Principal-induced stubbornness in experts*

Yuk-fai Fong[†], Xiaoxiao Hu[‡], Xiaoxuan Meng[§], Kirill Novoselov[¶]

August 2022

Abstract

A principal hires an expert to collect information and then make a decision, utilizing both the expert's private information and informative public opinion. The optimal contract induces the expert to sometimes defy public opinion even when public opinion is more informative than his private information. Our finding is robust to allowing for switching the arrival times of different signals, expert reporting his private information, expert's reputational concern and repeated interactions.

Keywords: Principal-induced stubbornness, Information acquisition, Optimal contract

JEL Codes: D82, D86

1 Introduction

In many principal-agent relationships, the principal relies on the agent to *both* analyze her problem and take an appropriate action to solve her problem. For example, shareholders rely on the CEO to assess the merits of various investment opportunities, and then choose one investment that supposedly maximizes shareholder value.

*The first author would like to acknowledge the RGC Grant (Project Number 640712) for its financial support and thank the organizing committees of the Econometric Society Australasian Meeting and the Singapore Economic Review Conference for invitations to present as an invited speaker. We also thank the audiences for useful comments at these conferences and the Midwest Economic Theory Conference, the 12^{th} Workshop on Industrial Organization and Management Strategy, the Asian Econometric Society Meeting and the Hong Kong University of Science and Technology. We thank Yu Fu Wong for valuable inputs.

[†]Faculty of Business and Economics, The University of Hong Kong

[‡]Economics and Management School, Wuhan University

[§]Faculty of Business and Economics, The University of Hong Kong

[¶]School of Accountancy, Shanghai University of Finance and Economics

Similarly, citizens rely on the political leader to analyze the pros and cons of different public policy options and then choose a policy that supposedly maximizes social welfare. For the agent, to effectively analyze the principal's problem, he has to exert effort to collect useful information. In many of these situations, apart from relying on information that he has collected on his own, the agent can also utilize publicly available information to guide his decision-making. Examples of such public information include commentaries by various pundits, such as Wall Street analysts or political commentators.[1] The principal would also like the agent to appropriately take into consideration publicly available information in his decision-making. So, motivating the agent to collect useful information is as important as motivating him to choose the right decision for the principal given the available information.

Inspired by the above observations, this paper analyzes the *optimal contract* in a principal-agent relationship where the principal relies on the agent to first generate an informative private signal relevant to the principal's interest, and then utilize both the *private signal* and a publicly observed signal, which we refer to as *public opinion*, to choose an investment project or a public policy on the principal's behalf. Given our focus on investigating the motivation of the agent's information acquisition and utilization, we make the simplifying assumption that the action of choosing an investment project or a public policy does not require effort.

We find that it is optimal for the principal to encourage the agent to overly rely on his private signal vis-à-vis public opinion. We refer to the agent's bias in favor of the private signal, induced by the principal, as *principal-induced stubbornness*. This distortion is created by rewarding the agent for defying public opinion. More specifically, conditional on an investment success, the agent is paid a higher bonus if he acted contrary to public opinion than if he acted in accordance with public opinion.

Due to the existence of informative public opinion, the expert has an incentive to under-provide effort and free ride on public opinion. The potential benefit of inducing stubbornness in the expert stems from the stochastic relationship between the effort level and the precision of the private signal. Due to the stochastic nature of the precision of the private signal, at any effort level, the agent may follow either the private signal or the public opinion depending on the realized precision of the private signal. If there is an additional bonus for an investment success when the expert defies public opinion, the agent will follow his private signal even if it is less (but not significantly less) informative than public opinion. Anticipating himself to rely more on the private signal, the agent is motivated to exert more efforts to improve his private signal. In contrast, if the agent's effort determines the private

[1] Such examples abound. For concrete ones, see "Analysts Suggest Google's Parent Should Acquire AIG," *Fortune.com*, February 16, 2016 and "The Case for a Liberal Scalia: Why President Obama Should Nominate a True Progressive to the High Court," *Slate.com*, February 17, 2016.

signal precision, then the agent can anticipate whether the realized private signal can affect his decision, and inducing stubbornness would not help motivating effort.[2]

Our findings are consistent with the not-uncommon observations that politicians disregard public opinion,[3] and that some political leaders with a reputation for being stubborn are rewarded and those who readily change their position in response to public opinion or new information are punished. For example, in the 2004 U.S. presidential election, voters elected George W. Bush, who had long been known for being stubborn on issues, such as tax and stem cell research, over John Kerry, who had changed positions on issues, such as the War in Iraq, in response to newly available information. Although one can argue that changing one's view in light of new information is simply Bayesian updating, and it is what an able politician should do, Kerry was accused of being a "flip-flopper" and lost votes because of this. An earlier often-cited example of voters' punishment of flip-flopping concerns Dick Gephardt, an unsuccessful candidate for the Democratic nomination in the 1988 presidential election. Many commentators felt that Gephardt's chance was destroyed by an advertisement aired by the campaign of his opponent, Governor Michael Dukakis, which accused him of "flip-flopping" using his voting record. Gephardt's defense—*"I'd rather change and be right than be rigid and be wrong"*—was not enough to save his campaign.[4]

Our paper provides a novel explanation for expert stubbornness. In the existing literature, other explanations of experts' stubbornness include the expert's career concern and expert's incentives to signal his high ability (e.g., Prendergast and Stole (1996); Levy (2004)). Both of these explanations require assuming that the expert has private information about his own abilities.[5] In this paper, we show that it can be optimal for the principal to induce expert stubbornness even if the expert's information does not reflect his talent and that he is not motivated by career concerns. It is also important to note that according to career concern, expert stubbornness is initiated by the expert and motivated by the expert's interest, whereas according to our theory, expert stubbornness is induced by the principal.

[2]In Section 4.4.2, we provide a detailed analysis for the case of the deterministic private signal precision.

[3]For instance, an article published on *Columbia University Record*, dated September 25, 2000, cited Political Scientists Robert Shapiro and Lawrence Jacobs as arguing that "presidents and members of Congress routinely disregard the policy preferences expressed in public opinion polls."

[4]Please see https://en.wikipedia.org/wiki/Flip-flop_(politics) and Rosenbarum, David (1988), "Of Political Flip-Flops in the '88 Democratic Race," *New York Times*.

[5]According to Prendergast and Stole (1996), where expert stubbornness is described as an exaggeration of the expert's information, "*exaggeration of information can occur only if the quality of the information itself reflects the manager's talent*" and according to Levy (2004), where stubbornness is described as anti-herding, "*Anti-herding results are derived in the literature in several contexts and under a variety of assumptions. I find that in terms of the decision maker's objectives, the decision maker must be motivated by career concerns*."

2 Literature

The literature on strategic information transmission of experts has investigated the expert's incentive to use stubbornness to signal his ability. Prendergast and Stole (1996) pioneered this idea, and Levy (2004) is another notable example. On the other hand, experts may also have opposite signaling incentives. For instance, Prat (2005) and Ottaviani and Sørensen (2006) have shown that, to improve their perceived abilities, experts may herd, or exhibit a common tendency to bias their predictions towards what a capable expert would predict instead of providing the most accurate forecasts Based on their private information. Li (2007) considers two consecutive private signals and showed that "mind changes" can signal quick improvement in ability. In contrast with these papers, our paper abstracts away from signaling incentives or career concerns by analyzing a setting in which there is only one type of expert. Our perspective on stubbornness is different. In our setting of pure moral hazard, the expert's stubbornness is intentionally and optimally induced by the principal for the principal's benefit instead of being initiated by the expert to advance the expert's benefit. The principal may choose to reward stubborn behavior even if the stubborn behavior is not a sign of higher ability. By showing how stubbornness arises in a new set of environments not previously reported, our findings complement existing ones to suggest that stubbornness in experts can be more widespread than previously considered and does not disappear even after the principal learns the expert's ability.

Our paper is also related to the literature on motivating information acquisition. The papers closest to ours are Prendergast (1993), Li (2001), Szalay (2005), Che and Kartik (2009) and a contemporaneous and complementary paper Terovitis (2018). Prendergast (1993) studies the trade-off between ex ante effort incentive and ex post optimal information transmission. In particular, Prendergast (1993) concerns the trade-off between motivating the worker to gather information, which inevitably causes the worker to conform when the firm uses subjective performance evaluation, and encouraging the worker to truthfully report his information. Most related to our paper are Li (2001), Szalay (2005) and Che and Kartik (2009), which also study using ex post sub-optimal mechanisms to motivate ex ante information acquisition. In Li (2001), by committing ex ante to a conservative decision rule, the committee members are less tempted to free-ride on the other members' effort. In Che and Kartik (2009), by hiring an agent with a different prior from her own, the principal would induce more effort from the agent. In Szalay (2005), by restricting the agent to only take extreme actions, the principal forces the agent to collect a more accurate signal. In line with these studies, in our setting, the principal distorts the agent's investment decision to motivate him to exert more effort on information acquisition. However, due to important differences in our settings, the distortion that we identify differs from the distortions identified in these studies, and as a result, our studies explain different phenomena and have different applications. Our setting differentiates from theirs in that (i) the agent's effort only affects the distribution of the signal precision rather than fully determines the accuracy of his private signal, (ii) the principal and the agent do not have a shared objective and their interests are only aligned through payments, (iii) we introduce informative, yet imprecise, public opinion which the agent can free ride on. The contemporaneous paper Terovitis (2018) studies a similar problem as ours in the financial recommendation setting. Terovitis (2018) assumes a binary state, a binary effort level and, in contrast to our model, there is no public opinion. Terovitis (2018) finds that the optimal contract induces the expert to act against the prior more often than that of the first-best case.

Two recent studies, Häfner and Taylor (2018) and Bardey et al. (2020), also analyze agent's information acquisition incentives. In Häfner and Taylor (2018)'s dynamic setting with a flow of entrepreneurs, every entrepreneur under-invests from the perspective of a social planner due to the fact that externalities on his successors are ignored. In Bardey et al. (2020)'s one-shot setting, the buyer switches between sellers to induce information acquisition. Due to the lack of free-riding problem in information acquisition and the deterministic nature of the signal precision, over-utilization of information, or stubbornness, is not generated in either paper.

Stubborn behaviors are also studied in other disciplines such as psychology, finance and management. They refer to such behavior as overconfidence, resistance to persuasion or intellectual arrogance. Most of the literature on overconfidence and intellectual arrogance takes such behavior as a primitive and studies its benefit and cost.[6] Building upon the evidence in psychology, behavioral economists study the motive of overconfidence. The seminal paper Bénabou and Tirole (2002) models an intra-personal game and demonstrates that overconfidence can be beneficial for a player with time-inconsistent preference. The literature on resistance to persuasion explains such behavior using behavioral motives, such as concerns of deception.[7] Our model is free from behavioral assumptions, and thus provides an alternative view—the expert's stubborn behavior can be the result of an optimal contract in the principal's interest.

3 Model

There are two players: a principal (she) and an expert (he). The principal relies on the expert to collect information and then make a decision. There is an unknown state of nature θ with two possible values, θ_L and θ_H. Both players share a uniform common prior regarding the state.

[6] See, for example, Bernardo and Welch (2001); Hsu, Novoselov, and Wang (2017); Owens, Johnson, and Mitchell (2013); Malmendier and Tate (2005, 2008).

[7] See Fransen, Smit, and Verlegh (2015) for a review.

Expert's private signal The expert can exert some effort level $e \geq 0$ at a cost $c(e)$ to generate an informative, yet imperfect signal s whose distribution depends on both the state and the expert's effort level. Specifically, signal s takes values s_L and s_H with

$$\Pr(s_L \mid \theta_L) = \Pr(s_H \mid \theta_H) = x,\ \Pr(s_H \mid \theta_L) = \Pr(s_L \mid \theta_H) = 1 - x$$

where $x \in X \equiv [0.5, 1]$ measures the precision of the signal and is drawn according to the cumulative distribution function (CDF):

$$H(x; e) = (1 - p(e))F(x) + p(e)G(x). \quad (1)$$

Here $H(\cdot\,; e)$ is a convex combination of two exogenous CDFs, $F(\cdot)$ and $G(\cdot)$. Both F and G admit PDFs, f and g respectively, that are strictly positive over X. Expert's effort level e, the realized signal s, and the signal precision x are all privately observed by the expert.

Throughout this paper, we impose the following assumptions:

(i) The distribution G first-order stochastically dominates the distribution F in that $F(x) > G(x)$ for all $x \in (0.5, 1)$.

(ii) The weight function $p(\cdot)$ is continuously differentiable and strictly increasing with $p(0) = 0$, $\lim_{e\to\infty} p(e) = 1$ and $p'(e) > 0$ for all $e \in (0, \infty)$.

(iii) The expert's cost function $c(e)$ is continuously differentiable and strictly increasing with $c(e) = 0$ and $c'(e) > 0$ for all $e \in (0, \infty)$.

(iv) The ratio $\frac{c'(e)}{p'(e)}$ is strictly increasing in e, $\lim_{e\to 0} \frac{c'(e)}{p'(e)} = 0$, $\lim_{e\to\infty} \frac{c'(e)}{p'(e)} \geq 1$ and $\lim_{e\to 0} p'(e)/\big(\frac{c'(e)}{p'(e)}\big)' = +\infty$.[8]

Public opinion After the expert privately observes the realized signal, Nature generates a public signal σ that takes values σ_L and σ_H according to

$$\Pr(\sigma_H \mid \theta_H) = \Pr(\sigma_L \mid \theta_L) = q,\ \Pr(\sigma_L \mid \theta_H) = \Pr(\sigma_H \mid \theta_L) = 1 - q.$$

We refer signal σ as the *public opinion*, whose distribution and realization are known to both players. The precision of public opinion is measured by the exogenous parameter $q \in (0.5, 1)$.

After Nature draws public opinion, the expert takes an action $a \in \{a_H, a_L\}$, generating payoffs of $r(a, \theta)$ to the principal with

$$r(a_L, \theta_L) = r(a_H, \theta_H) = 1, r(a_L, \theta_H) = r(a_H, \theta_L) = 0.$$

Payoffs from the state-matching action are normalized to one.

[8] Our assumptions on $p(\cdot)$ and $c(\cdot)$ guarantee that expert's optimal effort level should be positive.

Contracting The principal hires the expert to gather information and chooses an action on behalf of her. At the beginning of the game, the principal signs a court-enforceable contract with the expert. A contract specifies the principal's payments to the expert which can be contingent on all publicly observable information: the expert's chosen action a, the realized state θ and the realized public opinion σ. A contract is characterized by the principal's payment function $\hat{w} : \{a_L, a_H\} \times \{\theta_L, \theta_H\} \times \{\sigma_L, \sigma_H\} \to [0, +\infty)$. Here we assume that the expert is protected by limited liability; that is, $\hat{w}(a, \theta, \sigma)$ is always non-negative.

Payoffs The principal's payoff function is

$$u^P(a, \theta, \sigma) = r(a, \theta) - \hat{w}(a, \theta, \sigma) \tag{2}$$

and the expert's payoff function is

$$u^E(e, a, \theta, \sigma) = \hat{w}(a, \theta, \sigma) - c(e). \tag{3}$$

The values of both players' outside options are assumed to be 0. Note that the limited liability assumption implies that the expert's participation constraint always holds.

The timing of the game is as follows.

1. The principal commits to a payment scheme $\hat{w}$.
2. The expert exerts effort level $e \in [0, \infty)$. Then he privately observes the realized signal precision $x \in (0.5, 1)$ and the realized private signal $s \in \{s_L, s_H\}$.
3. Both players observe the realized public opinion, $\sigma \in \{\sigma_L, \sigma_H\}$.
4. The expert chooses an action $a \in \{a_L, a_H\}$.
5. The state of nature is realized, $\theta \in \{\theta_L, \theta_H\}$. Player's payoffs are specified in equations (2) and (3).

The timeline is summarized in Figure 1.

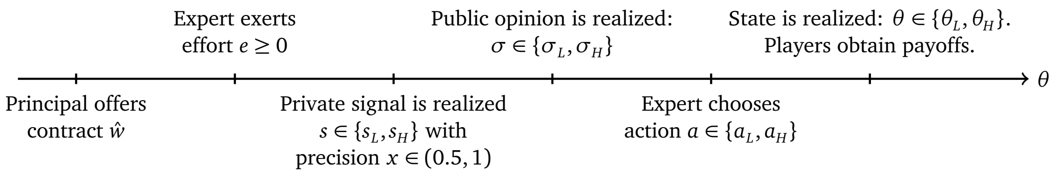

Figure 1: Timeline

3.1 Solution concept

Denote by $\rho \in (0,1)$ the expert's posterior belief regarding the state being θ_L after observing the realizations: x, s and σ. Fixing some payment scheme $\hat{w}$, the *fixed-payment equilibrium under* $\hat{w}$ consists of expert's belief updating $\hat{\rho} : x,s,\sigma \mapsto \hat{\rho}(x,s,\sigma) \in [0,1]$; expert's action rule $\hat{a} : \rho,\sigma \mapsto \hat{a}(\rho,\sigma) \in \{a_L, a_H\}$; expert's effort level $e^* \geq 0$; such that

1. Expert's belief updating process specified by $\hat{\rho}(x,s,\sigma)$ is consistent with Bayesian updating for any $x \in [0.5,1]$, $s \in \{s_L, s_H\}$ and $\sigma \in \{\sigma_L, \sigma_H\}$:

$$\hat{\rho}(x,s,\sigma) = \frac{\Pr(s \mid \theta_L, x)\Pr(\sigma \mid \theta_L)}{\Pr(s \mid \theta_L, x)\Pr(\sigma \mid \theta_L) + \Pr(s \mid \theta_H, x)\Pr(\sigma \mid \theta_H)}.$$

2. Expert's action $\hat{a}(\rho,\sigma)$ solves

$$\max_a \rho \hat{w}(a, \theta_L, \sigma) + (1-\rho)\hat{w}(a, \theta_H, \sigma)$$

   for any $\rho \in [0,1]$ and $\sigma \in \{\sigma_L, \sigma_H\}$. The expert's continuation value under the action rule $\hat{a}$ is

$$W(\rho,\sigma) := \rho \hat{w}(\hat{a}(\rho,\sigma), \theta_L, \sigma) + (1-\rho)\hat{w}(\hat{a}(\rho,\sigma), \theta_H, \sigma).$$

3. Given $\hat{a}$ and $\hat{\rho}$, expert's effort level e^* solves

$$\max_{e \geq 0} \int_{0.5}^{1} U^E(x)\, dH(x;e) - c(e),$$

   where $U^E(x)$ is the expert's continuation value after observing x and

$$U^E(x) = \mathbb{E}_{\{s,\sigma\}}[W\big(\hat{\rho}(x,s,\sigma),\sigma\big) \mid x].$$

Our solution concept is the *principal-preferred subgame-perfect equilibrium*: the principal commits to some payment scheme to maximize her expected payoff in the fixed-payment equilibrium. Her expected payoff in the fixed-payment equilibrium under $\hat{w}$ is

$$\hat{\pi}(\hat{w}) = \mathbb{E}_{\{\theta,\sigma,x,s\}}\Big[r\Big(\hat{a}(\hat{\rho}(x,\sigma,s),\sigma),\theta\Big);\, e^*\Big] - \int_{0.5}^{1} U^E(x)\, dH(x;e^*).$$

Whenever the expert is indifferent between some actions, he is assumed to take an action which maximizes the principal's expected payoff.

Some non-standard expressions Throughout this paper, we say some contract $\hat{w}$ *dominates* another contract $\hat{w}'$ if the principal obtains strictly higher expected payoffs in the fixed-payment equilibrium under $\hat{w}$: $\hat{\pi}(\hat{w}) > \hat{\pi}(\hat{w}')$; we say the action a_i *matches* the (realized) signal, s_j or σ_j, or the state θ_j if $i = j$ for $i, j \in \{L, H\}$; we say the expert's action rule is to *follow* the private signal or public opinion if he always chooses the action that matches the realization of the corresponding signal.

3.2 Symmetric contract

An expert's action is *Good* (or resp., *Bad*) if it matches (or resp., mismatches) the realized state; an expert's action is *Approved* (or resp., *Disapproved*) if it matches (or resp., mismatches) the realized public opinion. Based on these two notions, we call the payment scheme $\hat{w}$ a *symmetric contract* if the principal's payment is conditional on (i) whether the chosen action is Good or Bad and (ii) whether the chosen action is Approved or Disapproved.

Denote a symmetric contract by $\boldsymbol{w} \equiv (w_{GA}, w_{GD}, w_{BA}, w_{BD})$, whose meaning and relation to the payment function $\hat{w}(a, \theta, \sigma)$ are depicted in Table 1.

Expert's Action	*Contingencies*	*Payment*
Good, Approved	$\{(a_L, \theta_L, \sigma_L), (a_H, \theta_H, \sigma_H)\}$	w_{GA}
Good, Disapproved	$\{(a_L, \theta_L, \sigma_H), (a_H, \theta_H, \sigma_L)\}$	w_{GD}
Bad, Approved	$\{(a_L, \theta_H, \sigma_L), (a_H, \theta_L, \sigma_H)\}$	w_{BA}
Bad, Disapproved	$\{(a_L, \theta_H, \sigma_H), (a_H, \theta_L, \sigma_L)\}$	w_{BD}

Table 1: Symmetric contract $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$

Our main analysis focuses on symmetric contracts due to the symmetrical environment setting.[9] Symmetric contracts also seem more realistic and fairer compared to non-symmetric contracts. Somewhat surprisingly, the optimal contracts for the principal are non-symmetric. Our main result holds in both the optimal symmetric contract and the optimal general contacts.

4 Analysis

4.1 Benchmark I: First-best case

Consider first the benchmark case where the principal can acquire information and choose an action by herself. Given the realized private signal s_i with the realized

[9] Specifically, the prior distribution of the binary state is uniform; the conditional distributions of the private signal and public opinion are symmetric; the payoffs $r(a, \theta)$ are symmetric.

precision x and the realized public opinion σ_j, the principal's optimal response is to follow the signal with higher precision; that is, the optimal action rule is

$$\alpha^{FB}(x, s_i, \sigma_j) = \begin{cases} a_i & \text{if } x > q; \\ a_j & \text{otherwise,} \end{cases}$$

for $i \in \{L, H\}$ and $j \in \{L, H\}$. Under this action rule, the principal's continuation value after observing the signal s is $\max\{x, q\}$. Therefore, the first-best effort level e^{FB} solves

$$\max_{e \geq 0} \int_{0.5}^{1} \max\{x, q\}\, dH(x; e) - c(e). \tag{4}$$

Recall that $H(x; e) = p(e)G(x) + (1 - p(e))F(x)$. Our assumptions on $p(\cdot)$ and $c(\cdot)$ ensure that there exists a unique solution to the maximization problem (4). Using integration by parts, the first-order condition of the maximization problem is reduced to

$$\int_{q}^{1} (F(x) - G(x))dx = \frac{c'(e^{FB})}{p'(e^{FB})}. \tag{5}$$

To sum up, the first-best effort level is e^{FB} as determined in equation (5); the first-best action rule α^{FB} is to follow the more informative signal between the private signal and public opinion.

4.2 Benchmark II: Contracts independent of public opinion

Consider the principal's optimal contracting problem. We first focus on the family of symmetric contracts whose payments are not contingent on public opinion; that is, the payments are contingent solely on whether the action is Good or Bad. We call contracts of this kind *Contracts I*, defined as below.

Definition 1 (Contract I). A *Contract I* is a tuple (w_G, w_B), where $w_G (\geq 0)$ and $w_B (\geq 0)$ are payments for a Good action and a Bad action respectively.

Denote the optimal Contract I by $(\bar{w}_G, \bar{w}_B)$. To solve for the optimal contract I, it is without loss of generality to focus on those Contracts with $w_G > w_B$. The reason is that when $w_G < w_B$, the expert is encouraged to choose the Bad action, and thus his incentives will be severely misaligned with the principal's; moreover, Contracts I with $w_G = w_B$ are dominated by another Contract I, $(\varepsilon, 0)$, for some positive ε.

Lemma 1. *In the optimal Contract I, $\bar{w}_G > \bar{w}_B$.*

Proof. See Appendix A1. □

Under $w_G > w_B$, the expert follows the more informative signal and his action rule is still $\hat{a}^{FB}$. The expert chooses the effort level by solving the following maximization problem:

$$\max_{e\geq 0} \int_{0.5}^{q} (qw_G + (1-q)w_B)\, dH(x;e) + \int_{q}^{1} (xw_G + (1-x)w_B)\, dH(x;e) - c(e). \quad (6)$$

Our assumptions on $p(e)$ ensure that the maximization problem has a unique solution. The first-order condition yields:

$$(w_G - w_B)\int_{q}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}. \quad (7)$$

Equation (7) implies that only the payment difference $(w_G - w_B)$ matters in determining the expert's effort. As a result, it's optimal for the principal to pay nothing for a Bad action: $\bar{w}_B = 0$. Given that, the principal solves the following maximization problem:

$$\begin{aligned} &\max_{w_G\in(0,1)} \int_{0.5}^{q} (1-w_G)q\, dH(x;e) + \int_{q}^{1} (1-w_G)x\, dH(x;e) \\ &\text{subject to } w_G \int_{q}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}. \end{aligned} \quad (8)$$

Denote by $\bar{e}$ the expert's effort level under the optimal contract I, $(\bar{w}_G, 0)$. Since $\frac{c'(e)}{p'(e)}$ is strictly increasing and $\bar{w}_G < 1$, equations (5) and (7) imply $\bar{e} < e^{FB}$; that is, the expert under-provides efforts in the optimal Contract I compared to that of the first-best case. Proposition 1 summarizes our analysis of the benchmark case of Contracts I.

Proposition 1. *Under the optimal Contract I,*

(i) the expert follows public opinion if $x < q$ and follows his private signal if $x > q$, which is the same as that of the first-best case;

(ii) the expert under-provides efforts compared to that of the first-best case: $\bar{e} < e^{FB}$.

By focusing on the family of Contracts I, we have only considered the symmetric contracts which are independent of public opinion. Indeed, the identified optimal Contract I yields the highest payoff for the principal among general contracts whose payments are not contingent on public opinion. Furthermore, while there also exist other non-symmetric contracts independent of public opinion that yield the same payoff for the principal as that of the optimal Contract I, the induced effort levels and expert's action rules are the same as those in Proposition 1. See Appendix A2 for detailed analysis.

4.3 Optimal contracts

Consider symmetric contracts where the principal's payments can be contingent on both whether the action is Good or Bad and whether the action is Approved or Disapproved. Denote by $(\bar{w}_{GA}, \bar{w}_{GD}, \bar{w}_{BA}, \bar{w}_{BD})$ the principal's optimal contract. To solve for the optimal contract, it is without loss of generality to focus on contracts (w_{GA}, w_{GD}, w_{BA}, w_{BD}) that satisfy $w_{GA} > w_{BD}$ and $w_{GD} > w_{BA}$. Put in other words: if public opinion matches the state, the Good action should be rewarded more than the Bad action ($w_{GA} > w_{BD}$); if public opinion mismatches the state, the Good action should also be rewarded more than the Bad action ($w_{GD} > w_{BA}$). This result is formally stated as in Lemma 2, whose proof is similar to that of Lemma 1.

Lemma 2. *In the optimal symmetric contract,* $\bar{w}_{GA} > \bar{w}_{BD}$ *and* $\bar{w}_{GD} > \bar{w}_{BA}$.

Proof. See Appendix A3. □

4.3.1 Expert's strategy

We first derive the expert's action rule given a symmetric contract (w_{GA}, w_{GD}, w_{BA}, w_{BD}) satisfying $w_{GA} > w_{BD}$ and $w_{GD} > w_{BA}$. Consider the case where the expert's private signal differs from public opinion. When the expert observes a private signal s_L yet public opinion is σ_H, expert's posterior belief regarding the state being θ_L is

$$\hat{\rho}(x, s_L, \sigma_H) = \frac{x(1-q)}{x(1-q) + q(1-x)} =: \rho^*.$$

The expert will follow his private signal (and defy public opinion) if and only if $\rho^* w_{GD} + (1-\rho^*) w_{BD} \geq (1-\rho^*) w_{GA} + \rho^* w_{BA}$, which is equivalent to $x \geq x^*$ where

$$x^* = \frac{q(w_{GA} - w_{BD})}{(1-q)(w_{GD} - w_{BA}) + q(w_{GA} - w_{BD})}. \qquad (x^*)$$

For the symmetrical situation where the expert observes s_H and σ_L, by similar arguments the expert will follow his private signal if and only if $x > x^*$.

Then consider the case where the private signal agrees with public opinion. When the expert observes s_L and σ_L, his posterior belief regarding the state being θ_L is

$$\hat{\rho}(x, s_L, \sigma_L) = \frac{xq}{xq + (1-x)(1-q)} =: \rho^{**};$$

similarly, the expert follows his private signal (and his action is Approved) if and only if $x \geq x^{**}$ where

$$x^{**} = \frac{(1-q)(w_{GD} - w_{BA})}{(1-q)(w_{GD} - w_{BA}) + q(w_{GA} - w_{BD})}. \qquad (x^{**})$$

Both x^* and x^{**} lie in $(0,1)$ and $x^* + x^{**} = 1$. Since the signal precision $x \in [0.5,1]$, only one of the two constraints $x \geq x^*$ and $x \geq x^{**}$ can be valid. Indeed, under the optimal symmetric contract only constraint $x \geq x^*$ is valid, and $x \geq x^{**}$ always holds as $x^{**} < 0.5$.

Lemma 3. *Under the optimal symmetric contract,*

$$0 < \bar{w}_{GD} - \bar{w}_{BA} < \frac{q}{1-q}(\bar{w}_{GA} - \bar{w}_{BD}); \tag{Act}$$

the expert's action rule is to follow his private signal if $x > x^*$ and otherwise follow public opinion.

Proof. See Appendix A4. □

Lemma 3 implies that when private signal agrees with public opinion, under the optimal contract the expert will follow the common realization. This is consistent with the usual intuition as players' incentives will be severely misaligned otherwise. Moreover, when the private signal differs from public opinion, under the optimal symmetric contract the expert will follow his private signal if and only if the realized signal precision is above the cutoff x^*. This is in contrast with the previous two benchmark cases where the expert follows his private signal if and only if $x > q$. Later we show that $x^* < q$ under the optimal symmetric contract; that is, it's optimal for the principal to induce the expert to sometimes defy public opinion even though his private signal is less informative than public opinion.

Expert's effort level Lemma 3 implies that when $x > x^*$, the expert's action is Good with probability x and his continuation payoff is $x[qw_{GA} + (1-q)w_{GD}] + (1-x)[qw_{BD} + (1-q)w_{BA}]$; when $x < x^*$, the expert's action is Good with probability q and his continuation payoff is $(qw_{GA} + (1-q)w_{BA})$. Overall, the expert's ex ante expected payoff is:

$$V^E(e) := \int_{x^*}^{1} [x(qw_{GA} + (1-q)w_{GD}) + (1-x)(qw_{BD} + (1-q)w_{BA})]\, dH(x;e) + \int_{0.5}^{x^*} (qw_{GA} + (1-q)w_{BA})\, dH(x;e) - c(e). \tag{9}$$

The expert chooses $e \geq 0$ to maximize $V^E(e)$. First-order condition of this maximization problem yields:

$$\big(q(w_{GA} - w_{BD}) + (1-q)(w_{GD} - w_{BA})\big) \int_{x^*}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}. \tag{IC}$$

The incentive compatibility constraint (IC) determines expert's effort level in the fixed-payment equilibrium.

4.3.2 Characterizing optimal symmetric contract

Finally, the principal chooses a symmetric contract subject to three sets of constraints:

(i) the constraints regarding the expert's action rule as specified by (Act) and the cutoff is determined in (x^*);

(ii) the incentive compatibility constraint (IC) regarding the expert's effort level;

(iii) the limited liability constraints:

$$w_{GA} \geq 0, w_{BD} \geq 0, w_{GD} \geq 0, w_{BA} \geq 0. \tag{LL}$$

Formally, the principal chooses $(w_{GA}, w_{BD}, w_{GD}, w_{BA})$ to maximize her ex ante expected payoff $U^P = \int_{x^*}^{1} xh(x;e)\,dx + \int_{0.5}^{x^*} qh(x;e)\,dx - V^E(e) - c(e)$ subject to constraints (Act), (x^*), (IC) and (LL). An immediate observation is that constraints (Act), (x^*) and (IC) are only concerned with the payment differences, $w_{GD} - w_{BA}$ and $w_{GA} - w_{BD}$. Then if $w_{BA} > 0$ (or resp., $w_{BD} > 0$), the principal can lower w_{GD} and w_{BA} (or resp., w_{GA} and w_{BD}) simultaneously by some $\varepsilon > 0$ to obtain higher expected profits. Therefore, constraints $w_{BA} \geq 0$ and $w_{BD} \geq 0$ are binding in the optimal symmetric contract.

Lemma 4. *In the optimal symmetric contract, $\bar{w}_{BA} = \bar{w}_{BD} = 0$.*

By Lemma 4, the principal solves the following simplified maximization problem:

$$\max_{w_{GA}\geq 0, w_{GD}\geq 0} \int_{0.5}^{x^*} (q - qw_{GA})dH(x;e) + \int_{x^*}^{1} \left[x - \left(xqw_{GA} + x(1-q)w_{GD}\right)\right]dH(x;e)$$

subject to

$$\begin{aligned} &0 < w_{GD} < \frac{q}{1-q} w_{GA}, \quad x^* = \frac{qw_{GA}}{(1-q)w_{GD} + qw_{GA}}, \\ &\left(qw_{GA} + (1-q)w_{GD}\right) \int_{x^*}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}. \end{aligned} \tag{$\mathcal{M}$}$$

The solution to problem ($\mathcal{M}$) may be one of the three kinds: $w_{GA} = w_{GD}$, $w_{GA} < w_{GD}$ and $w_{GA} > w_{GD}$. Contracts of the first kind are specialized Contract I, and we refer the second and the third kind as *Contracts S* and *Contracts F* respectively. A Contract S induces the expert to be *stubborn* in the sense that he acts as if he has bias in favor of his private signal (i.e., $x^* < q$); a Contract F induces the expert to be a *flip-flopper* in the sense that he is prone to public opinion (i.e., $x^* > q$). Below we formally define Contract S and Contract F, and discuss their implications.

Definition 2 (Contracts S and F)**.** A Contract S is a tuple $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ satisfying $w_{GD} > w_{GA} > 0$ and $w_{BA} = w_{BD} = 0$. A Contract F is a tuple $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ satisfying $w_{GA} > w_{GD} > 0$ and $w_{BA} = w_{BD} = 0$.

Under a Contract S, a Good action will be rewarded more if it is also Disapproved. Recall that the expert follows his own signal if and only if $x \geq x^* = \frac{qw_{GA}}{(1-q)w_{GD}+qw_{GA}}$. When $w_{GD} > w_{GA}$, we have $x^* < q$. Therefore, compared to the first-best case, a Contract S induces the expert to be less reliant on public opinion. Similarly, under a Contract F, a Good action will be rewarded more if it is also Approved, which encourages the expert to conform more with public opinion.

By Proposition 2, the optimal symmetric contract is a Contract S. The proof sketch is as follows. For any Contract F, there exists some Contract I yielding higher payoffs for the principal. In addition, the optimal Contract I is dominated by some Contract S.

Proposition 2. *The optimal symmetric contract is a contract S, which induces the expert to be stubborn.*

Proof. See Appendix A5. □

By Proposition 2, it is beneficial for the principal to induce the expert to rely more on his private signal than public opinion: when public opinion differs from the private signal, the expert is sometimes incentivized to defy public opinion even though it is more informative than the private signal. We call this behavior *principal-induced stubbornness*. Our theory predicts when the expert's behavior exhibits stubbornness. Specifically, the expert is induced to inefficiently defy public opinion only when his private signal is slightly less informative than public opinion: $x \in (x^*, q)$.

The intuition of principal-induced stubbornness is as follows. Due to the existence of public opinion, the expert can free ride by exerting fewer efforts and following public opinion. To deal with the free-riding problem, the principal pays an additional bonus for a Good action which also defies public opinion by setting $w_{GD} > w_{GA}$, leading to the expert being stubborn. With the additional bonus, the expert's expected payoff is disproportionately changed when he follows public opinion or follows the private signal. More specifically, if the expert follows public opinion, he will not obtain the additional bonus with certainty; if he follows the private signal, he will obtain the additional bonus with probability $1-q$ conditional on his action being Good. As a result, the expert will rely more on the private signal. Anticipating that, he is motivated to exert a higher effort level ex ante. This effort-motivating aspect of a Contract S allows us to construct a Contract S based on the optimal Contract I such that a higher effort level is attained at a lower cost for the principal.[10]

[10] In particular, we take the optimal contract I $(\bar{w}_G, \bar{w}_B)$, and construct a Contract S that pays ε less for the Good and Approved action and $\frac{q}{1-q}\varepsilon$ more for the Good and Disapproved action, i.e., $\tilde{w}_{GA} = \bar{w}_G - \varepsilon$ and $\tilde{w}_{GD} = \bar{w}_G + \frac{q}{1-q}\varepsilon$ for some small $\varepsilon > 0$. See Appendix A5 for the details.

On the other hand, there is some efficiency loss associated with a Contract S since it shifts the cutoff x^* below q. However, a marginal shift in x^* only has a second-order effect given that q is the cutoff of the first-best action rule.[11] Therefore, the induced stubbornness in the expert can strictly increase the principal's expected payoffs.

Proposition 2 implies that Contracts F are never optimal. The intuition for the non-optimality of Contract F is as follows. Under a Contract F, a Good action will be rewarded more when it is approved by public opinion. Anticipating his own reduced reliance on the private signal the expert will exert less efforts ex ante. Moreover, the cutoff x^* is distorted upwards, leading to inefficient information utilization which further decreases the principal's payoffs.

4.3.3 Non-symmetric contracts

Symmetric contracts are more natural in our setting and seem more realistic as well. In the above analysis, we have only considered the family of symmetric contracts. In Appendix A6, we allow contracts to be non-symmetric and show that there exist non-symmetric contracts yielding strictly higher payoffs for the principal than that of the optimal symmetric contract (Claim 7). However, the optimal general contracts still induce expert stubbornness (Claim 8).

4.4 Discussions

We discuss two important assumptions, the assumption of limited liability and the precision of private signal being stochastic. These assumptions are necessary for the emergence of principal-induced stubbornness. For exposition convenience, our discussions are based on the family of symmetric contracts.

4.4.1 Limited liability

We have assumed that the expert is protected by limited liability; that is, the principal's payment is always non-negative. Without that assumption, the previous analysis on the expert's action rule and effort rule still apply, and both the cutoff x^* and optimal effort level are determined solely by $(w_{GD}-w_{BA})$ and $(w_{GA}-w_{BD})$. However, the expert's (ex ante) participation constraint[12] will be valid. By similar arguments for Lemma 4, the participation constraint is binding under the optimal symmetric contract. Therefore, the principal can extract all surplus. Let $\Delta w_{GD} = w_{GD} - w_{BA}$

[11] This second-order effect is reflected in equation (16) in Appendix A5.

[12] The expert's participation constraint is $\int_{x^*}^{1}[x(qw_{GA} + (1-q)w_{GD}) + (1-x)(qw_{BD} + (1-q)w_{BA})]dH(x;e) + \int_{0.5}^{x^*}(qw_{GA} + (1-q)w_{BA})dH(x;e) - c(e) \geq 0$.

and $\Delta w_{GA} = w_{GA} - w_{BD}$. Plugging the binding participation constraint into the principal's expected payoff, the principal's maximization problem becomes

$$\max_{\Delta w_{GD}\geq 0, \Delta w_{GA}\geq 0} \int_{0.5}^{x^*} q dH(x;e) + \int_{x^*}^{1} x dH(x;e) - c(e)$$

$$\text{subject to } x^* = \frac{q\Delta w_{GA}}{(1-q)\Delta w_{GD} + q\Delta w_{GA}}, \quad 0 < \Delta w_{GD} < \frac{q}{1-q}\Delta w_{GA},$$

$$\big(q\Delta w_{GA} + (1-q)\Delta w_{GD}\big)\int_{x^*}^{1}(F(x) - G(x))dx = \frac{c'(e)}{p'(e)}.$$

Setting $\Delta w_{GD} = \Delta w_{GA} = 1$, the cutoff action rule will be $x^* = q$ and the IC condition for effort will be the same as that in the first-best benchmark (Equation (5)). Therefore, the principal achieves the first-best payoffs, and such contracts[13] must be optimal for the principal.

To sum up, without the limited liability assumption, the effort level and action rule are the same as those of the first-best benchmark, and stubbornness is not induced in an optimal contract. We remark that if we consider long-term relationships between the principal and expert, then stubbornness can still be induced without imposing the limited liability assumption. See Section 5.4 for a brief discussion or the online Appendix for detailed analysis.

4.4.2 Stochastic precision of private signal

Suppose that the precision of private signal is deterministic. Specifically, given the exerted effort level $e \geq 0$, the precision of the expert's private signal is $x(e)$ deterministically. Assume $x(e)$ is a continuously differentiable and strictly increasing function satisfying $x(0) = 0.5$, $\lim_{e\to\infty} x(e) = 1$, $x'(e) > 0$ for all $e > 0$. Other similar assumptions are also imposed: $c'(e)/x'(e)$ is strictly increasing in e, $\lim_{e\to 0} c'(e)/x'(e) = 0$ and $\lim_{e\to+\infty} c'(e)/x'(e) > 1$.

To make the analysis nontrivial we also assume that the first-best case involves putting in positive amount of effort; that is,

$$x(e^*) - c(e^*) \geq q, \tag{A}$$

where $e^* > 0$ is the first-best effort given by $x'(e^*) = c'(e^*)$. In this case, the private signal is always followed since Assumption (A) implies that $x(e^*) > q$.

Consider the contracting problem between the principal and the expert. Denote by $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ a general symmetric contract. Due to the deterministic nature of the private signal precision, the principal will induce a positive effort level

[13] There exist multiple optimal contracts without the limited liability assumption. Any contract satisfying $\Delta w_{GD} = \Delta w_{GA} = 1$ and the binding participation constraint yields the first-best payoffs for the principal.

from the expert only if the generated private signal is more informative than public opinion. Therefore, under the optimal contract, there are two possible situations regarding the effort level:

1. The expert chooses some positive effort level and then follows his private signal:

   $$\max_e x(e)\, q w_{GA} + x(e)(1-q)\, w_{GD} + (1-x(e))(1-q)\, w_{BA} + (1-x(e))\, q w_{BD} - c(e).$$

   The expert's optimal effort level e^{**} is determined by

   $$q(w_{GA} - w_{BD}) + (1-q)(w_{GD} - w_{BA}) = c'(e^{**})/x'(e^{**}). \tag{10}$$

2. The expert chooses zero effort level and always follows public opinion. His expected payoff is $qw_{GA} + (1-q)w_{BD}$.

While by Assumption (A) the first-best case involves a positive effort level, it may not be optimal for the principal to induce expert's efforts under the optimal contract. The reason is as follows. Since inducing effort is costly, the principal will induce some positive effort level only if $x(e^{**}) > q$. However, $x(e^{**}) \leq q$ may still hold even though we impose Assumption (A). If a contract induces efforts, we must have $qw_{GA} + (1-q)w_{GD} < 1$ or otherwise the principal's expected payoff would be non-positive. Therefore, $e^{**} < e^{*}$ according to our assumptions on $x(e)$ and $c(e)$. It follows that when $q \in (x(e^{**}), x(e^{*}))$, the principal will not induce efforts in the optimal contract.

To sum up, without the assumption of stochastic private-signal precision, principal-induced stubbornness will not emerge under the optimal contract regardless of whether the principal induces efforts or not.

5 Extensions

In this section, we consider several variations of our main model and find that induced stubbornness is a robust feature of the optimal (symmetric) contracts.

5.1 Public opinion realized before expert exerting efforts

Suppose that public opinion is realized before the expert exerts efforts. This variation allows us to examine whether the principal has incentives to, if possible, delay the expert's effort choice. Figure 2 depicts the new timeline. The main distinction is that now the expert's effort choice can be contingent on public opinion. So a general

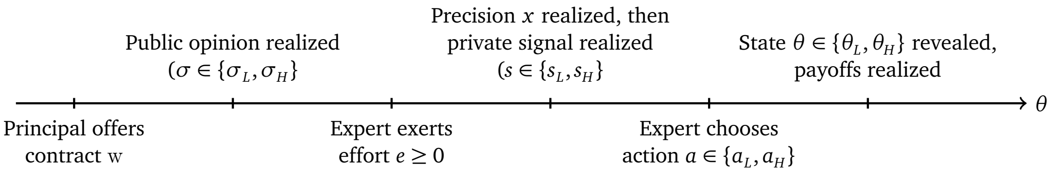

Figure 2: Timeline when public opinion is realized before expert's efforts

contract can induce two effort levels, e^H and e^L, and which one to be taken depends on the realized public opinion.

The principal offers two payment schemes, $\boldsymbol{w}^L$ and $\boldsymbol{w}^H$, contingent on the realized public signal and induce two possible effort levels. Denote by $(\bar{\boldsymbol{w}}^L, \bar{\boldsymbol{w}}^H)$ the principal's optimal contract. First, the expert's action rule under $(\bar{\boldsymbol{w}}^L, \bar{\boldsymbol{w}}^H)$ is still $\hat{a}(x, s, \sigma)$ as in Lemma 3 and the principal never rewards a Bad action. Given public opinion σ_i for $i \in \{L, H\}$, the expert will follow his private signal if $x > x^i = \frac{qw^i_{GA}}{qw^i_{GA} + (1-q)qw^i_{GD}}$ and otherwise will follow public opinion. Moreover, the incentive-compatible constraint for effort level given σ_i is

$$\left(qw^i_{GA} + (1-q)w^i_{GD}\right)\int_{x^i}^{1}(F(x) - G(x))dx = \frac{c'(e^i)}{p'(e^i)}.$$

The principal solves the following two maximization problems for $i \in \{L, H\}$:

$$\max_{w^i_{GA} \geq 0, w^i_{GD} \geq 0} \int_{0.5}^{x^i}(q - qw^i_{GA})dH(x; e^i) + \int_{x^i}^{1}\left[x - \left(xqw^i_{GA} + x(1-q)w^i_{GD}\right)\right]dH(x; e^i)$$

subject to

$$\begin{aligned} &0 < w^i_{GD} < \frac{q}{1-q}w^i_{GA}, \quad x^i = \frac{qw^i_{GA}}{(1-q)w^i_{GD} + qw^i_{GA}}, \\ &\left(qw^i_{GA} + (1-q)w^i_{GD}\right)\int_{x^i}^{1}(F(x) - G(x))dx = \frac{c'(e^i)}{p'(e^i)}. \end{aligned} \tag{$\mathcal{M}'$}$$

This maximization problem ($\mathcal{M}'$) is the same as that of solving for the optimal symmetric contract in the main model. Therefore, stubbornness is still induced under the optimal contract.

Proposition 3. *When public opinion is realized before the expert exerts efforts, the optimal contract induces stubbornness.*

Finally, the principal's highest payoff in this variation is the same as that of the optimal symmetric contract in the main model. Since the optimal contracts in the main model are non-symmetric, delaying the expert's effort choice harms the principal when she can commit to non-symmetric contracts.

5.2 Expert reporting private information

In this variation, we assume that the principal does not fully delegate the decision to the agent. Instead, the principal commits to some action rule at the beginning of the game, and later the expert reports to the principal his relevant private information—the private signal s and its precision x. Denote by $\hat{a}(\tilde{x}, \tilde{s}, \sigma)$ the action rule where $\tilde{x}$ and $\tilde{s}$ are the expert's reported signal precision and reported private signal respectively. Timing of the game is as follows.

1. The principal commits to some symmetric payment scheme $\boldsymbol{w}$ and action rule $\hat{a}$.
2. The expert exerts some effort level $e \geq 0$.
3. The expert privately observes the private signal $s \in \{s_L, s_H\}$ and its precision $x \in [0.5, 1]$. Public opinion $\sigma \in \{\sigma_L, \sigma_H\}$ is revealed to both players.
4. The expert reports $(\tilde{x}, \tilde{s})$ and the principal chooses action $\hat{a}(\tilde{x}, \tilde{s}, \sigma)$.
5. The state is revealed and players' payoffs are realized.

We call the combination $(\boldsymbol{w}, \hat{a})$ a *mechanism*. A mechanism is *incentive-compatible* if the induced expert's reports are truthful. Denote by $(\boldsymbol{w}^*, \hat{a}^*)$ the principal's optimal incentive-compatible mechanism. The principal's expected payoffs under $(\boldsymbol{w}^*, \hat{a}^*)$ must be weakly higher than that under the symmetric contract of the main model. First, we show that the action rule $\hat{a}^*$ takes the same cutoff form as that of the main model: the principal follows the reported private signal if the reported signal is above some cutoff $\tilde{x}^*$, and otherwise follows public opinion.

Lemma 5. *In the optimal incentive-compatible mechanism, the action rule $\hat{a}^*$ takes the cutoff form for some cutoff $\tilde{x}^* \in (0.5, 1)$ with*

$$\hat{a}^*(x, s_i, \sigma_j) = \begin{cases} a_i & \text{if } x > \tilde{x}^* \\ a_j & \text{otherwise} \end{cases}$$

for $i, j \in \{L, H\}$ and $\tilde{x}^* = \frac{q(w^*_{GA} - w^*_{BD})}{(1-q)(w^*_{GD} - w^*_{BA}) + q(w^*_{GA} - w^*_{BD})}$.

Proof. See Appendix A7. □

Given the functional form of the optimal action rule as in Lemma 5, the principal's maximization problem is the same as that of the main model.[14]

[14]In the optimal delegation literature, it's well noted that delegation of authority to the agent is equivalent to asking the agent for information and promising to behave in a particular way (Holmström, 1980). While we obtain a similar result, the equivalence in our setting is not obvious in foresight as we (i) allow monetary transfers between players and (ii) explicitly model the expert's information acquisition process.

Proposition 4. *In the optimal incentive-compatible mechanism, $\tilde{x}^* < q$.*

Proposition 4 shows that stubbornness is still induced in the optimal incentive-compatible mechanism. Such a result may be less intuitive compared to the same result in the main model. Now the principal can incentivize the expert through the payments as well as the committed action rule. It may seem optimal for the principal to commit to an ex post optimal action rule and only use the payments to incentivize the expert. However, due to the incentive compatibility requirement, the action rule and the payments cannot be independently designed. Indeed, the incentive-compatible action rule is already pinned down once a payment scheme $\hat{w}$ is fixed. Knowing this, the previous intuition for stubbornness applies: committing to an action rule that is ex post not optimal helps inducing a higher effort level from the expert ex ante.

5.3 Expert has reputational concern

In this variation, we maintain the timing of the main model but assume that the expert has extra reputational payoff from choosing a Good action. Specifically, the expert obtains an additional benefit $b > 0$ besides the principal's payment when the action is Good. We interpret it as the expert's reputational concern. Note that the principal does not pay for this benefit and that the benefit is a constant regardless of whether the Good action is Approved or Disapproved.

Denote by $\mathrm{w} \equiv \{w_{GA}, w_{GD}, w_{BA}, w_{BD}\}$ a symmetric contract. From the expert's perspective, a Good and Approved action rewards $w_{GA} + b$ and a Good and Disapproved action rewards $w_{GD} + b$. Lemma 4 still holds that under an optimal contract any Bad action will not be rewarded by the principal. The principal solves:

$$\max_{w_{GA}\geq 0, w_{GD}\geq 0} \int_{0.5}^{x^*} (q - q w_{GA}) dH(x;e) + \int_{x^*}^{1} \left[x - \left(xq w_{GA} + x(1-q) w_{GD}\right)\right] dH(x;e)$$

subject to

$$\begin{aligned} &x^* = \frac{q(w_{GA}+b)}{(1-q)(w_{GD}+b) + q(w_{GA}+b)}, \quad w_{GD} + b < \frac{q}{1-q}(w_{GA}+b), \\ &\left(q w_{GA} + (1-q) w_{GD} + b\right) \int_{x^*}^{1} (F(x) - G(x)) dx = \frac{c'(e)}{p'(e)}. \end{aligned} \tag{11}$$

Intuitively, when b is high enough, the principal may "free ride" on the expert's reputational concern and never reward a Good action. However, as long as the principal pays for a Good action under the optimal symmetric contract, then the expert stubbornness will still be induced.

Proposition 5. *When the expert has reputational concern, stubbornness is induced in the optimal symmetric contract as long as it involves a positive payment.*

Proof. See Appendix A8. □

The intuition for the stubbornness result in Proposition 5 is the same as that of the main model: stubbornness motivates a higher effort level from the expert.

5.4 Repeated interactions

In many real-life settings, explicit incentive contracts are not plausible and the players rely on the future value of the relationship to enforce the desired behavior. In this variation, we consider such repeated interactions between the principal and the expert where no court enforceable contracts are available and show that similar stubbornness results could be obtained. Following the literature, we call the equilibria in a repeated game with transfers the *relational contracts* (Levin, 2003). In this subsection, we only briefly discuss the properties of the optimal relational contract. Please refer to the online Appendix for the descriptions of the repeated interactions and the analysis.

We drop the limited liability assumption as there is an endogenous "limited liability constraint" in the relational contract setting: the maximum difference between possible payments is bounded above by the discounted future value of the relationship.[15] The discount factor δ and the precision of public opinion q critically affect this endogenous limited liability constraint, and thus they determine whether or not stubbornness is induced in an optimal relational contract.

Figure 3 illustrates the kinds of the optimal relational contracts under different values of δ and q. When δ is sufficiently close to one, the endogenous limited liability is not binding and the first-best case is achievable. The exact cutoffs of such δ depend on q since a higher q leads to a higher surplus in the relationship. When the first-best case is not achievable, unlike the one-shot base game in which stubbornness is always induced, the optimal relational contract may be a Contract I when q is below some cutoff q^*. The reason is that inducing stubbornness has an additional cost for the principal in the relational contract setting. To induce stubbornness, the difference between the average payments for a Good and a Bad action must be reduced as required by the endogenous limited liability constraint, which weakens the expert's incentives to exert efforts. Lastly, there is an additional channel to induce stubbornness in the relational contract setting. Besides paying different rewards for Good actions conditional on whether they are Approved or not, the principal may also reward a Bad and Disapproved action in an optimal relational contract.

[15] This is called the dynamic enforcement (DE) constraint in Levin (2003).

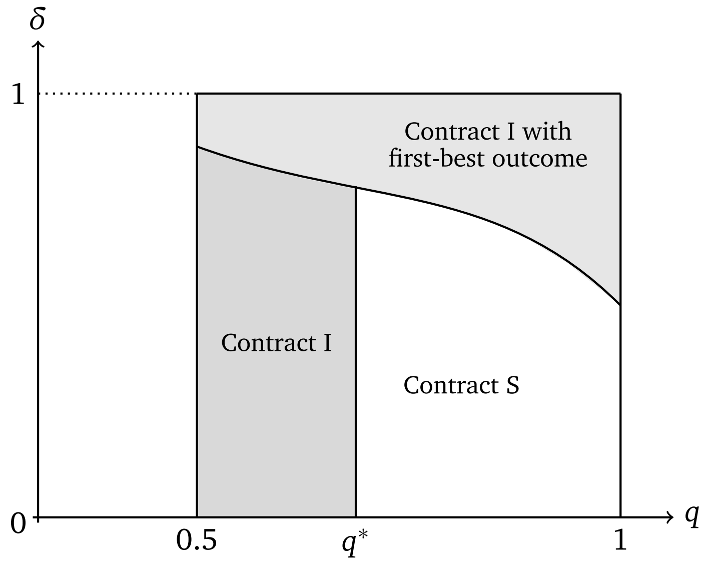

Figure 3: Different kinds of optimal relational contracts

6 Concluding remarks

In this paper, we present a principal-agent model in which it is optimal for the principal to induce the agent to be resistant to public opinion, the phenomenon we call *principal-induced stubbornness* of the agent. Stubbornness is induced by offering the agent more generous payments if his action is good and also defies public opinion. On the other hand, we show that it is never optimal for the principal to induce underutilization of the private signal. We further analyze several variations of the model and show that inducing stubbornness is a robust feature of the optimal contracts.

Our theory may explain why citizens and shareholders support stubborn leaders. As an application, our model explains why voters preferred Bush over Kerry. An alternative explanation of Bush's stubbornness is that he used it to signal his capability. However, this explanation may not apply, as after serving as the U.S. president for one term, voters should possess a fairly accurate sense of his capability.

References

[1] Bardey, David, Denis Gromb, David Martimort, and Jérôme Pouyet. 2020. “Controlling sellers who provide advice: Regulation and competition.” *Journal of Industrial Economics* 68 (3):409–444.

[2] Bénabou, Roland and Jean Tirole. 2002. “Self-confidence and personal motivation.” *Quarterly Journal of Economics* 117 (3):871–915.

[3] Bernardo, Antonio and Ivo Welch. 2001. “On the evolution of overconfidence and entrepreneurs.” *Journal of Economics and Management Strategy* 10 (3):301–330.

[4] Che, Yeon-Koo and Navin Kartik. 2009. “Opinions as incentives.” *Journal of Political Economy* 117 (5):815–860.

[5] Fransen, Marieke, Edith Smit, and Peeter Verlegh. 2015. “Strategies and motives for resistance to persuasion: an integrative framework.” *Frontiers in Psychology* 6:1201.

[6] Häfner, Samuel and Curtis Taylor. 2018. “Advice is cheap: Information is not!” Working paper, available at SSRN (ID 3057479).

[7] Holmström, Bengt. 1980. “On the theory of delegation.” Discussion Paper, No. 438, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL.

[8] Hsu, Charles, Kirill Novoselov, and Rencheng Wang. 2017. “Does accounting conservatism mitigate the shortcomings of CEO overconfidence?” *Accounting Review* 92 (6):77–101.

[9] Levin, Jonathan. 2003. “Relational incentive contracts.” *American Economic Review* 93 (3):835–857.

[10] Levy, Gilat. 2004. “Anti-herding and strategic consultation.” *European Economic Review* 48 (3):503–525.

[11] Li, Hao. 2001. “A theory of conservatism.” *Journal of Political Economy* 109 (3):617–636.

[12] Li, Wei. 2007. “Changing one’s mind when the facts change: incentives of experts and the design of reporting protocols.” *Review of Economic Studies* 74 (4):1175–1194.

[13] Malmendier, Ulrike and Geoffrey Tate. 2005. "CEO overconfidence and corporate investment." *Journal of Finance* 60 (6):2661–2700.

[14] ———. 2008. "Who makes acquisitions? CEO overconfidence and the market's reaction." *Journal of Financial Economics* 89 (1):20–43.

[15] Ottaviani, Marco and Peter Norman Sørensen. 2006. "Reputational cheap talk." *RAND Journal of Economics* 37 (1):155–175.

[16] Owens, Bradley, Michael Johnson, and Terence Mitchell. 2013. "Expressed humility in organizations: Implications for performance, teams, and leadership." *Organization Science* 24 (5):1517–1538.

[17] Prat, Andrea. 2005. "The wrong kind of transparency." *American Economic Review* 95 (3):862–877.

[18] Prendergast, Canice. 1993. "A theory of 'yes men'." *American Economic Review* 83 (4):757–770.

[19] Prendergast, Canice and Lars Stole. 1996. "Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning." *Journal of Political Economy* 104 (6):1105–1134.

[20] Szalay, Dezsö. 2005. "The economics of clear advice and extreme options." *Review of Economic Studies* 72 (4):1173–1198.

[21] Terovitis, Spyros. 2018. "Motivating information acquisition under delegation." Working paper, available at SSRN (ID 3264177).

A Appendix

A1 Proof of Lemma 1

Proof. Consider some fixed Contract I $(\tilde{w}_G, \tilde{w}_B)$. We show that it cannot be optimal when $\tilde{w}_G \leq \tilde{w}_B$.

Case I: $\tilde{w}_G < \tilde{w}_B$ Given $\tilde{w}_G < \tilde{w}_B$, the expert's optimal response is *not* to follow the more informative signal between public opinion and private signal; that is, the expert chooses a_H (or resp., a_L) when the more informative signal suggests that θ_L (or resp., θ_H) is more likely. Given that, the expert's optimal effort level $\tilde{e}$ solves

$$\max_{e \geq 0} \int_{0.5}^{q} (q\tilde{w}_B + (1-q)\tilde{w}_G)\, dH(x;e) + \int_{q}^{1} (x\tilde{w}_B + (1-x)\tilde{w}_G)\, dH(x;e) - c(e).$$

The principal's expected payoff under $(\tilde{w}_G, \tilde{w}_B)$ is

$$\widetilde{U} = \int_{0.5}^{q} [(1-q) - q\tilde{w}_B - (1-q)\tilde{w}_G]\, dH(x;\tilde{e}) + \int_{q}^{1} [(1-x) - x\tilde{w}_B - (1-x)\tilde{w}_G]\, dH(x;\tilde{e}).$$

Consider another Contract I $(\tilde{w}_B, \tilde{w}_G)$ that pays $\tilde{w}_B$ when the action is Good or otherwise pays $\tilde{w}_G$. In this case, the expert will follow the more informative signal. It can be verified that the induced effort level will still be $\tilde{e}$, and the principal's expected payoff will be

$$\int_{0.5}^{q} [q - q\tilde{w}_B - (1-q)\tilde{w}_G]\, dH(x;\tilde{e}) + \int_{q}^{1} [x - x\tilde{w}_B - (1-x)\tilde{w}_G]\, dH(x;\tilde{e}),$$

which is strictly higher than $\widetilde{U}$.

Case II: $\tilde{w}_G = \tilde{w}_B > 0$ In this case, the expert has no incentives to exert efforts and always follows public opinion. The principal benefits by not paying the expert at all. That is, $(\tilde{w}_G, \tilde{w}_B)$ is dominated by $(0, 0)$.

Case III: $\tilde{w}_G = \tilde{w}_B = 0$ Again, the expert does not exert efforts and follows the more informative signal. The principal's payoff is $\widetilde{U}_0 = \int_{0.5}^{q} q\, dH(x; e=0) + \int_{q}^{1} x\, dH(x; e = 0)$.

Consider another Contract I $(\varepsilon, 0)$ that pays $\varepsilon > 0$ for the Good action or 0 otherwise. The expert solves

$$\max_{e \geq 0} \int_{0.5}^{q} q\varepsilon\, dH(x;e) + \int_{q}^{1} x\varepsilon\, dH(x;e) - c(e).$$

and the effort level $\tilde{e}$ is determined by

$$\varepsilon \int_q^1 (F(x) - G(x))\, dx = \frac{c'(\tilde{e})}{p'(\tilde{e})}. \tag{12}$$

The principal's payoff will be

$$U(\varepsilon) = \int_{0.5}^{q} (q - q\varepsilon)\, dH(x; \tilde{e}) + \int_q^1 (x - x\varepsilon)\, dH(x; \tilde{e}).$$

As $\widetilde{U}_0 = U(0)$, it suffices to show that $U'(0) > 0$. Note that $\frac{dU}{d\varepsilon} = \frac{\partial U}{\partial \varepsilon} + \frac{\partial U}{\partial \tilde{e}} \frac{d\tilde{e}}{d\varepsilon}$ where

1. $\frac{\partial U}{\partial \varepsilon} = -\int_{0.5}^{q} q\, dH(x; \tilde{e}) - \int_q^1 x\, dH(x; \tilde{e})$ is finite.
2. by implicit function theorem,

$$\frac{\partial U}{\partial \tilde{e}} \frac{d\tilde{e}}{d\varepsilon} = \Big(\int_q^1 (F(x) - G(x))\, dx \Big)^2 p'(\tilde{e}) \Big/ \Big(\frac{c'(\tilde{e})}{p'(\tilde{e})} \Big)'$$

Therefore, $U'(0) = +\infty$. □

A2 General contracts independent of public opinion

Denote by w_{ij} the payment when the state is i and the action is j.

Claim 1. *In the optimal contract,* $w_{LL} \geq w_{LH}$ *and* $w_{HH} \geq w_{HL}$.

Proof. Assume by contradiction that Claim 1 does not hold. Then either (i) $\tilde{w}_{LL} \geq \tilde{w}_{LH}$ and $\tilde{w}_{HH} < \tilde{w}_{HL}$, (ii) $\tilde{w}_{LL} < \tilde{w}_{LH}$ and $\tilde{w}_{HH} \geq \tilde{w}_{HL}$, or (iii) $\tilde{w}_{LL} < \tilde{w}_{LH}$ and $\tilde{w}_{HH} < \tilde{w}_{HL}$.

For cases (i) and (ii), the expert always chooses action a_L and a_H respectively. And consequently, he exerts no effort ex ante. Therefore, the principal benefits by setting $\tilde{w}'_{LL} = \tilde{w}'_{LH} = \tilde{w}'_{HH} = \tilde{w}'_{HL} = 0$.

For case (iii), the expert always tries to take a Bad action. Consider a corresponding contract $(\tilde{w}'_{LL}, \tilde{w}'_{LH}, \tilde{w}'_{HH}, \tilde{w}'_{HL})$ with $\tilde{w}'_{LL} = \tilde{w}_{LH}$, $\tilde{w}'_{LH} = \tilde{w}_{LL}$, $\tilde{w}'_{HH} = \tilde{w}_{HL}$, and $\tilde{w}'_{HL} = \tilde{w}_{HH}$. Under $(\tilde{w}'_{LL}, \tilde{w}'_{LH}, \tilde{w}'_{HH}, \tilde{w}'_{HL})$, by symmetry, the expert's optimal effort level and the principal's expected payment will be the same; moreover, the expert will always try to take a Good action that matches the state. Therefore, the principal benefits by choosing $(\tilde{w}'_{LL}, \tilde{w}'_{LH}, \tilde{w}'_{HH}, \tilde{w}'_{HL})$. □

Under $w_{LL} \geq w_{LH}$ and $w_{HH} \geq w_{HL}$, the expert follows the more informative signal, and chooses the effort level according to

$$\begin{aligned}\max_{e\geq 0} \frac{1}{2}\Big[&\int_{0.5}^{q} (q(w_{LL}+w_{HH})+(1-q)(w_{LH}+w_{HL}))\,dH(x;e) \\ +&\int_{q}^{1} (x(w_{LL}+w_{HH})+(1-x)(w_{LH}+w_{HL}))\,dH(x;e)\Big]-c(e).\end{aligned} \tag{13}$$

The first-order condition yields:

$$\frac{1}{2}((w_{LL}+w_{HH})-(w_{LH}+w_{HL}))\int_{q}^{1}(F(x)-G(x))dx = \frac{c'(e)}{p'(e)}. \tag{14}$$

Equation (14) implies that only $(w_{LL}+w_{HH})-(w_{LH}+w_{HL})$ matters in determining the expert's effort. As a result, it's optimal for the principal to pay nothing for a Bad action: $w_{LH} = w_{HL} = 0$.

The principal solves the maximization problem:

$$\max_{w_{LL},w_{HH}\in(0,1)} \int_{0.5}^{q}\Big(1-\frac{1}{2}(w_{LL}+w_{HH})\Big)q\,dH(x;e)+\int_{q}^{1}\Big(1-\frac{1}{2}(w_{LL}+w_{HH})\Big)x\,dH(x;e)$$

$$\text{subject to } \frac{1}{2}(w_{LL}+w_{HH})\int_{q}^{1}(F(x)-G(x))dx = \frac{c'(e)}{p'(e)}.$$

Therefore, only $w_{LL}+w_{HH}$ matters in the maximization problem. Comparing this maximization problem with (8) reveals that $w_{LL}+w_{HH} = 2\bar{w}_G$ solves the problem. Consequently, even though there are multiple ways to achieve the maximum expected payoff for the principal, this payoff is the same as under Contract I.

A3 Proof of Lemma 2

Proof. The proof is composed of two parts.

Claim 2. *$\tilde{w}_{GA} \geq \tilde{w}_{BD}$ and $\tilde{w}_{GD} \geq \tilde{w}_{BA}$.*

Assume by contradiction that Claim 2 does not hold. Then either (i) $\tilde{w}_{GA} < \tilde{w}_{BD}$ and $\tilde{w}_{GD} \geq \tilde{w}_{BA}$, (ii) $\tilde{w}_{GA} \geq \tilde{w}_{BD}$ and $\tilde{w}_{GD} < \tilde{w}_{BA}$, or (iii) $\tilde{w}_{GA} < \tilde{w}_{BD}$ and $\tilde{w}_{GD} < \tilde{w}_{BA}$.

For case (i), the expert always gets higher interim expected payoffs by defying public opinion. Knowing that, he will exert no effort and choose the action that defies public opinion. Therefore, the principal benefits by paying nothing in all contingencies.

For case (ii), the expert always gets higher interim expected payoffs by agreeing with public opinion. Knowing that, he will exert no effort and choose the action that approves public opinion. Therefore, the principal benefits by paying nothing in all contingencies.

For case (iii), the expert always tries to take a Bad action that mismatches the state. Consider the corresponding contract $(\tilde{w}'_{GA}, \tilde{w}'_{GD}, \tilde{w}'_{BA}, \tilde{w}'_{BD})$ with $\tilde{w}'_{GA} = \tilde{w}_{BD}$, $\tilde{w}'_{GD} = \tilde{w}_{BA}$, $\tilde{w}'_{BA} = \tilde{w}_{GD}$, and $\tilde{w}'_{BD} = \tilde{w}_{GA}$. Under $(\tilde{w}'_{GA}, \tilde{w}'_{GD}, \tilde{w}'_{BA}, \tilde{w}'_{BD})$, by symmetry, the expert's optimal effort level and the principal's expected payment will be the same; moreover, the expert will always try to take a Good action that matches the state. Therefore, $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ is dominated by the corresponding contract $(\tilde{w}'_{GA}, \tilde{w}'_{GD}, \tilde{w}'_{BA}, \tilde{w}'_{BD})$.

Claim 3. *$\bar{w}_{GA} > \bar{w}_{BD}$ and $\bar{w}_{GD} > \bar{w}_{BA}$.*

Assume by contradiction that Claim 3 does not hold. Then combined with Claim 1, we have either (i) $\bar{w}_{GA} = \bar{w}_{BD}$ and $\bar{w}_{GD} = \bar{w}_{BA}$, (ii) $\bar{w}_{GA} > \bar{w}_{BD}$ and $\bar{w}_{GD} = \bar{w}_{BA}$, or (iii) $\bar{w}_{GA} = \bar{w}_{BD}$ and $\bar{w}_{GD} > \bar{w}_{BA}$.

For case (i), consider a contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ with $\tilde{w}_{GA} = \tilde{w}_{BD} = w_1$ and $\tilde{w}_{GD} = \tilde{w}_{BA} = w_2$. Then the expert will exert no effort and follow public opinion. When either w_1 or w_2 is strictly positive, the principal benefits by paying nothing in all contingencies; when $w_1 = w_2 = 0$, by Lemma 1 the principal benefits by offering $(\varepsilon, \varepsilon, 0, 0)$ for some positive ε.

For case (ii), consider a contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ with $\tilde{w}_{GA} > \tilde{w}_{BD}$ and $\tilde{w}_{GD} = \tilde{w}_{BA}$. Following the analysis in Section 4.3.1, the expert's action rule and optimal effort level are determined by the two differences: $\tilde{w}_{GA} - \tilde{w}_{BD}$ and $\tilde{w}_{GD} - \tilde{w}_{BA}$. As a result, contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ is dominated by $(\tilde{w}'_{GA}, 0, 0, 0)$ with $\tilde{w}'_{GA} = \tilde{w}_{GA} - \tilde{w}_{BD}$. However, under $(\tilde{w}'_{GA}, 0, 0, 0)$ the expert will exert no effort and always follow public opinion. So $(\tilde{w}'_{GA}, 0, 0, 0)$ is dominated by $(0, 0, 0, 0)$.

For case (iii), the argument is similar to that of case (ii). □

A4 Proof of Lemma 3

Proof. It suffices to show that a contract inducing $x^{**} \geq 0.5$ can not be optimal. Otherwise, under such a contract $x^* \leq 0.5$ and

(i) when expert's private signal and public opinion are different, the expert will follow the private signal (since $x \geq x^*$ always holds) and defy public opinion;

(ii) when expert's private signal and public opinion are the same, the expert will follow public opinion when $x \geq x^{**}$ and defy public opinion otherwise.

Equivalently, the expert defies public opinion when $x < x^{**}$ and follows the private signal otherwise. Conditional on $x < x^{**}$, the expert's expected payoff is:

$$qw_{BD} + (1-q)w_{GD}.$$

Conditional on $x \geq x^{**}$, the expert's expected payoff is:

$$xqw_{GA} + x(1-q)w_{GD} + (1-x)qw_{BD} + (1-x)(1-q)w_{BA}.$$

It follows that the expert's ex ante expected payoff is:

$$\int_{x^{**}}^{1} [xqw_{GA} + x(1-q)w_{GD} + (1-x)qw_{BD} + (1-x)(1-q)w_{BA}]\, dH(x;e)$$
$$+ \int_{0.5}^{x^{**}} [qw_{BD} + (1-q)w_{GD}]\, dH(x;e) - c(e).$$

The expert chooses effort e to maximize his expected payoff. The first-order condition yields:

$$[q(w_{GA} - w_{BD}) + (1-q)(w_{GD} - w_{BA})] \int_{x^{**}}^{1} [F(x) - G(x)]dx = \frac{c'(e)}{p'(e)}. \tag{15}$$

For every such contract $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ inducing $x^{**} \in [0.5, 1)$, consider another contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ with

$$\tilde{w}_{GA} = \frac{1-q}{q}(w_{GD} - w_{BA}),$$
$$\tilde{w}_{GD} = \frac{q}{1-q}(w_{GA} - w_{BD}),$$
$$\tilde{w}_{BA} = \tilde{w}_{BD} = 0.$$

We show that $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ is dominated by the contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$.

Claim 4. *Under the constructed contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$,*

(i) $\tilde{x}^* = x^{**} \in [0.5, 1)$, where $\tilde{x}^*$ is the cutoff value under $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ when two signals differ;

(ii) the expert follows public opinion when $x < \tilde{x}^*$ and follows the private signal otherwise;

(iii) the induced effort is the same as that under $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$.

For (i) and (ii):

$$\begin{aligned}\hat{x}^* &= \frac{q(\tilde{w}_{GA}-\tilde{w}_{BD})}{(1-q)(\tilde{w}_{GD}-\tilde{w}_{BA})+q(\tilde{w}_{GA}-\tilde{w}_{BD})} \\ &= \frac{(1-q)(w_{GD}-w_{BA})}{q(w_{GA}-w_{BD})+(1-q)(w_{GD}-w_{BA})} = x^{**}.\end{aligned}$$

For (iii), the expert's incentive-compatible constraint of effort level under $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ is equation (IC), which yields:

$$[(1-q)(w_{GD}-w_{BA})+q(w_{GA}-w_{BD})]\int_{x^{**}}^{1}(F(x)-G(x))\,dx = \frac{c'(e)}{p'(e)}.$$

This coincides with the FOC of the expert's problem under $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ as in equation (15).

Claim 4 implies that under $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$, the probability of the induced action being Good is higher than that under $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$, for the following reasons:

1. the two contracts induce the same level of effort and hence the same distribution of private signals;

2. the two contracts induce the same cutoff ($\tilde{x}^*$ and x^{**}), and the induced action will be the same when x is higher than the cutoff. However, when x is lower than the cutoff, the contract $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$ induces the expert to follow public opinion whereas the contract $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ induces the expert to defy public opinion.

On the other hand, the principal's expected payments under $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$

is weakly lower than that under $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$:

$$\begin{aligned}
&\int_{0.5}^{\tilde{x}^*} \big(q\tilde{w}_{GA} + (1-q)\tilde{w}_{BA}\big) dH(x;e) \\
&+ \int_{\tilde{x}^*}^{1} \Big[x\big(q\tilde{w}_{GA} + (1-q)\tilde{w}_{GD}\big) + (1-x)\big(q\tilde{w}_{BD} + (1-q)\tilde{w}_{BA}\big)\Big] dH(x;e) \\
= &\int_{0.5}^{x^{**}} (1-q)(w_{GD} - w_{BA}) dH(x;e) \\
&+ \int_{x^{**}}^{1} \Big[x\big((1-q)(w_{GD} - w_{BA}) + q(w_{GA} - w_{BD})\big)\Big] dH(x;e) \\
\leq &\int_{0.5}^{x^{**}} [(1-q)w_{GD} + qw_{BD}] dH(x;e) \\
&+ \int_{x^{**}}^{1} [xqw_{GA} + x(1-q)w_{GD} + (1-x)qw_{BD} + (1-x)(1-q)w_{BA}] dH(x;e)
\end{aligned}$$

Therefore, the principal's expected profits under $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ is strictly lower than that under $(\tilde{w}_{GA}, \tilde{w}_{GD}, \tilde{w}_{BA}, \tilde{w}_{BD})$. □

A5 Proof of Proposition 2

Proof. An optimal contract is either a Contract I ($w_{GA} = w_{GD}$), Contract F ($w_{GA} > w_{GD}$) or Contract S ($w_{GA} < w_{GD}$). We first show that every Contract F is dominated by some Contract I. Then we show that the optimal Contract I is dominated by some Contract S, which concludes the proof of Proposition 2 that the optimal contract is a Contract S.

Claim 5. *For any Contract F, there exists a Contract I that gives the principal a strictly higher payoff.*

Note that for a Contract F with $qw_{GA} + (1-q)w_{GD} > 1$, the principal pays $w_{GA} > qw_{GA} + (1-q)w_{GD} > 1$ when the expert follows public opinion and $qw_{GA} + (1-q)w_{GD} > 1$ when the expert follows the private signal. Since the return from the Good action is only 1, the principal's payoff would be negative. Therefore, such a Contract F is clearly dominated by the optimal Contract I.

Now, we focus on Contract F with $qw_{GA} + (1-q)w_{GD} \leq 1$. For any such Contract F, consider a Contract I with $\tilde{w}_G = qw_{GA} + (1-q)w_{GD}$ and $\tilde{w}_B = 0$. Let the cutoffs of the Contract F and the Contract I be x^F and x^I respectively. We have

$$x^F = \frac{qw_{GA}}{w_G} > q \text{ and } x^I = q.$$

Expert's effort levels under the two contracts, e^F and e^I, are determined by:

$$w_G \int_{x^F}^{1} (F(x) - G(x))dx = \frac{c'(e^F)}{p'(e^F)}, \text{ and}$$
$$w_G \int_{q}^{1} (F(x) - G(x))dx = \frac{c'(e^I)}{p'(e^I)}.$$

Since $x^F > q$ and $F(x) > G(x)$ for all $x \in (0.5, 1)$, we have $\int_{x^F}^1 (F(x) - G(x))dx < \int_q^1 (F(x) - G(x))dx$. Then since $\frac{c'(e)}{p'(e)}$ is strictly increasing, we have $e^F < e^I$. The principal's ex ante expected payoff under Contract F and Contract I, Π^F and Π^I respectively, are

$$\Pi^F = \int_{0.5}^{x^F} [q(1 - w_{GA})]dH(x; e^F) + \int_{x^F}^{1} [x(1 - w_G)]dH(x; e^F)$$
$$\Pi^I = \int_{0.5}^{q} [q(1 - w_G)]dH(x; e^I) + \int_{q}^{1} [x(1 - w_G)]dH(x; e^I)]$$

Rearranging the terms and using integration by parts,

$$\Pi^I - \Pi^F = \int_{0.5}^{x^F} [q(w_{GA} - w_G)]dH(x; e^F) + \int_{q}^{x^F} [(x - q)(1 - w_G)]dH(x; e^F)$$
$$+ (p(e^I) - p(e^F))(1 - w_G) \int_{q}^{1} (F(x) - G(x))dx.$$

$\Pi^I - \Pi^F$ can be decomposed into three parts.

(i) The first term $\int_{0.5}^{x^F} [q(w_{GA} - w_G)]dH(x; e^F)$ is positive. The amount $w_{GA} - w_G$ is saved for the principal when the Good Action is Approved conditional on $x \in [0.5, x^F]$.

(ii) The second term $\int_q^{x^F} [(x-q)(1-w_G)]dH(x; e^F)$ is non-negative since $w_G \leq 1$. This term measures principal's gain through the expert's more efficient action rule under Contract I. Specifically, conditional on $x \in [q, x^F]$, the expert follows the more informative private signal under Contract I while follows the less informative public opinion under Contract F.

(iii) The third term is non-negative since $p(e^I) > p(e^F)$, $w_G \leq 1$ and $F(x) > G(x)$ for all $x \in (0.5, 1)$. This term measures principal's gain through the expert's higher effort level under Contract I.

Therefore, $\Pi^I - \Pi^F > 0$.

Next, we show that the optimal Contract I is dominated by some Contract S.

Claim 6. *There exists a Contract S that gives the principal a strictly higher payoff than that of the optimal Contract I.*

A special type of Contract S is constructed as follows:

$$w_{GA} = \bar{w}_G - \varepsilon \text{ and } w_{GD} = \bar{w}_G + \frac{q}{1-q}\varepsilon \text{ for some } \varepsilon > 0.$$

As long as $\varepsilon < \frac{2q-1}{2q}\bar{w}_G$, the constructed Contract S satisfies the constraints of the maximization problem ($\mathcal{M}$). We prove the claim by showing that at $\varepsilon = 0$, a marginal increase of ε results in an increase in the principal's ex ante expected payoff. The principal's ex ante expected payoff is

$$\begin{aligned} &\int_{0.5}^{x^*} [q - qw_{GA}]dH(x;e) + \int_{x^*}^{1} [x - (xqw_{GA} + x(1-q)w_{GD})]dH(x;e) \\ =&\int_{0.5}^{x^*} [q(1-\bar{w}_G+\varepsilon)]dH(x;e) + \int_{x^*}^{1} [x(1-\bar{w}_G)]dH(x;e) \equiv \Pi(\varepsilon), \end{aligned}$$

where x^* and e are determined by the constraints. ε affects $\Pi(\varepsilon)$ directly, and also indirectly through x^* and e. The derivative of $\Pi(\varepsilon)$ could be derived as follows:

$$\frac{d\Pi(\varepsilon)}{d\varepsilon} = \frac{\partial\Pi(\varepsilon)}{\partial\varepsilon} + \frac{\partial\Pi(\varepsilon)}{\partial x^*}\frac{dx^*}{d\varepsilon} + \frac{\partial\Pi(\varepsilon)}{\partial e}\frac{de}{d\varepsilon}.$$

The direct effect $\frac{\partial\Pi(\varepsilon)}{\partial\varepsilon}$ is positive:

$$\frac{\partial\Pi(\varepsilon)}{\partial\varepsilon} = \int_{0.5}^{x^*} qdH(x;e) > 0.$$

Next, we analyze the effect though x^*. The cutoff x^* is:

$$x^* = \frac{qw_{GA}}{(1-q)w_{GD} + qw_{GA}} = \frac{q(\bar{w}_G - \varepsilon)}{\bar{w}_G}.$$

We have $\lim_{\varepsilon\to 0} x^* = q$. Differentiate x^* with respect to ε gives:

$$\frac{dx^*}{d\varepsilon} = -\frac{q}{\bar{w}_G} < 0.$$

The partial derivative $\frac{\partial \Pi(\varepsilon)}{\partial x^*}$ is:

$$\begin{aligned}\frac{\partial \Pi(\varepsilon)}{\partial x^*} &= [q(1-\bar{w}_G+\varepsilon)-x^*(1-\bar{w}_G)][p(e)g(x^*)+(1-p(e))f(x^*)] \\ &= \frac{q\varepsilon}{\bar{w}_G}[p(e)g(x^*)+(1-p(e))f(x^*)]\end{aligned}$$

Notice that

$$\lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial x^*} = 0. \tag{16}$$

That is, shifting x^* downward at $x^* = q$ only leads to a second-order loss.

Lastly, we analyze the effect through e. The expert chooses e according to:

$$\begin{aligned} &(qw_{GA}+(1-q)w_{GD})\int_{x^*}^{1}(F(x)-G(x))dx = \frac{c'(e)}{p'(e)} \\ \Longrightarrow\ &\bar{w}_G\int_{x^*}^{1}(F(x)-G(x))dx = \frac{c'(e)}{p'(e)}. \end{aligned} \tag{17}$$

Let

$$D := \bar{w}_G\int_{x^*}^{1}(F(x)-G(x))dx > 0,$$

then,

$$\frac{dD}{d\varepsilon} = -\bar{w}_G\frac{dx^*}{d\varepsilon}\left(F(x^*)-G(x^*)\right) = q\left(F(x^*)-G(x^*)\right) > 0.$$

Therefore,

$$\frac{de}{d\varepsilon} = \frac{p'(e)\frac{dD}{d\varepsilon}}{c''(e)-p''(e)D} > 0.$$

The partial derivative $\frac{\partial \Pi(\varepsilon)}{\partial e}$ is:

$$\frac{\partial \Pi(\varepsilon)}{\partial e} = p'(e)\left(-\frac{q\varepsilon}{\bar{w}_G}(F(x^*)-G(x^*))+(1-\bar{w}_G)\int_{x^*}^{1}(F(x)-G(x))dx\right)$$

Then,

$$\lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial e} = p'(e)(1-\bar{w}_G)\int_{q}^{1}(F(x)-G(x))dx > 0.$$

Finally,

$$\begin{aligned}\lim_{\varepsilon\to 0}\frac{d\Pi(\varepsilon)}{d\varepsilon} &= \lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial \varepsilon}+\lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial x^*}\lim_{\varepsilon\to 0}\frac{dx^*}{d\varepsilon}+\lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial e}\lim_{\varepsilon\to 0}\frac{de}{d\varepsilon} \\ &= \int_{0.5}^{q} q\,dH(x;e)+0+\lim_{\varepsilon\to 0}\frac{\partial \Pi(\varepsilon)}{\partial e}\lim_{\varepsilon\to 0}\frac{de}{d\varepsilon} > 0\end{aligned}$$

which concludes the proof. □

A6 Non-symmetric contracts

Denote a general contract by $(w^L_{GA}, w^L_{GD}, w^L_{BA}, w^L_{BD}, w^H_{GA}, w^H_{GD}, w^H_{BA}, w^H_{BD})$, where the superscript L (H) indicates public opinion being σ_L (σ_H). The previous result that only Good actions should be rewarded still applies; that is

$$w^L_{GA} \geq 0, w^L_{GD} \geq 0, w^H_{GA} \geq 0, w^H_{GD} \geq 0 \tag{LL-g}$$

and

$$w^L_{BA} = w^L_{BD} = w^H_{BA} = w^H_{BD} = 0.$$

Since the expert's action decision is made after public opinion is realized, the general contract would induce two action rules based on whether public opinion is L or H:

$$x^*_L = \begin{cases} q & \text{if } w^L_{GA} = w^L_{GD} = 0; \\ \frac{qw^L_{GA}}{qw^L_{GA}+(1-q)w^L_{GD}} & \text{otherwise.} \end{cases} \tag{x^*_L}$$

$$x^*_H = \begin{cases} q & \text{if } w^H_{GA} = w^H_{GD} = 0; \\ \frac{qw^H_{GA}}{qw^H_{GA}+(1-q)w^H_{GD}} & \text{otherwise.} \end{cases} \tag{x^*_H}$$

As a result, the expert chooses his effort level to solve:

$$\begin{aligned} \max_{e\geq 0} \frac{1}{2}\Big[\int_{x^*_L}^{1} x(qw^L_{GA} + (1-q)w^L_{GD})\, dH(x;e) + \int_{0.5}^{x^*_L} qw^L_{GA}\, dH(x;e)\Big] \\ + \frac{1}{2}\Big[\int_{x^*_H}^{1} x(qw^H_{GA} + (1-q)w^H_{GD})\, dH(x;e) + \int_{0.5}^{x^*_H} qw^H_{GA}\, dH(x;e)\Big] - c(e) \end{aligned} \tag{18}$$

Using integration by parts, the FOC of the maximization problem (18) reduces to

$$\begin{aligned} \frac{1}{2}\Big[(qw^L_{GA} + (1-q)w^L_{GD}) \int_{x^*_L}^{1} (F(x) - G(x))\, dx + \\ (qw^H_{GA} + (1-q)w^H_{GD}) \int_{x^*_H}^{1} F(x) - G(x)\, dx\Big] = \frac{c'(e)}{p'(e)} \end{aligned} \tag{IC-g}$$

Denote the effort level that solves (IC-g) by e^g.

The principal solves the following maximization problem:

$$\begin{aligned} \max_{w^L_{GA}, w^L_{GD}, w^H_{GA}, w^H_{GD}} \frac{1}{2}\Big[\int_{x^*_L}^{1} x(1 - qw^L_{GA} - (1-q)w^L_{GD})\, dH(x;e) + \int_{0.5}^{x^*_L} q(1 - w^L_{GA})\, dH(x;e)\Big] \\ + \frac{1}{2}\Big[\int_{x^*_H}^{1} x(1 - qw^H_{GA} - (1-q)w^H_{GD})\, dH(x;e) + \int_{0.5}^{x^*_H} q(1 - w^H_{GA})\, dH(x;e)\Big] \end{aligned}$$

subject to (x^*_L), (x^*_H), (IC-g), (LL-g). (19)

Denote by $\Pi(w_{GA}^L, w_{GD}^L, w_{GA}^H, w_{GD}^H)$ the principal's expected payoffs. We first show that there exists some non-symmetric contract that dominates the optimal symmetric contract.

Claim 7. *The optimal contract is non-symmetric.*

Proof. Denote by $(\bar{w}_{GA}, \bar{w}_{GD}, 0, 0)$ the optimal symmetric contract. We show that there exists a general contract $(\tilde{w}_{GA}^L, \tilde{w}_{GD}^L, 0, 0, \tilde{w}_{GA}^H, \tilde{w}_{GD}^H, 0, 0)$ with

$$\tilde{w}_{GA}^L = 2\bar{w}_{GA},\ \tilde{w}_{GD}^L = 2\bar{w}_{GD},\ \tilde{w}_{GA}^H = \tilde{w}_{GD}^H = 0$$

yielding strictly higher expected payoff than that of the optimal symmetric contract for the principal.

Under this constructed contract, upon observing σ_L, the expert's action rule is determined by the cutoff $x_L^* = x^*$; upon observing σ_H, the expert's action rule is determined by cutoff q. Overall, the expert's effort level under the contract is determined by:

$$\begin{aligned}
&\frac{1}{2}\Big[(q\tilde{w}_{GA}^L + (1-q)\tilde{w}_{GD}^L)\int_{x^*}^1 F(x) - G(x)\,dx\Big] = \frac{c'(e)}{p'(e)} \\
\Longleftrightarrow\ &(q\tilde{w}_{GA}^* + (1-q)\tilde{w}_{GD}^*)\int_{x^*}^1 F(x) - G(x)\,dx = \frac{c'(e)}{p'(e)},
\end{aligned}$$

which is the same as the effort level induced by the symmetric contract $(\bar{w}_{GA}, \bar{w}_{GD}, 0, 0)$.

The principal's expected payoff under the constructed contract is:

$$\begin{aligned}
&\frac{1}{2}\Big[\int_{x^*}^1 x(1 - 2q\bar{w}_{GA} - 2(1-q)\bar{w}_{GD})\,dH(x;e) + \int_{0.5}^{x^*} q(1 - 2\bar{w}_{GA})\,dH(x;e)\Big] \\
&+ \frac{1}{2}\Big[\int_q^1 x\,dH(x;e) + \int_{0.5}^q q\,dH(x;e)\Big] \\
= &\int_{x^*}^1 x(1 - q\bar{w}_{GA} - (1-q)\bar{w}_{GD})\,dH(x;e) + \int_{0.5}^{x^*} q(1 - \bar{w}_{GA})\,dH(x;e) \\
&+ \frac{1}{2}\int_{x^*}^q (q - x)\,dH(x;e),
\end{aligned}$$

which is higher than the expected payoff under the optimal symmetric contract by $\frac{1}{2}\int_{x^*}^q (q-x)\,dH(x;e) > 0$. This concludes the proof. □

Claim 8. *The optimal contract induces stubbornness.*

Proof. Assume $x_L^* \leq x_H^*$ (the symmetric case where $x_L^* \leq x_H^*$ can be similarly analyzed). We first prove that $x_H^* \leq q$. Suppose not, that is, an optimal contract $(w_{GA}^L, w_{GD}^L, w_{GA}^H, w_{GD}^H)$ induces $x_H^* > q$, which is equivalent to $w_{GA}^H > w_{GD}^H$. We will construct a new contract that induces $\tilde{x}_H^* = q$ while keeping x_L^* and e^g, and show that this contract yields a strictly higher expected payoff for the principal. The new contract $(\tilde{w}_{GA}^L, \tilde{w}_{GD}^L, \tilde{w}_{GA}^H, \tilde{w}_{GD}^H)$ specifies:

$$\tilde{w}_{GA}^L = w_{GA}^L,\ \tilde{w}_{GD}^L = w_{GD}^L,\ \tilde{w}_{GA}^H = \tilde{w}_{GD}^H =: \tilde{w}_G^H,$$

where $\tilde{w}_G^H$ is set to keep the effort level at e^g. That is, $\tilde{w}_G^H$ solves

$$\frac{1}{2}\Big[(qw_{GA}^L+(1-q)w_{GD}^L)\int_{x_L^*}^1 F(x)-G(x)\,dx+\tilde{w}_G^H\int_q^1 F(x)-G(x)\,dx\Big]=\frac{c'(e^g)}{p'(e^g)}. \quad \text{(IC-n)}$$

Comparing (IC-g) and (IC-n), we have

$$\tilde{w}_G^H < qw_{GA}^H + (1-q)w_{GD}^H. \tag{20}$$

Since $w_{GA}^H > w_{GD}^H$, this further implies $\tilde{w}_G^H < w_{GA}^H$.

The principal's expected payoff is strictly higher under the new contract:

$$\begin{aligned}
&\Pi(\tilde{w}_{GA}^L, \tilde{w}_{GD}^L, \tilde{w}_{GA}^H, \tilde{w}_{GD}^H) - \Pi(w_{GA}^L, w_{GD}^L, w_{GA}^H, w_{GD}^H)\\
=&\frac{1}{2}\Big[\int_q^1 x(1-\tilde{w}_G^H)\,dH(x;e) + \int_{0.5}^q q(1-\tilde{w}_G^H)\,dH(x;e)\Big]\\
&-\frac{1}{2}\Big[\int_{x_H^*}^1 x(1-qw_{GA}^H-(1-q)w_{GD}^H)\,dH(x;e) + \int_{0.5}^{x_H^*} q(1-w_{GA}^H)\,dH(x;e)\Big]\\
=&\frac{1}{2}\Big[\int_{0.5}^q q(w_{GA}^H-\tilde{w}_G^H)\,dH(x;e) + \int_q^{x_H^*}[(x-q)+q(w_{GA}^H-\tilde{w}_G^H)]\,dH(x;e)\\
&+\int_{x_H^*}^1 x(qw_{GA}^H+(1-q)w_{GD}^H-\tilde{w}_G^H))\,dH(x;e)\Big] > 0.
\end{aligned}$$

This concludes the proof of the claim. And thus, $x_L^* \leq x_H^* \leq q$.

1. If $x_H^* < q$, then both x_L^* and x_H^* are distorted downwards from q and stubbornness is induced in both cases.

2. If $x_H^* = q$, we already know that contract I that specifies $x_L^* = x_H^* = q$ is not optimal. In this case, we must have $x_L^* < x_H^* \leq q$, and stubbornness is induced when the realized public opinion is L.

Therefore, stubbornness must be induced in an optimal contract. □

A7 Proof of Lemma 5

Proof. First, note that the principal's expected payoffs under $(\boldsymbol{w}^*, \hat{a}^*)$ must be weakly higher than that under the symmetric contract of the main model. So under the optimal incentive-compatible mechanism both actions a_L and a_H can be induced; we have either $w_{GA} \neq w_{BD}$ or $w_{GA} \neq w_{BD}$, or both.

We first show that an incentive-compatible action rule takes the following form: given the reported private signal $\tilde{s}$ and public opinion σ, the principal follows the public signal or private signal when the reported $\tilde{x}$ is above some cutoff and otherwise follows the other signal. When the cutoff is less or equal to 0.5, the principal will always follow either public opinion or the reported private signal.

Claim 9. *In the optimal incentive-compatible mechanism,*

$$\hat{a}(\tilde{x}, \tilde{s}_i, \sigma_j) = \begin{cases} a_\ell & \text{if } i \neq j, \tilde{x} > \tilde{x}^* \\ a_{\ell'} & \text{if } i \neq j, \tilde{x} < \tilde{x}^* \\ a_k & \text{if } i = j, \tilde{x} > \tilde{x}^{**} \\ a_{k'} & \text{if } i = j, \tilde{x} < \tilde{x}^{**} \end{cases}$$

for $i, j, \ell, k, \ell', k' \in \{L, H\}$, $\ell \neq \ell'$, $k \neq k'$, $x^* = \frac{q(w_{GA}-w_{BD})}{(1-q)(w_{GD}-w_{BA})+q(w_{GA}-w_{BD})}$ *and* $x^{**} = \frac{(1-q)(w_{GD}-w_{BA})}{(1-q)(w_{GD}-w_{BA})+q(w_{GA}-w_{BD})}$.

Let $(\boldsymbol{w}, \hat{a})$ be an incentive-compatible mechanism with $\boldsymbol{w} = (w_{GA}, w_{GD}, w_{BA}, w_{BD})$. Suppose the observed public signal differs from private signal, say (s_L, σ_H). Assume that both actions can be induced under the action rule $\hat{a}$. The difference of expert's expected payments between recommending action a_L and a_H is $\hat{\rho}(x, s_L, \sigma_H)(w_{GD} - w_{BA}) - (1 - \hat{\rho}(x, s_L, \sigma_H))(w_{GA} - w_{BD})$. Recall that $\hat{\rho}(x, s_L, \sigma_H)$ is strictly monotonic in x. So it's in the expert's interests to recommend the principal to follow some cutoff-form action rule, where the cutoff x^* is determined by $x^* = \frac{q(w_{GA}-w_{BD})}{(1-q)(w_{GD}-w_{BA})+q(w_{GA}-w_{BD})}$. Similarly, when the expert obverses (s_H, σ_L), he will still follow public opinion or private signal when $x > x^*$ and will follow the other signal otherwise.

When the private signal agrees with public opinion, it follows by similar arguments that the incentive-compatible action rule takes the cutoff-form with the cutoff being $\tilde{x}^{**} = \frac{(1-q)(w_{GD}-w_{BA})}{(1-q)(w_{GD}-w_{BA})+q(w_{GA}-w_{BD})}$.

Both $\tilde{x}^*$ and $\tilde{x}^{**}$ lie in $(0, 1)$ and $x^* + x^{**} = 1$, and only one of the two constraints $x \geq x^*$ and $x \geq x^{**}$ can be valid. Indeed, only $\tilde{x}^*$ is a valid cutoff and $\tilde{x}^{**} < 0.5$.

Claim 10. *In the optimal incentive-compatible mechanism,* $x^* \in (0.5, 1)$.

The proof of Claim 10 is similar to that of Lemma 3 in the main model. □

A8 Proof of Proposition 5

Proof. The optimal Contract I solves

$$\begin{aligned} &\max_{w_G} \int_{0.5}^{q} (q - q w_G) dH(x;e) + \int_{q}^{1} (x - x w_G) dH(x;e) \\ &\text{subject to } (w_G + b) \int_{q}^{1} (F(x) - G(x)) dx = \frac{c'(e)}{p'(e)}. \end{aligned} \tag{21}$$

Denote the solution by $\bar{w}_G$. $\bar{w}_G < 1$ must still hold since otherwise the principal's expected payoff would be non-positive.

Firstly, we show that every Contract F ($w_{GA} > w_{GD}$) is dominated by some Contract I.

Claim 11. *For any Contract F, there exists a Contract I that gives the principal a strictly higher payoff.*

Similar to the proof for Proposition 2, a Contract F with $q w_{GA} + (1-q) w_{GD} > 1$ yields negative expected payoffs for the principal and is thus clearly dominated by the optimal Contract I.

We focus on Contract F with $q w_{GA} + (1-q) w_{GD} \leq 1$. For any such Contract F, consider a Contract I with $w_G = q w_{GA} + (1-q) w_{GD} \leq 1$. The cutoff of Contract F is modified to be $\hat{x}^F = \frac{q(w_{GA}+b)}{w_G+b}$, which still has the property $\hat{x}^F > q$. The cutoff of Contract I is still q.

The expert's effort choices under the two contracts be $\hat{e}^F$ and $\hat{e}^I$ are determined by:

$$\begin{aligned} (w_G + b) \int_{\hat{x}^F}^{1} (F(x) - G(x)) dx &= \frac{c'(\hat{e}^F)}{p'(\hat{e}^F)} \\ \text{and } (w_G + b) \int_{q}^{1} (F(x) - G(x)) dx &= \frac{c'(\hat{e}^I)}{p'(\hat{e}^I)}. \end{aligned}$$

Still, $\hat{e}^F < \hat{e}^I$ because $\hat{x}^F > q$ and $F(x) > G(x)$ for all $x \in (0.5, 1)$.

Let Π^F and Π^I be the principal's ex ante expected payoff under Contract F and Contract I respectively. Then the difference is the same as before after taking into account the modified $\hat{x}^F$, $\hat{e}^F$ and $\hat{e}^I$:

$$\begin{aligned} \Pi^I - \Pi^F = &\int_{0.5}^{\hat{x}^F} q(w_{GA} - w_G) dH(x;\hat{e}^F) + \int_{q}^{x^F} (x-q)(1-w_G) dH(x;\hat{e}^F) \\ &+ (p(\hat{e}^I) - p(\hat{e}^F))(1-w_G) \int_{q}^{1} (F(x) - G(x)) dx. \end{aligned}$$

The first term is still positive and the last two terms are still non-negative since $\hat{x}^F > q$, $\hat{e}^F < \hat{e}^I$ and $w_G \leq 1$. Therefore, $\Pi^I - \Pi^F > 0$ and Claim 11 is proved.

Next, we show that the optimal Contract I is dominated by some Contract S ($w_{GD} > w_{GA}$) when $\bar{w}_G > 0$.

Claim 12. *There exists a Contract S that gives the principal a strictly higher payoff than the optimal Contract I if $\bar{w}_G > 0$.*

Similar to the proof of Proposition 2, we construct a special type of Contract S:

$$w_{GA} = \bar{w}_G - \varepsilon \text{ and } w_{GD} = \bar{w}_G + \frac{q}{1-q}\varepsilon \text{ for some } \varepsilon > 0.$$

As long as $\varepsilon < \min\{\frac{2q-1}{2q}(\bar{w}_G + b), \bar{w}_G\}$, the constructed Contract S satisfies the constraints of the maximization problem (11). Still, we prove the claim by showing that at $\varepsilon = 0$, a marginal increase of ε results in an increase in the principal's ex ante expected payoff. The principal's ex ante expected payoff is

$$\int_{0.5}^{x^*} q(1 - \bar{w}_G + \varepsilon) dH(x; e) + \int_{x^*}^{1} x(1 - \bar{w}_G) dH(x; e) \equiv \Pi(\varepsilon),$$

where x^* and e are determined by the constraints. The derivative of $\Pi(\varepsilon)$ could be derived as follows:

$$\frac{d\Pi(\varepsilon)}{d\varepsilon} = \frac{\partial \Pi(\varepsilon)}{\partial \varepsilon} + \frac{\partial \Pi(\varepsilon)}{\partial x^*}\frac{dx^*}{d\varepsilon} + \frac{\partial \Pi(\varepsilon)}{\partial e}\frac{de}{d\varepsilon}.$$

The direct effect $\frac{\partial \Pi(\varepsilon)}{\partial \varepsilon}$ is positive:

$$\frac{\partial \Pi(\varepsilon)}{\partial \varepsilon} = \int_{0.5}^{x^*} q dH(x; e) > 0.$$

Next, we analyze the effect though x^*. The cutoff x^* is:

$$x^* = \frac{q(w_{GA} + b)}{(1-q)w_{GD} + q w_{GA} + b} = \frac{q(\bar{w}_G + b - \varepsilon)}{\bar{w}_G + b}.$$

We have $\lim_{\varepsilon \to 0} x^* = q$. Differentiate x^* with respect to ε gives:

$$\frac{dx^*}{d\varepsilon} = -\frac{q}{(\bar{w}_G + b)} < 0.$$

The partial derivative $\frac{\partial \Pi(\varepsilon)}{\partial x^*}$ is:

$$\frac{\partial \Pi(\varepsilon)}{\partial x^*} = \frac{q\varepsilon(1+b)}{(\bar{w}_G + b)}[p(e)g(x^*) + (1 - p(e))f(x^*)]$$

Notice that

$$\lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial x^*} = 0. \tag{22}$$

Lastly, we analyze the effect through e. The expert chooses e according to:

$$(\bar{w}_G + b) \int_{x^*}^{1} (F(x) - G(x)) dx = \frac{c'(e)}{p'(e)}. \tag{23}$$

Therefore, $\frac{de}{d\varepsilon} > 0$ because $\frac{dx^*}{d\varepsilon} < 0$, $p''(e) < 0$, $c''(e) > 0$ and $F(x) > G(x)$ for all $x \in (0.5, 1)$.

The partial derivative $\frac{\partial \Pi(\varepsilon)}{\partial e}$ is:

$$\frac{\partial \Pi(\varepsilon)}{\partial e} = p'(e) \left(-\frac{q\varepsilon}{\bar{w}_G}(F(x^*) - G(x^*)) + (1 - \bar{w}_G) \int_{x^*}^{1} (F(x) - G(x)) dx \right)$$

Then,

$$\lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial e} = p'(e)(1 - \bar{w}_G) \int_{q}^{1} (F(x) - G(x)) dx > 0.$$

Finally:

$$\begin{aligned} \lim_{\varepsilon \to 0} \frac{d\Pi(\varepsilon)}{d\varepsilon} &= \lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial \varepsilon} + \lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial x^*} \lim_{\varepsilon \to 0} \frac{dx^*}{d\varepsilon} + \lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial e} \lim_{\varepsilon \to 0} \frac{de}{d\varepsilon} \\ &= \int_{0.5}^{q} q dH(x; e) + 0 + \lim_{\varepsilon \to 0} \frac{\partial \Pi(\varepsilon)}{\partial e} \lim_{\varepsilon \to 0} \frac{de}{d\varepsilon} > 0. \end{aligned}$$

Combining the results of Claim 11 and Claim 12, when $\bar{w}_G > 0$, the optimal contract belongs to Contract S.

Lastly, we consider the case $\bar{w}_G = 0$. By Claim 11, any Contract F must be dominated by the optimal Contract I. Excluding Contract F from possible optimal contracts, we are left with Contract I and Contract S. Therefore, the optimal contract either pays nothing, or it induces stubbornness. □