty set, note that $h(b)$ increases in b for $b > 0$. Since $h(1/4) > 0$ and $h(\tilde{b}) = 0$, we have $\tilde{b} < \frac{1}{4}$. □

Now, we compute the expected payoff of the decision maker under Disclosure and No Disclosure.

Proposition 6. *There exists a unique cutoff* $\hat{\hat{b}} \in (\frac{\sqrt{1-q}}{1+\sqrt{1-q}}, \tilde{b})$ *such that the decision maker's payoff is higher under No Disclosure than under Disclosure if and only if* $b < \hat{\hat{b}}$.

Proof. Proof for Proposition 6. First, consider $b \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$. By step 2 of the proof of Proposition 4, $y_\emptyset^b = \frac{1}{2} - \frac{2qb^2}{1-q}$ and $y_\emptyset^{-b} = \frac{1}{2} + \frac{2qb^2}{1-q}$. Hence, the decision maker's expected payoff under No Disclosure is

$$\begin{aligned} E^d(v) &= -p\left[(1-q)\int_0^1 (y_\emptyset^b - \theta)^2\, d\theta + q\int_{y_\emptyset^b - 2b}^{y_\emptyset^b} (y_\emptyset^b - \theta)^2\, d\theta\right] \\ &\quad -(1-p)\left[(1-q)\int_0^1 (y_\emptyset^{-b} - \theta)^2\, d\theta + q\int_{y_\emptyset^{-b}}^{y_\emptyset^{-b}+2b} (y_\emptyset^{-b} - \theta)^2\, d\theta\right] \\ &= -(1-q)\int_0^1 (y_\emptyset^b - \theta)^2 - q\int_0^{2b} t^2 dt, \qquad (17) \end{aligned}$$

where the last equality follows from $y_\emptyset^{-b} = 1 - y_\emptyset^b$ and $t \equiv y_\emptyset^b - \theta$.

Now, consider Disclosure. We can rewrite $\tilde{b} = \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}k$, where $k \equiv \frac{\sqrt{1-2q(1-p)}-\sqrt{1-q}}{(2p-1)(1-\sqrt{1-q})}$. We first show $\frac{\sqrt{1-q}}{2(1+\sqrt{1-q})} < \tilde{b}$. Recall that $\tilde{b}$ is the solution for (14) when the equality holds. By the implicit function theorem

$$\frac{\partial \tilde{b}}{\partial p} = -\frac{8q(\tilde{b})^2}{8q(2p-1)\tilde{b} + 4(1-q)} < 0.$$

When $p = 1$, $k = 1$ and $\tilde{b} = \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$. Hence, $\tilde{b} > \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$ for $p \in [\frac{1}{2}, 1)$. By Proposition 5, $y_\emptyset = \frac{1}{2} - \frac{2b^2(2p-1)q}{1-q}$, and

$$\begin{aligned} E(v) &= -(1-q)\int_0^1 (y_\emptyset - \theta)^2 \, d\theta \\ &\quad -pq\int_{y_\emptyset - 2b}^{y_\emptyset} (y_\emptyset - \theta)^2 \, d\theta - (1-p)q\int_{y_\emptyset}^{y_\emptyset + 2b} (y_\emptyset - \theta)^2 \, d\theta \\ &= -(1-q)\int_0^1 (y_\emptyset - \theta)^2 \, d\theta - q\int_0^{2b} t^2 dt. \end{aligned} \tag{18}$$

Take the difference

$$E^d(v) - E(v) = -(1-q)\left[\left(y^b_\emptyset - \frac{1}{2}\right)^2 - \left(y_\emptyset - \frac{1}{2}\right)^2\right]$$

Because $y^b_\emptyset < y_\emptyset \leq \frac{1}{2}$ for $p \in [1/2, 1)$, $E^d(v) < E(v)$ for $b \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}$.

Next, consider $\frac{\sqrt{1-q}}{2(1+\sqrt{1-q})} < b \leq \tilde{b}$. Step 2 of the proof of Proposition 4 shows that $y^b_\emptyset = \frac{\sqrt{1-q}}{1+\sqrt{1-q}}$ and $y^{-b}_\emptyset = \frac{1}{1+\sqrt{1-q}}$. The decision maker's payoff under Disclosure is

$$E^d(v) = -(1-q)\int_0^1 (y^b_\emptyset - \theta)^2 d\theta - \frac{q(y^b_\emptyset)^3}{3}. \tag{19}$$

Because $y^b_\emptyset$ is independent of b, $E^d(v)$ is constant in b. The decision maker's payoff under No Disclosure is (18) and is decreasing in b because $y_\emptyset = \frac{1}{2} - \frac{2b^2(2p-1)q}{1-q}$ decreases in b. Take the difference between (19) and (18),

$$\begin{aligned}
E^d(v) - E(v) &= -(1-q)\int_0^1 (y_\emptyset^b - \theta)^2 d\theta - \frac{q(y_\emptyset^b)^3}{3} \\
&\quad +(1-q)\int_0^1 (y_\emptyset - \theta)^2 \, d\theta + \frac{q(2b)^3}{3} \\
&= -(1-q)(y_\emptyset^b - y_\emptyset)(y_\emptyset^b + y_\emptyset - 1) + \frac{q}{3}((2b)^3 - (y_\emptyset^b)^3) \qquad (20)
\end{aligned}$$

Next, we evaluate $E^d(v) - E(v)$ at $\tilde{b}$. Note that $2\tilde{b} = y_\emptyset = y_\emptyset^b k$. Substituting $y_\emptyset^b k$ for $2\tilde{b}$ and $y_\emptyset$, we have

$$\begin{aligned}
E^d(v) - E(v)|_{\tilde{b}} &= -(1-q)(y_\emptyset^b - y_\emptyset^b k)(y_\emptyset^b + y_\emptyset^b k - 1) + \frac{q(y_\emptyset^b)^3}{3}(k^3 - 1) \\
&= y_\emptyset^b(k-1)\left((1-q)(y_\emptyset^b + y_\emptyset^b k - 1) + \frac{q(y_\emptyset^b)^2(k^2+k+1)}{3}\right) \\
&= \frac{y_\emptyset^b(k-1)(1-q)}{3(1+\sqrt{1-q})}\underbrace{\left((1-\sqrt{1-q})k^2 + (1+2\sqrt{1-q})k - (2+\sqrt{1-q})\right)}_{A},
\end{aligned}$$

where the third equality holds after substituting $y_\emptyset^b = \frac{\sqrt{1-q}}{1+\sqrt{1-q}}$ into the big bracket in the second equation. Because $k > 1$, the sign of $E^d(v) - E(v)|_{\tilde{b}}$ depends on the sign of A, which is quadratic in k and is increasing for $k > 0$. Recall that k decreases in p and $k = 1$ when $p = 1$. So, $k > 1$ for $p \in [\frac{1}{2}, 1)$. Because $A = 0$ at $k = 1$, $A > 0$ for $k > 1$. Hence, $E^d(v) - E(v)|_{\tilde{b}} > 0$ for $p \in [\frac{1}{2}, 1)$.

Recall that $E^d(v)$ is constant in b whereas $E(v)$ decreases in b for $\frac{\sqrt{1-q}}{1+\sqrt{1-q}} < b \leq \tilde{b}$. The difference $E^d(v) - E(v)$ is increasing in b. Because $E^d(v) < E(v)$ at $b = \frac{\sqrt{1-q}}{1+\sqrt{1-q}}$ and $E^d(v) > E(v)$ at $b = \tilde{b}$, there exists a unique cutoff $\hat{\hat{b}} \in (\frac{\sqrt{1-q}}{1+\sqrt{1-q}}, \tilde{b})$ such that $E^d(v) < E(v)$ if $b \leq \hat{\hat{b}}$ and $E^d(v) > E(v)$ if $\hat{\hat{b}} < b \leq \tilde{b}$.

Lastly, consider $b > \tilde{b}$. By Proposition 5, the decision maker's payoff under No

Disclosure is

$$\begin{aligned} E(v) &= -(1-q)\int_0^1 (y_\emptyset-\theta)^2\, d\theta \\ &\quad -q\left(p\int_0^{y_\emptyset}(y_\emptyset-\theta)^2\, d\theta+(1-p)\int_{y_\emptyset}^{y_\emptyset+2b}(y_\emptyset-\theta)^2\, d\theta\right) \\ &= -(1-q)\int_0^1 (y_\emptyset-\theta)^2 d\theta - q\left(p\int_0^{y_\emptyset}(y_\emptyset-\theta)^2\, d\theta+(1-p)\int_0^{2b} t^2 dt\right) \end{aligned}$$

Take the derivative

$$\begin{aligned} \frac{\partial E(v)}{\partial b} &= -8q(1-p)b^2+\left(-2(1-q)\int_0^1(y_\emptyset-\theta)d\theta-2pq\int_0^{y_\emptyset}(y_\emptyset-\theta)d\theta\right)\frac{\partial y_\emptyset}{\partial b} \\ &= -8q(1-p)b^2+\underbrace{\left(-pq(y_\emptyset)^2-2(1-q)y_\emptyset+(1-q)\right)}_{B}\frac{\partial y_\emptyset}{\partial b} \end{aligned}$$

The expression B is concave in $y_\emptyset$ with the larger root at $\bar{y_\emptyset}=\frac{\sqrt{(1-q)^2+pq(1-q)}-(1-q)}{pq}$. Since

$$y_\emptyset=\frac{\sqrt{(1-q)^2+pq(1-q+4q(1-p)b^2)}-(1-q)}{pq}>\bar{y_\emptyset},$$

$B<0$. Because $\frac{\partial y_\emptyset}{\partial b}>0$, $\frac{\partial E(v)}{\partial b}<0$. Given that $E^d(v)$ is constant in b whereas $E(v)$ decreases in b, $E^d(v)>E(v)$ at $\tilde{b}$ implies $E^d(v)>E(v)$ for $b>\tilde{b}$. We conclude that $E^d(v)<E(v)$ if and only if $b<\hat{\hat{b}}$. □

Proposition 7. *(i) The cutoff $\hat{\hat{b}}$ decreases in p and q. (ii) For a given $b<\hat{\hat{b}}$, the decision maker's payoff difference $E(v)-E^d(v)$ decreases in p for $p\in[\frac{1}{2},1)$.*

Proof. The cutoff $\hat{\hat{b}}$ is determined by setting (20) to zero. Let

$$G\equiv-(1-q)(y_\emptyset^b-y_\emptyset)(y_\emptyset^b+y_\emptyset-1)+\frac{q}{3}((2\hat{\hat{b}})^3-(y_\emptyset^b)^3),$$

where $y_\emptyset=\frac{1}{2}-\frac{2(\hat{\hat{b}})^2(2p-1)q}{1-q}$ and $y_\emptyset^b=\frac{\sqrt{1-q}}{1+\sqrt{1-q}}$. By the implicity function theorem,

$$\frac{\partial\hat{\hat{b}}}{\partial p}=-\frac{\frac{\partial G}{\partial y_\emptyset}\frac{\partial y_\emptyset}{\partial p}}{\frac{\partial G}{\partial\hat{\hat{b}}}+\frac{\partial G}{\partial y_\emptyset}\frac{\partial y_\emptyset}{\partial\hat{\hat{b}}}}<0,$$

where the inequality follows from $\frac{\partial G}{\partial y_\emptyset} = -(1-q) < 0$, $\frac{\partial y_\emptyset}{\partial p} = -\frac{4(\hat{\hat{b}})^2 q}{1-q} < 0$, $\frac{\partial G}{\partial \hat{\hat{b}}} = 8q(\hat{\hat{b}})^2 > 0$, $\frac{\partial y_\emptyset}{\partial \hat{\hat{b}}} = -\frac{4q(2p-1)\hat{\hat{b}}}{1-q} < 0$.

$$\frac{\partial \hat{\hat{b}}}{\partial q} = -\frac{\frac{\partial G}{\partial q} + \frac{\partial G}{\partial y_\emptyset}\frac{\partial y_\emptyset}{\partial q} + \frac{\partial G}{\partial y_\emptyset^b}\frac{\partial y_\emptyset^b}{\partial q}}{\frac{\partial G}{\partial \hat{\hat{b}}} + \frac{\partial G}{\partial y_\emptyset}\frac{\partial y_\emptyset}{\partial \hat{\hat{b}}}}.$$

We have shown that the denominator is positive. The numerator is positive because $\frac{\partial G}{\partial q} = \frac{(2\hat{\hat{b}})^3 - (y_\emptyset^b)^3}{3(1-q)} > 0$, $\frac{\partial y_\emptyset}{\partial q} = -\frac{2b^2(2p-1)}{(1-q)^2} < 0$, $\frac{\partial G}{\partial y_\emptyset^b} = (1-q)(1-2y_\emptyset^b) - q(y_\emptyset^b)^2 = 0$, $\frac{\partial y_\emptyset^b}{\partial q} = -\frac{(1-q)^{-1/2}}{2(1+\sqrt{1-q})^2} < 0$. As a result $\frac{\partial \hat{\hat{b}}}{\partial q} < 0$.

Next, we show (ii). Consider $b < \hat{\hat{b}}$, so $E(v) > E^d(v)$. By (20), we have

$$E(v) - E^d(v) = (1-q)(y_\emptyset^b - y_\emptyset)(y_\emptyset^b + y_\emptyset - 1) - \frac{q}{3}((2b)^3 - (y_\emptyset^b)^3).$$

Take the derivative

$$\frac{\partial(E(v) - E^d(v))}{\partial p} = (1-q)\frac{\partial y_\emptyset}{\partial p} = -4qb^2 < 0. \qquad \square$$

Corollary 7.1. *Mandatory disclosure policy improves the decision maker's welfare if the expert is more likely to be informed or has a large bias, and reduces the decision maker's welfare if the expert is less likely to be informed or has a small bias.*

The cutoff bias $\hat{b}(q)$ is decreasing in q. We illustrate the decision maker's utility with and without the disclosure policy in Figure 1. Concealing the expert's bias dominates disclosing the bias in the region below the downward sloping curve $\hat{b}(q)$ and vice versa in the region above the curve.

Corollary 7.2. *Ex ante, the expert prefers to not disclose his bias to the decision maker.*

Proof. We first compare the positive-biased expert's expected payoff under disclosure with that under nondisclosure. The argument for the negative-biased expert is similar.

First, consider the interior-cutoff equilibrium under disclosure. Note that if an interior-cutoff equilibrium occurs under disclosure, it must be that Type 1 equilibrium occurs under nondisclosure. In either case, the expert's expected payoff is the same when he is informed, as he hides θ in an interval of length $2b$ immediately to the left of the null action and fully reveals θ outside of that interval. However, under disclosure, the expert's expected payoff is lower when he is uninformed, as the null action is distorted further downward than it is under nondisclosure, which hurts a positive-biased expert. Hence, we conclude the positive-biased expert's expected payoff is higher under nondisclosure than under disclosure.

Now, consider the Type B equilibrium under disclosure. Consider first the extreme case that $q = 1$. Under disclosure, the expert fully reveals all information. Under nondisclosure, however, he only reveals information outside of the length $2b$ interval to the left of the null action. By revealed preferences, his payoff is higher under nondisclosure.

When $q \neq 1$, a similar revealed preferences argument applies when the expert is informed, as he manages to hide a longer interval of the state of the world under nondisclosure, where the induced action is at most b away from his ideal action. In the meantime when the expert is uninformed, he is also worse off under disclosure, for the same reason as that for the interior-cutoff equilibrium.

Given that the ex-interim payoffs of the expert is strictly lower under disclosure, regardless of his bias. His ex ante expected payoff is also strictly lower. □

Note that the argument does not rely on symmetry, so is true regardless of the bias distribution.

6 Discussion

In this subsection, we consider equilibrium characterization and welfare comparisons for other possible parameter combinations. In particular, we will discuss alternative distributions of bias β and then alternative distributions of state of the world θ.

6.1 General distribution of bias β

In this subsection, we consider the case where the expert's bias β has a general distribution on $[b_l, b_h]$, with distribution function G.[5] Without loss of generality, we consider the case where $b_h \geq |b_l|$. We limit our attention to the case where b_h is such that under disclosure the equilibrium is an interior-cutoff one, namely,

$$b_h \leq \frac{\sqrt{1-q}}{2(1+\sqrt{1-q})}.$$

Consider first the case where $b_l \geq 0$. The equilibrium under nondisclosure is characterized by the following equation:

$$(1-q)(\mu - y_\emptyset) = q\int_{b_l}^{b_h}\int_{y_\emptyset - 2b}^{y_\emptyset}(y_\emptyset - \theta)dF(\theta)dG(b). \tag{21}$$

For the disclosure case, the equilibrium under disclosure is characterized by the equation:

$$(1-q)(\mu - y_\emptyset^b) = q\int_{y_\emptyset^b - 2b}^{y_\emptyset^b}(y_\emptyset^b - \theta)dF(\theta), \tag{22}$$

To evaluate the welfare of the decision maker, we need to observe that the expected payoff of the decision maker is equal to

$$\begin{aligned} E(v) &= -(1-q)\int_0^1 (y_\emptyset - \theta)^2\, dF(\theta) \\ &\quad - q\int_{b_l}^{b_h}\int_{y_\emptyset - 2b}^{y_\emptyset}(y_\emptyset - \theta)^2 dF(\theta)\, dG(b). \end{aligned} \tag{23}$$

[5] Although we conduct our analysis assuming that β follows a continuous distribution, we can extend the analysis to discrete and hybrid distributions.

In the disclosure case, the decision maker's expected payoff is

$$E^d(v) = \int_{b_l}^{b_h} \left[-(1-q)\int_0^1 (y_\emptyset^b - \theta)^2 \, dF(\theta) - q \int_{y_\emptyset^b - 2b}^{y_\emptyset^b} (y_\emptyset^b - \theta)^2 dF(\theta) \right] dG(b) \tag{24}$$

Under uniform distribution and after some algebraic calculations, the above equations become

$$\begin{aligned} E(v) &= -(1-q)\int_0^1 \left(\frac{1}{2} - \theta\right)^2 d\theta - (1-q)\left(y_\emptyset - \frac{1}{2}\right)^2 \\ &\quad -qE\left(\frac{8}{3}b^3\right). \end{aligned}$$

In the disclosure case, the decision maker's expected payoff is

$$\begin{aligned} E^d(v) &= -(1-q)\int_0^1 \left(\frac{1}{2} - \theta\right)^2 d\theta - (1-q)\int_{b_l}^{b_h} \left(y_\emptyset^b - \frac{1}{2}\right)^2 dG(b) \\ &\quad -qE\left(\frac{8}{3}b^3\right). \end{aligned}$$

Thus, the difference in the decision maker's payoff between the nondisclosure case and the disclosure case can be written

$$E(v) - E^d(v) = -(1-q)\left[\left(y_\emptyset - \frac{1}{2}\right)^2 - \int_{b_l}^{b_h} \left(y_\emptyset^b - \frac{1}{2}\right)^2 dG(b)\right].$$

Note, from the equilibrium conditions (21) and (**??**), we have

$$\begin{aligned} \left(y_\emptyset - \frac{1}{2}\right)^2 &= \frac{q^2}{(1-q)^2}\left[E(2b^2)\right]^2, \\ &\leq \frac{q^2}{(1-q)^2} E\left[(2b^2)^2\right], \\ &= \int_{b_l}^{b_h} \left(y_\emptyset^b - \frac{1}{2}\right)^2 dG(b), \end{aligned}$$

where the equalities are implied by the equilibrium conditions and the inequality uses the convexity of the square function. Furthermore, the inequality is strict unless the distribution of b is degenerate. Hence, the decision maker's payoff under nondisclosure is higher.

In the case $b_l \leq 0$, note that

$$\left(y_\emptyset - \frac{1}{2}\right)^2 \;\leq\; \left(\bar{y}_\emptyset - \frac{1}{2}\right)^2,$$

while

$$\left(y_\emptyset^b - \frac{1}{2}\right)^2 \;=\; \left(\bar{y}_\emptyset^h - \frac{1}{2}\right)^2,$$

where the $\bar{y}$ expressions correspond to the value when b is replaced by $|b|$ for all $b \in [b_l, b_h]$ (thus creating a new distribution function $\bar{G}$). Hence, the welfare comparisons are further enforced.

6.2 Welfare comparison: Non-uniform distribution

In this subsection, we consider an example in which the dominance of nondisclosure over disclosure fails, even when under the latter type A equilibrium occurs.

Let $p = 1/2$ and $b_h = -b_l = b$. Instead of assuming that $\theta \in [0,1]$ follows the uniform distribution, let its density function f be

$$f(x) = \begin{cases} t + 2(2-2t)x, & x \in [0, 1/2], \\ t + 2(2-2t)(1-x), & x \in [1/2, 1], \end{cases}$$

where $t \in [0,1]$. Thus, the density function is symmetric around 1/2 and single-peaked.[6] When $t = 1$, θ follows the uniform distribution; when $t = 0$, θ has zero density at 0. Assume further that b is such that under disclosure the equilibrium is of type A.

The equilibrium under nondisclosure is characterized by the following equation:

$$(1-q)(\mu - y_\emptyset) \;=\; q\left[\frac{1}{2}\int_{y_\emptyset - 2b}^{y_\emptyset}(y_\emptyset - \theta)dF(\theta) + \frac{1}{2}\int_{y_\emptyset}^{y_\emptyset + 2b}(y_\emptyset - \theta)dF(\theta)\right]. \quad (25)$$

[6] If $t \in [1,2]$, then the distribution is single-troughed.

For the disclosure case, the equilibrium under disclosure is characterized by the equation:

$$(1-q)(\mu - y_\emptyset^h) \;\; = \;\; q\int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta) dF(\theta) \tag{26}$$

Note that (25) is satisfied by $y_\emptyset = \mu = 1/2$. The LHS of (26) is clearly decreasing in $y_\emptyset^h$, while the RHS of (26) is increasing in $y_\emptyset^h$ for all $y_\emptyset^h \leq 1/2$. To see the latter, one may use Leibniz's Rule or note that

$$\int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta) dF(\theta) \;\; = \;\; \int_0^{2b} x f(y_\emptyset^h - x) dx,$$

which is clearly increasing in $y_\emptyset^h \in [0, 1/2]$ as f is increasing on $[0, 1/2]$. Thus, there is a unique $y_\emptyset^h \in (0, 1/2)$ that solves (26).

Now, we compare the receiver's expected payoffs under disclosure and nondisclosure. Note that, as before and by symmetry, under nondisclosure the receiver's expected payoff is

$$\begin{aligned} E(v) \;\; &= \;\; -(1-q)\int_0^1 (y_\emptyset - \theta)^2 \, dF(\theta) - q\int_{y_\emptyset - 2b}^{y_\emptyset} (y_\emptyset - \theta)^2 dF(\theta), \\ &= \;\; -(1-q)\int_0^1 \left(\frac{1}{2} - \theta\right)^2 \, dF(\theta) - (1-q)\left(y_\emptyset - \frac{1}{2}\right)^2 - q\int_{y_\emptyset - 2b}^{y_\emptyset} (y_\emptyset - \theta)^2 dF(\theta). \end{aligned}$$

In the disclosure case, the receiver's expected payoff is

$$\begin{aligned} E^d(v) \;\; &= \;\; -(1-q)\int_0^1 (y_\emptyset^h - \theta)^2 \, dF(\theta) - q\int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta)^2 dF(\theta), \\ &= \;\; -(1-q)\int_0^1 \left(\frac{1}{2} - \theta\right)^2 \, dF(\theta) - (1-q)\left(y_\emptyset^h - \frac{1}{2}\right)^2 - q\int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta)^2 dF(\theta). \end{aligned}$$

Note that $y_\emptyset = 1/2$ and $y_\emptyset^h < 1/2$. Thus, substituting the pdf f into the above expressions, the difference in the receiver's payoff between the nondisclosure case and the disclosure case can be written

$$E(v) - E^d(v) \;\; = \;\; (1-q)\left(y_\emptyset^h - \frac{1}{2}\right)^2 - q\int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta)^2 \cdot 2(2-2t)\left(y_\emptyset - y_\emptyset^h\right) d\theta.$$

Using $y_\emptyset = 1/2$, $\mu = 1/2$, and (26), we have

$$\begin{aligned}
\frac{E(v) - E^d(v)}{1/2 - y_\emptyset^h} &= q \int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta) \left[t + 2(2-2t)\theta\right] d\theta - q \int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta)^2 \cdot 2(2-2t) d\theta, \\
&= q \int_{y_\emptyset^h - 2b}^{y_\emptyset^h} (y_\emptyset^h - \theta) \left[4(2-2t)\theta - 2(2-2t) y_\emptyset^h + t\right] d\theta.
\end{aligned}$$

Note that when $t = 1$, the above expression is always strictly positive, as previously shown. However, when $t = 0$, the above expression can be written

$$q \int_0^{2b} x \left[-8x + 4y_\emptyset^h\right] dx,$$

which is strictly negative $(-16b^3/3)$ for $y_\emptyset^h = 2b$ (the threshold between type A and type B equilibria). Hence, it is no longer true that nondisclosure always dominates disclosure when b is such that type A equilibrium occurs under disclosure.

7 Conclusion

In this paper, we study the effect of disclosure of conflicts of interest on communication between an expert and a client. We show that even when the expert's information is verifiable, there may be circumstances under which disclosure of conflicts of interest is not conducive to informative communication. In particular, we identify an important tradeoff between disclosure of conflicts of interest and utilization of the expert's information. If the expert is not necessarily informed, then an informed expert would have the option of feigning ignorance and not being forced to reveal all the information he has, as the typical unravelling argument implies. This, however, would cause the client to draw unwarranted inference from the uninformed expert's inability to provide information. On the other hand, when the direction of the bias of the expert is not disclosed, this negative effect is smaller.

Our paper makes a contribution to the theoretical literature on the effect of mandatory disclosure policies, as well as that on communication of verifiable information when the expert's bias is uncertain. We identify new environments in which mandatory disclosure policies are counterproductive, in addition to that of Li and Madarasz (2008).

It is worth noting that our analysis is focused on an environment in which both the expert and the decision maker are fully rational. There is evidence that a client tends not to make the full negative inference about the expert's action to conceal information.[7] This may undermine the unravelling argument needed for the full revelation equilibrium in the case of disclosure of conflicts of interest and possibly tilt the comparison more in favour of nondisclosure of conflicts of interest.

References

Bhattacharya, S. and A. Mukherjee (2013). Strategic information revelation when experts compete to influence. *The RAND Journal of Economics 44*(3), 522–544.

Brown, A. L., C. F. Camerer, and D. Lovallo (2012). To review or not to review? Limited strategic thinking at the movie box office. *American Economic Journal: Microeconomics 4*(2), 1–26.

Cain, D. M., G. Loewenstein, and D. A. Moore (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies 34*(1), 1–25.

Chung, W. and R. Harbaugh (2019). Biased recommendations from biased and unbiased experts. *Journal of Economics & Management Strategy 28*(3), 520–540.

[7] See Brown et al. (2012) for empirical evidence and Jin et al. (2015) for experimental evidence.

Crawford, V. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1431–1452.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research*, 123–145.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics 24*(3), 461–483.

Hagenbach, J., F. Koessler, and E. Perez-Richet (2014). Certifiable Pre-Play Communication: Full Disclosure. *Econometrica 83*(3), 1093–1131.

Ismayilov, H. and J. Potters (2013). Disclosing advisor's interests neither hurts nor helps. *Journal of Economic Behavior & Organization 93*, 314 – 320.

Jin, G. Z., M. Luca, and D. Martin (2015). Is no news (perceived as) bad news? An experimental investigation of information disclosure. Technical report, National Bureau of Economic Research.

Jung, W.-O. and Y. K. Kwon (1988). Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research 26*(1), 146–153.

Li, M. and K. Madarasz (2008). When mandatory disclosure hurts: Expert advice and conflicting interests. *Journal of Economic Theory 139*(1), 47–74.

Loewenstein, G., C. R. Sunstein, and R. Golman (2014). Disclosure: Psychology changes everything. *Annual Review of Economics 6*(1), 391–419.

Mezzetti, C. (2020). Manipulative disclosure. Technical report.

Milgrom, P. (1981, Autumn). Good news and bad news: Representation theorems and applications. *Bell Journal of Economics 12*(2), 380–391.

Seidmann, D. J. and E. Winter (1997). Strategic information transmission with verifiable messages. *Econometrica 65*(1), 163–170.

Shavell, S. (1989). Sharing of information prior to settlement or litigation. *The RAND Journal of Economics 20*(2), 183–195.

Wolinsky, A. (2003). Information transmission when the sender's preferences are uncertain. *Games and Economic Behavior 42*(2), 319–326.