"Principal-induced stubbornness in experts"
Online Appendix

Yuk-fai Fong, Xiaoxiao Hu, Xiaoxuan Meng, and Kirill Novoselov

1 Relational contract model

There are two players: a principal (she) and an expert (he). Both players are risk-neutral and infinitely lived. Periods are discrete and indexed by $t \in \{1, 2, ..., \infty\}$. Players discount payoffs from the following period by a factor of $\delta \in (0, 1)$. In each period, the principal relies on the expert to both study the state of nature and then take an action to match the state of nature. The state of nature θ_t can take two possible values, θ_L and θ_H. The state of each period is independently drawn and the *ex-ante* probability of each possible state is $1/2$. Both players share a common prior regarding the state.

Expert's private signal The expert can exert some effort level $e_t \geq 0$ to generate an informative, yet imperfect, signal of the state s_t whose distribution depends on both the state and the expert's effort level. Specifically, signal s_t takes values s_L and s_H with

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

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

$$H(x_t; e_t) = (1 - p(e_t))F(x_t) + p(e_t)G(x_t). \tag{1}$$

Here $H(\cdot\,; e_t)$ 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_t, the realized signal s_t, and the signal precision x_t are all privately observed by the expert.

Throughout this paper, we impose the following assumptions:

(i) The PDFs f and g satisfy the monotone likelihood ratio property (MLRP), i.e., $g(x)/f(x)$ increases on $[0.5, 1]$. Note that MLRP further implies that G first-order stochastically dominates F, i.e., $G(x) < F(x)$ *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 $\lim_{e\to 0} p'(e)/\left(\frac{c'(e)}{p'(e)}\right)'=+\infty$.[1]

Public opinion After the expert privately observes the realized signal, Nature generates a public signal σ_t 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_t as *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:

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

Compensation package At the beginning of each period t, the principal proposes a compensation package to the expert. Compensation consists of a base wage w_t and a discretionary payment, which we refer to as a performance-based bonus, b_t. The base wage can be contingent on all publicly observable information until period t. The bonus can be contingent on all publicly observable information in period t: the expert's chosen action a_t, the realized state θ_t and the realized public opinion σ_t. The bonus function is characterized by $\hat{b}_t:\{a_L,a_H\}\times\{\theta_L,\theta_H\}\times\{\sigma_L,\sigma_H\}\to\mathbb{R}$.

At the end of period t, the principal is only obligated to pay w_t. The players decide whether to pay $b_t(a_t,\theta_t,\sigma_t)$: if $b_t>0$, the principal makes the decision; if $b_t<0$, the expert makes the decision. Let the total payment from the principal to the expert in the period be W_t.

Payoffs The principal's ex-post payoff in period t is

$$u_t^P=r(a_t,\theta_t)-W_t \tag{2}$$

and the expert's ex-post payoff is

$$u_t^E=W_t-c(e_t). \tag{3}$$

The values of both players' outside options are assumed to be 0.

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

Timing The timing of the period-t game is as follows.

1. The principal offers a compensation package to the expert.
2. If accepts the compensation package, the expert exerts effort level $e_t \geq 0$ and observes the realized signal precision $x_t \in (0.5, 1)$ and the realized private signal $s_t \in \{s_L, s_H\}$.
3. Both players observe the realized public opinion, $\sigma_t \in \{\sigma_L, \sigma_H\}$.
4. The expert chooses an action $a_t \in \{a_L, a_H\}$.
5. The state of nature is realized, $\theta_t \in \{\theta_L, \theta_H\}$. And the players decide whether to honor the discretionary payment b_t. Player's payoffs are specified by Equations (2) and (3).

1.1 Solution concept

In this game, the principal's observations are all public in the sense that she does not observe anything that the expert does not observe. Therefore, this is a game of *imperfect public monitoring*. We follow the literature to use *perfect public equilibrium* (PPE) as the solution concept. PPE requires the strategy profile following any public history forms a Nash Equilibrium.[2] We also follow the literature in calling such an equilibrium in a repeated game with transfers a *relational contract*.

1.2 Symmetric contract

Again, we focus on *symmetric contracts* in which the bonus payment is conditional on (a) whether the chosen action is Good or Bad and (b) whether the chosen action is Approved or Disapproved. Denote a symmetric contract by $\mathbf{b} \equiv (b_{GA}, b_{GD}, b_{BA}, b_{BD})$, whose meaning and relation to the bonus function $b(a, \theta, \sigma)$ are depicted in Table 1.

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

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

[2]See Appendix A1 for a detailed description of PPE in this particular game.

2 Analysis

A relational contract is optimal if and only if it generates the highest expected surplus. The reason is that the surplus can be redistributed between the two players by changing the fixed wage in the initial period. Such a change does not affect the incentives except for the participation constraints in the initial period.

According to Levin (2003), for the purpose of characterizing the *optimal relational contract* in our environment, it is without loss of generality to restrict attention to stationary contracts, in which the base wage and the bonus function are the same in each period. That is, for each t, $w_t = w$ and $\hat{b}_t$ is $\hat{b} : \{a_L, a_H\} \times \{\theta_L, \theta_H\} \times \{\sigma_L, \sigma_H\} \to \mathbb{R}$.

Lemma 1. *(Theorem 2 of Levin (2003)) If an optimal contract exists, there are stationary contracts that are optimal.*

Proof. See Appendix A1. □

Due to the stationary nature of the optimal contract, we will suppress the subscript t in the analysis.

2.1 First-best case

The first-best case is the same as in the one-shot model. To achieve the first-best outcome, the principal must (i) follow the public signal if $x < q$ and follow the private signal if $x \geq q$; and (ii) choose an effort that solves:

$$\int_q^1 (F(x) - G(x))dx = \frac{c'(e^{FB})}{p'(e^{FB})}. \tag{4}$$

2.2 Characterizing the optimization problem

We first characterize the expert's strategy provided that the discretionary bonuses are paid. Then, we derive conditions for the enforcement of the discretionary bonuses. Finally, we characterize the maximization problem.

2.2.1 Expert's strategy

The analysis of the optimal action rule in the one-shot model goes through and Lemma 2 follows.[3]

[3] One subtle point is that when proving the non-optimality of a contract with $x^{**} \in [0.5, 1)$ (similar to Appendix A1 in the paper), we need to ensure that the constructed contract does not pay an excessive bonus. This could be shown in two steps. First, the maximum bonus under the constructed

Lemma 2. *Under the optimal symmetric contract,*

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

and there is there is a cutoff $x^* \in (0.5, 1)$ defined by

$$x^* = \frac{q(b_{GA} - b_{BD})}{(1-q)(b_{GD} - b_{BA}) + q(b_{GA} - b_{BD})}, \tag{x^*}$$

such that the expert follows his private signal if $x > x^*$ and follows public opinion otherwise.

Expert's effort level The expert chooses e according to:

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

Same as in the one-shot model, the constraints (Act), (x^*) and (IC) are only concerned with the payment differences, $b_{GD} - b_{BA}$ and $b_{GA} - b_{BD}$.

2.2.2 Dynamic enforcement

Since the bonus payment is discretionary, there is an additional dynamic enforcement constraint. Denote the surplus of the relationship under the symmetric contract **b** by

$$S(\mathbf{b}) = \frac{\int_{0.5}^{x^*} q dH(x;e) + \int_{x^*}^{1} x dH(x;e) - c(e)}{1-\delta}$$

where x^* and e are determined by (x^*) and (IC) respectively. Given $S(\mathbf{b})$, the most extreme outcomes at the end of each period are that either the principal or the expert obtains the entire surplus of the game going forward. Therefore, the maximum difference in bonuses is bounded above by the size of the future surplus discounted by one period. We can normalize the minimum bonus to zero. Then:

$$\begin{aligned} &\min\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} = 0, \\ &\max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} \leq \delta S(\mathbf{b}). \end{aligned} \tag{DE}$$

2.2.3 Characterizing optimal symmetric contract

The optimal symmetric contract solves:

$$\max_{b_{GA}, b_{BD}, b_{GD}, b_{BA}} S(\mathbf{b})$$
$$\text{subject to (Act), } (x^*), \text{ (IC), (DE)}$$

contract is no higher than that under the original contract. Second, the value of the relationship is higher under the constructed contract.

2.3 Sustainability of the first-best surplus

If there is no bound on the bonus payment, the principal could choose a bonus schedule such that (i) $x^* = q$; and (ii) effort induced by (IC) coincides with the first-best effort given by (4). Under this bonus schedule, the first-best surplus is achieved. The following lemma identifies the required bonus schedule:

Lemma 3. *The first-best surplus is achieved if and only if*

$$b_{GA} = b_{GD} = 1,\ b_{BA} = b_{BD} = 0.$$

Proof. *See Appendix A2.* □

From Lemma 3, the bonus schedule that induces the first-best surplus requires:

$$\max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} = 1.$$

Since the restriction on maximum bonus is given by (DE), and the first-best surplus of the relationship is given by:

$$S^{FB} = \frac{\int_{0.5}^{q} q dH(x; e^{FB}) + \int_{q}^{1} x dH(x; e^{FB}) - c(e^{FB})}{1-\delta},$$

we obtain that the first-best surplus may only be achieved when the players are sufficiently patient:

$$\delta \geq \frac{1}{1 + \int_{0.5}^{q} q dH(x; e^{FB}) + \int_{q}^{1} x dH(x; e^{FB}) - c(e^{FB})} \equiv \delta^{FB}. \tag{5}$$

A more precise public opinion corresponds to higher total surplus, which further translates into a lower cutoff discount factor required to sustain the first-best surplus. Therefore, the sustainability of the first-best outcome varies systematically in δ and q. The result is summarized in Lemma 4 and illustrated in Figure 1.

Lemma 4.

(1) *The first-best surplus is never achievable for all values of* $q \in (0.5, 1)$ *when* $\delta \leq 0.5$*;*

(2) *The first-best surplus is achievable for all values of* $q \in (0.5, 1)$ *when* $\delta \geq \lim_{q \to 0.5} \delta^{FB}$*;*

(3) *For every* δ *with* $\delta \in \left(0.5, \lim_{q \to 0.5} \delta^{FB}\right)$*, there exists* $q^{**} \in (0.5, 1)$*, such that*

   (a) *for* $q < q^{**}$*, the first-best surplus is not achievable, and*

   (b) *for* $q \geq q^{**}$*, the first best is achievable.*

Proof. *See Appendix A3.* □

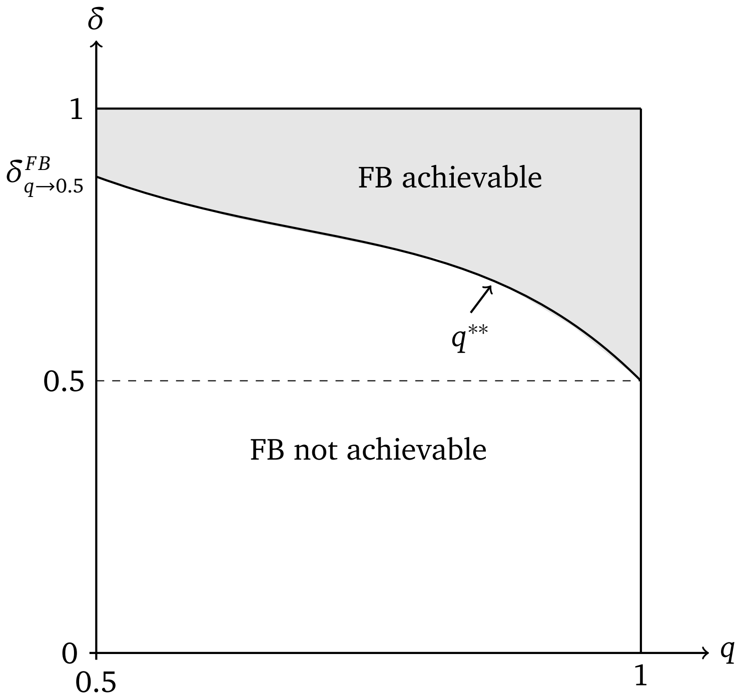

Figure 1: First-best surplus (FB)

2.4 When the first-best surplus is not sustainable

When the first-best surplus is not attainable, the optimal contract may be one with distortion in information acquisition and/or utilization. In addition, the optimal contract must share the feature that the maximum bonus amount is equal to the upper bound. The lemma below formalizes the idea:

Lemma 5. *When the first-best surplus is not attainable, the optimal contract must satisfy:*

$$\begin{aligned} \min\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} &= 0, \\ \max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} &= \delta S(\mathbf{b}). \end{aligned}$$

Proof. *See Appendix A4.* □

By Lemma 2, $b_{GD} > b_{BA}$ and $b_{GA} > b_{BD}$. Accordingly, the maximum bonus must be one of b_{GD} and b_{GA}; the minimum bonus must be one of b_{BA} and b_{GD}. Furthermore, since the expert's cutoff action rule and the IC condition for effort are only concerned with the payment differences $b_{GA} - b_{BD}$ and $b_{GD} - b_{BA}$, we can further restrict attention to the following two types of contracts.

Lemma 6. *It is without loss of generality to focus on the two types of contracts:*

$$b_{BA} = 0,\ b_{GD} = \delta S(\mathbf{b}),\ \text{and}\ b_{BD}, b_{GA} \in [0, \delta S(\mathbf{b})],\ \text{or}$$
$$b_{BD} = 0,\ b_{GA} = \delta S(\mathbf{b}),\ \text{and}\ b_{BA}, b_{GD} \in [0, \delta S(\mathbf{b})].$$

Proof. See Appendix A5. □

We refer to these contracts as *Contract S* and *Contract F*, respectively, as the former contract may induce the expert to be *Stubborn* in the sense that it may induce bias in favor of his private signal over public opinion (i.e. $x^* < q$) and the latter contract may induce the expert to be a *Flip-flopper* in the sense that it may induce bias in favor of public opinion over the private signal (i.e. $x^* > q$). These two classes of contracts are nontrivial and we will analyze them in detail in the next subsection (Section 2.5). Below we formally define Contract S and Contract F, and discuss their implications.

Definition 1 (Contract S)**.** *A Contract S is a tuple* $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ *satisfying*

$$b_{BA} = 0,\ b_{GD} = \delta S(\mathbf{b}),\ b_{BD} = \tilde{\varepsilon},\ b_{GA} = \delta S(\mathbf{b}) - \varepsilon,$$

for $\varepsilon, \tilde{\varepsilon} \in [0, \delta S(\mathbf{b})]$.

Under Contract S, the expert will be rewarded a smaller bonus $\delta S(\mathbf{b}) - \varepsilon$ when his state-matching action also matches public opinion, and/or rewarded a positive bonus $\tilde{\varepsilon}$ when the expert defies public opinion but fails. We know that only $(b_{GA} - b_{BD})$ and $(b_{GD} - b_{BA})$ matter in determining x^* and e. In Contract S:

$$b_{GA} - b_{BD} = \delta S(\mathbf{b}) - \varepsilon - \tilde{\varepsilon},$$
$$b_{GD} - b_{BA} = \delta S(\mathbf{b}).$$

Therefore, the optimal contract would be invariant of reallocation of ε and $\tilde{\varepsilon}$ as long as $(\varepsilon + \tilde{\varepsilon})$ is kept unchanged. In the analysis that follows, it is without loss of generality to focus on $\hat{\varepsilon} = \varepsilon + \tilde{\varepsilon}$, instead of analyzing ε and $\tilde{\varepsilon}$ separately.

From Lemma 2, the expert will follow his own signal if:

$$x \geq x^* = \frac{q(\delta S(\mathbf{b}) - \hat{\varepsilon})}{\delta S(\mathbf{b}) - q\hat{\varepsilon}}.$$

It is easy to see that:

$$x^* \leq q,$$

and:

$$\left.\frac{\partial x^*}{\partial \hat{\varepsilon}}\right|_{S(\mathbf{b})} = -\frac{(1-q)q\delta S(\mathbf{b})}{(\delta S(\mathbf{b}) - q\hat{\varepsilon})^2} < 0. \tag{6}$$

Therefore, Contract S indeed induces the expert to rely less on public opinion. Moreover, when there is more distortion in bonuses, reliance on public opinion further decreases. Clearly, when $\hat{\varepsilon} = 0$, $x^* = q$.

Definition 2 (Contract F)**.** *A Contract F is a tuple* $(w_{GA}, w_{GD}, w_{BA}, w_{BD})$ *satisfying*

$$b_{BD} = 0,\ b_{GA} = \delta S(\mathbf{b}),\ b_{BA} = \tilde{\varepsilon},\ b_{GD} = \delta S(\mathbf{b}) - \varepsilon,$$

for $\varepsilon, \tilde{\varepsilon} \in [0, \delta S(\mathbf{b})]$.

Similar to Contract S, in Contract F, only $(\varepsilon + \tilde{\varepsilon})$ matters, and so it is without loss of generality to focus on $\hat{\varepsilon} = \varepsilon + \tilde{\varepsilon}$. The expert will follow his own signal if:

$$x \geq x^* = \frac{q\delta S(\mathbf{b})}{\delta S(\mathbf{b}) - \hat{\varepsilon} + q\hat{\varepsilon}}.$$

It is easy to see that:

$$x^* \geq q,$$

and:

$$\left.\frac{\partial x^*}{\partial \hat{\varepsilon}}\right|_{S(\mathbf{b})} = \frac{(1-q)q\delta S(\mathbf{b})}{(\delta S(\mathbf{b}) - \hat{\varepsilon} + q\hat{\varepsilon})^2} > 0. \tag{7}$$

Therefore, Contract F encourages the expert to rely more on public opinion compared to a contract with no distortion in information utilization. In addition, when the expert's incentive in information utilization is more distorted, he follows public opinion more frequently. Similar to Contract S, when $\varepsilon = \tilde{\varepsilon} = 0$, Contract F does not induce bias in the utilization of the private information and public opinion.

2.5 Main result

From the above analysis, we know that when the first-best surplus is not attainable, the optimal contract must belong to one of two classes: Contract S or Contract F. By analyzing the two classes of contracts separately, we find that the optimal contract with the form of Contract S sometimes induces distortion, while the optimal contract with the form of Contract F always remains a contract with no distortion in information utilization. Therefore, Contract S weakly dominates Contract F, and we have the following proposition:

Proposition 1. *There exist* δ^**,*

(1) when $\delta \in (0, 0.5]$*, the first-best surplus is never attainable, and there exists a unique cutoff* q^* *such that if* $q > q^*$*, the optimal contract is Contract S with* $\varepsilon > 0$ *or* $\tilde{\varepsilon} > 0$*, or both; and if* $q \leq q^*$*, the optimal contract is the contract with no distortion in information utilization (*$\varepsilon = \tilde{\varepsilon} = 0$ & $x^* = q$*).*

(2) when $\delta \in (0.5, \delta^*)$*, there exist unique cutoffs* q^* *and* q^{**} *with* $q^* < q^{**}$ *such that if* $q \leq q^*$*, or* $q \geq q^{**}$*, the optimal contract is the contract with no distortion in information utilization (*$\varepsilon = \tilde{\varepsilon} = 0$ & $x^* = q$*). For* $q \in (q^*, q^{**})$*, the optimal contract is Contract S with* $\varepsilon > 0$ *or* $\tilde{\varepsilon} > 0$*, or both.*

(3) when $\delta \in [\delta^*, 1]$, the optimal contract is the contract with no distortion in information utilization ($\varepsilon = \tilde{\varepsilon} = 0$ & $x^* = q$).

Proof. See Appendix A6. □

Figure 2 below illustrates the optimal contracts identified in Proposition 1.

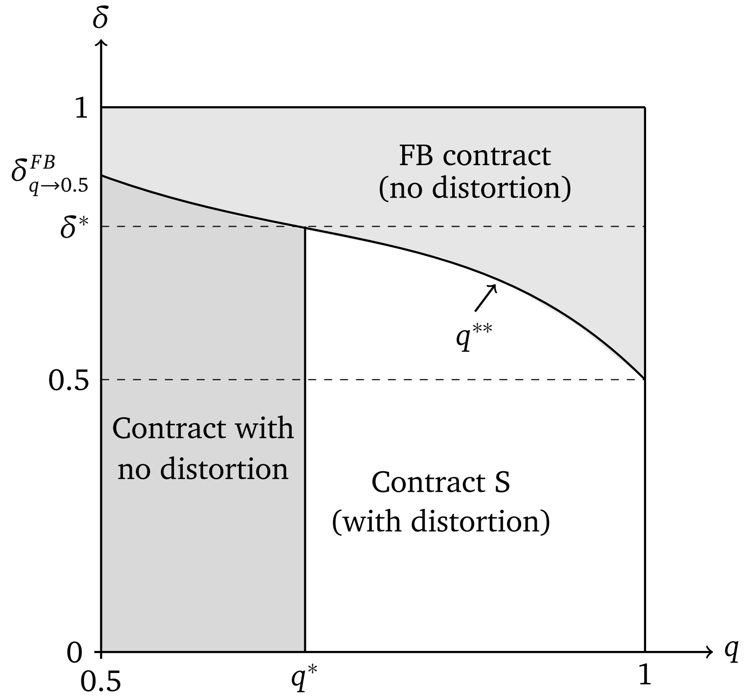

Figure 2: Optimal contract

According to Part (2) of Proposition 1, for the intermediate range of discount factors, when the precision of public opinion is either very high or very low, it is optimal for the principal to offer the expert a contract which induces him to utilize the more precise signal between the private signal and public opinion. However, when the precision of public opinion falls into the intermediate range, it is beneficial for the principal to induce the expert to rely more on his private signal than public opinion, sometimes acting on his private signal even when it is less informative than public opinion. We interpret this behavior as a form of *stubbornness* because the socially optimal choice should be to follow the more informative signal. Obviously, the optimal relational contract induces the expert to inefficiently defy public opinion only when his private information is slightly less informative than public opinion. Therefore, our theory generates very specific predictions about when the expert's behavior exhibits stubbornness.

Now, we provide the intuition for our main finding. The motivation for the expert to exert effort comes from the increased likelihood of matching the state and receiving the bonus. By lowering the bonus when the expert's state-matching action also matches public opinion (and/or increasing the bonus when the state-mismatching action defies public opinion) in Contract S, there are several effects. The first effect is that, on average, the bonus difference of the state-matching and state-mismatching actions is lowered, and this weakens the expert's effort incentive. Notice that this effect is specific to the relational contracting environment where there is an endogenous limited liability constraint on the principal's side. More specifically, in a relational contract, any bonus is bounded above by the future value of the relationship discounted by one period. When we introduce distortion in Contract S, some bonus of the state-matching action must be set below the maximum bonus and/or some bonus of the state-mismatching action must be set above 0 to create the differential in bonuses, causing the expected bonus difference of the state-matching and state-mismatching actions to be reduced.

On the other hand, Contract S with distortion also motivates the expert's incentive by cutting the expert's expected bonus disproportionately when he exerts low effort. More specifically, if the expert follows public opinion, which he is more likely to do if he exerts low effort, the bonus for matching the state is reduced with certainty. If he generates a more informative private signal and follows that instead, when his action matches the state, the bonus is reduced only with probability $q < 1$. According to Proposition 1, the positive effect on effort dominates the negative effect on effort only when public opinion is accurate enough, i.e., when q is larger than q^*. This is because when q is sufficiently close to half, or when public opinion is sufficiently uninformative, x is going to be almost always larger than q whether or not the expert exerts effort. As a result, the effort-motivating effect disappears, causing the net effect of Contract S with distortion on effort to be negative.

When δ is relatively high, i.e., when $\delta > 0.5$, as public opinion becomes sufficiently precise, i.e., q is sufficiently close to one, the first-best surplus becomes achievable.[4] Note that it is impossible to achieve the first-best surplus using a contract with distortion. This implies that for $q > q^{**}$, the optimal contract must not induce distortion in information utilization. When δ is relatively low, i.e., when $\delta \leq 0.5$, the first-best surplus is not attainable for all $q \in (0.5, 1)$[5] In this case, Contract S that induces stubbornness becomes optimal for any $q > q^*$.

Adopting Contract S with distortion also has another two effects on the expert's effort incentive but they do not affect the main finding qualitatively. First, Contract S with distortion in information utilization shifts the cutoff value x^*, causing it to fall

[4]When $\delta \geq \lim_{q \to 0.5} \delta^{FB} > 0.5$, the first-best surplus is always attainable for all $q \in (0.5, 1)$.

[5]Note that there is a discontinuity at $q = 1$. When $q = 1$, the optimal effort is zero and as a result, the first-best surplus is trivially achievable. We do not consider the case of $q = 1$ of practical interest because the outcomes of important decisions are rarely absolutely certain.

below q, and potentially leads to efficiency loss. However, since the first-best cutoff is $x^* = q$, as long as ε (and/or $\tilde{\varepsilon}$) is small, the small shift in x^* only has a second-order effect. This second-order effect is reflected in Equation (16) in the Appendix. Second, when a change in ε (and/or $\tilde{\varepsilon}$) leads to a change in effort, the surplus of the relationship and thus the maximum bonus payable will also change in the same direction, further amplifying the effect of the change in ε (and/or $\tilde{\varepsilon}$). However, since the changes are in the same direction, they only reinforce the change in effort, and do not change the sign of the net effect on effort.

From Proposition 1, we already know that Contract F with distortion is never optimal. We formally state the result in the following corollary:

Corollary 1. *It is never optimal to induce $x^* > q$.*

The intuition for the non-optimality to induce under-utilization of the private signal is as follows. Just like in the case of Contract S, there are four effects of using Contract F with distortion. First, to introduce distortion in Contract F, as under Contract S, some bonus of the state-matching action is set below the maximum bonus and/or some bonus for the state-mismatching action is set at a positive level, causing the expected bonus difference of the state-matching and state-mismatching actions to be lower than that under a contract with no distortion. Second, the anticipation of his own reduced reliance on the private signal motivates the expert to work less hard on the generation of a precise private signal. Third, setting x^* above q leads to inefficient over-utilization of public opinion and thus lowers the principal's expected payoff. Fourth, the maximum bonus payable will change in the same direction as the surplus of the relationship, further amplifying the effect of the change in ε (and/or $\tilde{\varepsilon}$). Since all four effects are negative to the principal's payoff, it is never optimal to induce under-utilization of the expert's private signal.

References

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

A Appendix

A1 Proof of Lemma 1

A1.1 Solution concept

Let U_t^P and U_t^E be the principal's and the expert's expected total discounted payoffs in period t. Suppose that a relational contract specifies an initial payment w_1, $b_1(a_1, \theta_1, \sigma_1)$ and induces the first period effort e_1. Continuation contract gives payoffs $U_2^P(a_1, \theta_1, \sigma_1)$ and $U_2^E(a_1, \theta_1, \sigma_1)$ for each $a_1 \in \{a_L, a_H\}$, $s_1 \in \{s_L, s_H\}$ and $\sigma_1 \in \{\sigma_L, \sigma_H\}$. We assume the most severe punishment: any deviation results in reversion to the static equilibrium in which no trade is possible and the players both obtain their outside options (which are assumed to be 0).

$$\begin{aligned}
U_1^E =& \mathbb{E}_{x,s,\sigma,\theta}[w_1 + b_1\big(\hat{a}_1(x_1, s_1, \sigma_1), \theta_1, \sigma_1\big) - c(e_1); e_1] \\
&+ \delta \mathbb{E}_{x,s,\sigma,\theta}[U_2^E\big(\hat{a}_1(x_1, s_1, \sigma_1), \theta_1, \sigma_1\big); e_1] \\
U_1^P =& \mathbb{E}_{x,s,\sigma,\theta}[r_1(a_1, \theta_1) - w_1 - b_1\big(\hat{a}_1(x_1, s_1, \sigma_1), \theta_1, \sigma_1\big); e_1] \\
&+ \delta \mathbb{E}_{x,s,\sigma,\theta}[U_2^P\big(\hat{a}_1(x_1, s_1, \sigma_1), \theta_1, \sigma_1\big); e_1]
\end{aligned}$$

and

$$S^* = \mathbb{E}_{x,s,\sigma,\theta}[r_1(a_1, \theta_1) - c(e_1); e_1] + \delta \mathbb{E}_{x,s,\sigma,\theta}[S_2\big(\hat{a}_1(x_1, s_1, \sigma_1), \theta_1, \sigma_1\big); e_1]$$

where $S_2\big(a_1, \theta_1, \sigma_1\big)$ is the continuation surplus following outcome $(a_1, \theta_1, \sigma_1)$.

Denote $\rho_t \in (0, 1)$ the expert's posterior belief regarding the state being θ_L in period t. For the strategy profile to form a PPE, it requires

(i) Expert's belief updating process $\hat{\rho}_1(x_1, s_1, \sigma_1)$ is consistent with Bayesian updating.

(ii) Expert's action $a_1(\rho_1, \sigma_1)$ solves

$$W(\rho_1, \sigma_1) \equiv \max_a \; w_1 + \rho_1 \hat{b}(a, \theta_L, \sigma_1) + (1 - \rho_1)\hat{b}(a, \theta_H, \sigma_1) + \delta U_2^E\big(a, \theta_1, \sigma_1\big)$$

for any $\rho_1 \in [0, 1]$ and $\sigma_1 \in \{\sigma_L, \sigma_H\}$.

(iii) Expert's effort level e_1 solves:

$$\max_{e \geq 0} \int_{0.5}^{1} \mathbb{E}_{\{s,\sigma\}}[W\big(\hat{\rho}(x, s, \sigma), \sigma\big) \mid x] \, dH(x; e) - c(e).$$

(iv) Both parties are willing to make the discretionary payment:

$$b_1(a_1,\theta_1,\sigma_1)+\delta U_2^E\big(a,\theta_1,\sigma_1\big)\geq 0;$$
$$-b_1(a_1,\theta_1,\sigma_1)+\delta U_2^P\big(a,\theta_1,\sigma_1\big)\geq 0$$

for all $a_1\in\{a_L,a_H\}$, $s_1\in\{s_L,s_H\}$ and $\sigma_1\in\{\sigma_L,\sigma_H\}$.

(v) Participation constraints:

$$U_1^E\geq 0,\quad U_1^P\geq 0.$$

(vi) The strategies in each continuation game constitute a PPE. In particular, for each (a_1,θ_1,σ_1), the pair $U_2^E\big(a_1,\theta_1,\sigma_1\big)$, $U_2^P\big(a_1,\theta_1,\sigma_1\big)$ constitute PPE payoffs.

A1.2 Stationary contracts

We show the optimality of stationary contracts. Let S^* be the maximum surplus.

We use upper bar to denote the players' actions and the payoffs in an optimal contract. Then $\bar{U}_1^E+\bar{U}_1^P=S^*$. We first show that, $\bar{U}_2^E\big(a_1,\theta_1,\sigma_1\big)+\bar{U}_2^P\big(a_1,\theta_1,\sigma_1\big)=S^*$ for all $\big(a,\theta_1,\sigma_1\big)$. Suppose not, then there exists $\big(\tilde{a},\tilde{\theta}_1,\tilde{\sigma}_1\big)$ such that $\bar{U}_2^E\big(\tilde{a}_1,\tilde{\theta}_1,\tilde{\sigma}_1\big)+\bar{U}_2^P\big(\tilde{a}_1,\tilde{\theta}_1,\tilde{\sigma}_1\big)<S^*$. Then consider increase $\bar{U}_2^P\big(\tilde{a}_1,\tilde{\theta}_1,\tilde{\sigma}_1\big)$ by some small amount $\varepsilon>0$ so that $\bar{U}_2^E\big(\tilde{a}_1,\tilde{\theta}_1,\tilde{\sigma}_1\big)+\bar{U}_2^P\big(\tilde{a}_1,\tilde{\theta}_1,\tilde{\sigma}_1\big)+\varepsilon<S^*$ still holds. So, constraints (vi) is still satisfied. Moreover, such a change does not affect the constraints (i), (ii) and (iii), and relaxes the constraints (iv) and (v). In addition, U_1^P increases, and consequently the total surplus becomes higher than S^*, which is impossible.

Since $\bar{U}_2^E\big(a_1,\theta_1,\sigma_1\big)+\bar{U}_2^P\big(a_1,\theta_1,\sigma_1\big)=S^*$ for all $\big(a,\theta_1,\sigma_1\big)$, we have

$$S^*=\frac{1}{1-\delta}\mathbb{E}_{x,s,\sigma,\theta}[r_1(a_1,\theta_1)-c(e_1);e_1].$$

Then we show that there exist stationary contracts that yield S^*. Define the stationary discretionary payment:

$$b^*(a_1,\theta_1,\sigma_1):=\bar{b}_1(a_1,\theta_1,\sigma_1)+\delta\bar{U}_2^E(a_1,\theta_1,\sigma_1)-\delta\bar{U}_1^E,$$

for all $a_1\in\{a_L,a_H\}$, $s_1\in\{s_L,s_H\}$ and $\sigma_1\in\{\sigma_L,\sigma_H\}$. Define the fixed payment w^*:

$$w^*=(1-\delta)\bar{U}_1^E-\mathbb{E}_{x,s,\sigma,\theta}[b^*\big(\hat{a}_1(x_1,s_1,\sigma_1),\theta_1,\sigma_1\big)-c(\bar{e}_1);\bar{e}_1]$$

The stationary contract has the principal propose w^* and $b^*(a_1,\theta_1,\sigma_1)$ in each period. This contract yields payoffs (U_1^E,U_1^P) with

$$U_1^E=\frac{1}{1-\delta}\mathbb{E}_{x,s,\sigma,\theta}[w^*+b^*\big(\hat{a}_1(x_1,s_1,\sigma_1),\theta_1,\sigma_1\big)-c(e_1);e_1]=\bar{U}_1^E;$$
$$U_1^P=S^*-\bar{U}_1^E=\bar{U}_1^P$$

to the expert and the principal respectively.

We check that the proposed stationary contract forms a PPE. Note that the discretionary payments are defined so that

$$b^*(a_1,\theta_1,\sigma_1)+\delta\bar{U}_1^E=\bar{b}_1(a_1,\theta_1,\sigma_1)+\delta\bar{U}_2^E(a_1,\theta_1,\sigma_1), \tag{8}$$

for all $a_1\in\{a_L,a_H\}$, $s_1\in\{s_L,s_H\}$ and $\sigma_1\in\{\sigma_L,\sigma_H\}$. We could directly check that (ii), (iii) and (iv) are satisfied by substituting (8) into the respective functions. (vi) is satisfied since the stationary contract repeats in each following period.

A2 Proof of Lemma 3

1. "If"

   With $b_{GA}=b_{GD}=1$, $b_{BA}=b_{BD}=0$, (x^*) becomes:

   $$x^*=q,$$

   and (IC) becomes:

   $$\int_q^1(F(x)-G(x))dx=\frac{c'(e)}{p'(e)}$$

   which is the same as the first-best effort given by (4).

2. "Only if"

   The first-best surplus requires: the action rule $x^*=q$ and the optimal effort from (IC) is the same as the first-best effort given by (4).

   (i) For the action rule, from (x^*):

   $$\begin{aligned} x^* &= \frac{q[b_{GA}-b_{BD}]}{(1-q)[b_{GD}-b_{BA}]+q[b_{GA}-b_{BD}]}=q \\ \Rightarrow\quad & b_{GA}-b_{BD}=b_{GD}-b_{BA}. \end{aligned} \tag{9}$$

   (ii) For the optimal effort, with $x^*=q$ and $b_{GA}-b_{BD}=b_{GD}-b_{BA}$, (IC) becomes:

   $$(b_{GA}-b_{BD})\int_q^1(F(x)-G(x))dx=\frac{c'(e)}{p'(e)}. \tag{10}$$

   The only way for (10) and (4) to induce the same level of effort is to set:

   $$b_{GA}-b_{BD}=1.$$

Together with (9), we have:

$$b_{GA} - b_{BD} = b_{GD} - b_{BA} = 1.$$

Since:

$$\min\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} = 0,$$

we have $b_{BD} = 0$, or $b_{BA} = 0$, or both. If $b_{BD} > 0$ and $b_{BA} = 0$, we could always find another bonus schedule with:

$$b'_{BD} = b'_{BA} = 0, \, b'_{GA} = b'_{GD} = 1,$$

that induces the same incentive. Note that in the original contract, $b_{GA} = b_{BD} + 1 > 1$, and thus the sustainability for the newly constructed contract is higher since:

$$\max\{b'_{GA}, b'_{GD}, b'_{BA}, b'_{BD}\} = 1 < b_{GA} = \max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\}.$$

Therefore, a contract with $b_{BD} > 0$ and $b_{BA} = 0$ is dominated. Similarly, a contract with $b_{BD} = 0$ and $b_{BA} > 0$ is also dominated. So:

$$b_{BD} = b_{BA} = 0,$$

and thus:

$$b_{GA} = b_{GD} = 1.$$

A3 Proof of Lemma 4

Notice that higher q results in higher per period reward from the relationship:

$$\text{Envelope Theorem } \frac{d}{dq}\left[\int_{0.5}^{q} q dH(x; e^{FB}) + \int_{q}^{1} x dH(x; e^{FB}) - c(e^{FB})\right] = H(q; e^{FB}) > 0.$$

This higher reward will translate into a larger set of δ sustainable for the relationship since the lower bound δ^{FB} decreases with q. To see this, from (5):

$$\frac{d\delta^{FB}}{dq} = \frac{1}{(1 + \int_{0.5}^{q} q dH(x; e^{FB}) + \int_{q}^{1} x dH(x; e^{FB}) - c(e^{FB}))^2} \cdot (-H(q; e^{FB})) < 0. \quad (11)$$

Part (1): the maximum bonus of 1 is never credible if it is not credible for the highest value of q, i.e., $q \to 1$, and this requires:

$$\delta \leq \delta^{FB}_{q\to 1} = \frac{1}{1 + \int_{0.5}^{1} 1 dH(x; e^{FB}_{q\to 1}) - c(e^{FB}_{q\to 1})} = \frac{1}{2 - c(e^{FB}_{q\to 1})} = \frac{1}{2}.$$

Part (2): the maximum bonus of 1 is always credible if it is credible for the lowest value of q, i.e., $q \to 0.5$, which requires:

$$\delta \geq \delta^{FB}_{q\to 0.5} = \frac{1}{1+\int_{0.5}^{1} x dH(x; e^{FB}_{q\to 0.5}) - c(e^{FB}_{q\to 0.5})} = \frac{1}{2-\int_{0.5}^{1} H(x; e^{FB}_{q\to 0.5})dx - c(e^{FB}_{q\to 0.5})},$$

where $e^{FB}_{q\to 0.5}$ is defined by:

$$\int_{0.5}^{1} (F(x) - G(x))dx = \frac{c'(e^{FB}_{q\to 0.5})}{p'(e^{FB}_{q\to 0.5})}.$$

Part (3): According to Part (1) and (2), when:

$$\delta \in \left(\frac{1}{2}, \delta^{FB}_{q\to 0.5}\right),$$

the first-best surplus is achievable for some, but not all, values of q. Every $\delta \in \left(0.5, \delta^{FB}_{q\to 0.5}\right)$ is the lower bound for some q^{**}. For $q \geq q^{**}$, the first-best surplus is achievable while for $q < q^{**}$, the first-best surplus is not achievable. By (11), the lower bound decreases with q.

A4 Proof of Lemma 5

We prove by contradiction. Suppose, on the contrary:

$$\begin{aligned} \min\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} &= 0, \\ \max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} &< \delta S(\mathbf{b}). \end{aligned}$$

We will then show that this contract is dominated by a scaled contract. Define a scaled contract:

$$b'_{GA} = kb_{GA}, b'_{GD} = kb_{GD}, b'_{BA} = kb_{BA}, b'_{BD} = kb_{BD}$$

with

$$0 < k \leq \frac{\delta S(\mathbf{b})}{\max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\}}.$$

Since

$$\frac{\delta S(\mathbf{b})}{\max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\}} > 1,$$

The constant k could either be smaller or greater than 1.

By construction, the constraint (DE) is still satisfied as long as the total surplus under the newly constructed contract is weakly higher than that under the original contract. We will show below that the condition holds.

From (x^*), the cutoff value x^* is preserved under scaling. However, the induced effort e changes. From (IC), after scaling, FOC of the expert's problem becomes:

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

FOC of the original problem is given by (IC). Since the ratio $\frac{c'(e)}{p'(e)}$ is strictly increasing in e, if $k>1$ (scale up), e under the scaled contract is larger than e under the original contract; if $k<1$ (scale down), e under the scaled contract is smaller than e under the original contract.

We now show that the effect of a marginal change of e while keeping x^* fixed on the value of the relationship

$$S(\mathbf{b})=\frac{\left[\int_{0.5}^{x^*}q\,dH(x;e)+\int_{x^*}^{1}x\,dH(x;e)\right]-c(e)}{1-\delta}$$

is not neutral, i.e.,

$$\left.\frac{dS(\mathbf{b})}{de}\right|_{x^*}=\frac{p'(e)\left[[G(x^*)-F(x^*)](q-x^*)-\int_{x^*}^{1}[G(x)-F(x)]dx\right]-c'(e)}{1-\delta}\neq 0.$$

There are two cases to consider.

1. $x^*\geq q$

$$\begin{aligned}
\left.\frac{dS(\mathbf{b})}{de}\right|_{x^*} &= \frac{p'(e)\left[(G(x^*)-F(x^*))(q-x^*)-\int_{x^*}^{1}(G(x)-F(x))dx\right]-c'(e)}{1-\delta}\\
&= \frac{p'(e)}{1-\delta}\left\{\begin{array}{c}(F(x^*)-G(x^*))(x^*-q)\\ +\{1-[q(b_{GA}-b_{BD})+(1-q)(b_{GD}-b_{BA})]\}\int_{x^*}^{1}(F(x)-G(x))dx\end{array}\right\}\\
&\geq \frac{p'(e)}{1-\delta}\left\{[1-\{q[b_{GA}-b_{BD}]+(1-q)[b_{GD}-b_{BA}]\}]\int_{x^*}^{1}(F(x)-G(x))dx\right\}\\
&> \frac{p'(e)}{1-\delta}[1-\{q+(1-q)\}]\int_{x^*}^{1}(F(x)-G(x))dx=0.
\end{aligned}$$

   The second inequality follows from $b_{GA}-b_{BD}<1$ and $b_{GD}-b_{BA}<1$.

2. $x^*<q$

   Suppose, on the contrary:

$$\left.\frac{dS(\mathbf{b})}{de}\right|_{x^*}=0. \tag{12}$$

We show that we could marginally increase b_{BA} to achieve a higher surplus. First, a marginal increase of x^* increases total surplus:

$$\left.\frac{dS(\mathbf{b})}{dx^*}\right|_e = \frac{(q-x^*)[p(e)g(x)+(1-p(e))f(x)]}{1-\delta} > 0. \tag{13}$$

Second, an increase in b_{BA} induces an increase in x^*:

$$\begin{aligned}\frac{dx^*}{db_{BA}} &= \frac{d}{db_{BA}}\left[\frac{q(b_{GA}-b_{BD})}{(1-q)(b_{GD}-b_{BA})+q(b_{GA}-b_{BD})}\right] \\ &= \frac{q(1-q)(b_{GA}-b_{BD})}{\{(1-q)(b_{GD}-b_{BA})+q(b_{GA}-b_{BD})\}^2} > 0.\end{aligned}$$

So:

$$\frac{dx^*}{db_{BA}} > 0. \tag{14}$$

Then, it follows from (13), (14) and the assumption (12):

$$\begin{aligned}\frac{dS(\mathbf{b})}{db_{BA}} &= \left.\frac{dS(\mathbf{b})}{de}\right|_{x^*}\cdot\frac{de}{db_{BA}} + \left.\frac{dS(\mathbf{b})}{dx^*}\right|_e\cdot\frac{dx^*}{db_{BA}} \\ &= \left.\frac{dS(\mathbf{b})}{dx^*}\right|_e\cdot\frac{dx^*}{db_{BA}} > 0.\end{aligned}$$

Since:

$$\max\{b_{GA}, b_{GD}, b_{BA}, b_{BD}\} < \delta S(\mathbf{b}),$$

we can achieve the higher surplus by increasing the bonus of b_{BA} slightly. This contradicts with the fact that the bonus schedule is optimal. Therefore, the initial supposition must be wrong and we must have:

$$\left.\frac{dS(\mathbf{b})}{de}\right|_{x^*} \neq 0.$$

So, total surplus could be increased by either increasing or decreasing effort through scaling the bonuses in a manner which does not alter x^*. Base wage could be adjusted so that the expert still obtains his outside option of 0.

A5 Proof of Lemma 6

By Lemma 5, the maximum bonus must be $\delta S(\mathbf{b})$. By Lemma 2, $b_{GD} > b_{BA}$ and $b_{GA} > b_{BD}$. So, the maximum bonus must be one of b_{GD} and b_{GA}; the minimum bonus must be one of b_{BA} and b_{GD}. There are four such possibilities. Now, we show that we can restrict attention to two of such pairs: 1) $b_{GA} = \delta S(\mathbf{b})$ and $b_{BD} = 0$; 2) $b_{GD} = \delta S(\mathbf{b})$ and $b_{BA} = 0$.

Suppose, to the contrary, $b_{GA} = \delta S(\mathbf{b})$ and $b_{BD} > 0$. We could construct a new contract $\{b'_{GA}, b'_{GD}, b_{BA}, b_{BD}\}$:

$$b'_{GA} = \delta S(\mathbf{b}) - b_{BD}, b'_{GD} = b_{GD}, b'_{BA} = b_{BA}, b'_{BD} = 0.$$

The new contract induces the same x^* and e. So, the surplus under the new contract is the same as that under the initial contract. If:

$$\max\{b_{GD}, b_{BA}\} < \delta S(\mathbf{b}),$$

then, in the new contract:

$$\max\{b'_{GA}, b'_{GD}, b'_{BA}, b'_{BD}\} < \delta S(\mathbf{b}).$$

by Lemma 5, it must not be optimal. Or if:

$$\max\{b_{GD}, b_{BA}\} = \delta S(\mathbf{b}),$$

then, the initial contract can be replaced by the new contract $\{b'_{GA}, b'_{GD}, b'_{BA}, b'_{BD}\}$ where the maximum bonus is not b'_{GA}. Therefore, if $b_{GA} = \delta S(\mathbf{b})$ and $b_{BD} > 0$, the contract is either not optimal or it can be replaced by a contract of which the maximum bonus is not b_{GA}. A similar logic is applicable to the other pair. Therefore, we could restrict attention to the following two types of contracts:

$$\begin{aligned} b_{BA} &= 0,\ b_{GD} = \delta S(\mathbf{b}),\ \text{and } b_{BD}, b_{GA} \in [0, \delta S(\mathbf{b})],\ \text{or} \\ b_{BD} &= 0,\ b_{GA} = \delta S(\mathbf{b}),\ \text{and } b_{BA}, b_{GD} \in [0, \delta S(\mathbf{b})]. \end{aligned}$$

A6 Proof of Proposition 1

There are two classes of potentially optimal contracts: a contract that induces stubbornness (Contract S) and a contract that induces the flip-flopper behavior (Contract F). We consider the two classes of contracts separately, and analyze the optimal contract within each class of contracts. Then, we combine the results and arrive at Proposition 1.

For Contract S:

1. When the first-best surplus is attainable, Contract S with no distortion outperforms Contract S with distortion. Given δ and q, the attainability of the first-best surplus is given by Lemma 4.

2. When the first-best surplus is not attainable, we obtain a cutoff q^* where there is a switch of dominance between Contract S with $\hat{\varepsilon} = 0$ and Contract S with $\hat{\varepsilon} > 0$.

The difficulty in analyzing these contracts stems from the interplay of the bonus schedule and the value of the relationship. The bonus schedule affects both agent's information acquisition and information utilization incentives, which further affect the surplus of the relationship. In turn, by Lemma 5, the upper bound of the bonus schedule is determined by the discounted value of the relationship. To tackle this problem, we first consider a fixed upper bound of a bonus schedule $r \leq 1$,[6] and then show that allowing for a bonus schedule to adjust only enhances our result.

The bonus schedule for Contract S is thus written as:

$$b_{BA} = 0,\ b_{GD} = r,\ b_{BD} = \tilde{\varepsilon},\ b_{GA} = r - \varepsilon.$$

From (IC), the expert's effort choice is given by:

$$(r - q\hat{\varepsilon}) \int_{x^*}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}, \text{ where } \hat{\varepsilon} = \varepsilon + \tilde{\varepsilon}. \tag{15}$$

Let:

$$D := (r - q\hat{\varepsilon}) \int_{x^*}^{1} (F(x) - G(x))dx,$$

then, (15) becomes:

$$p'(e)D = c'(e),$$

and

$$\frac{de}{d\hat{\varepsilon}} = \left(\frac{p'(e)}{c''(e) - p''(e)D} \right) \frac{dD}{d\hat{\varepsilon}}.$$

Since $p'(e) > 0$, $c''(e) > 0$, $p''(e) < 0$ and $D > 0$, the two derivatives $de/d\hat{\varepsilon}$ and $dD/d\hat{\varepsilon}$ share the same sign. Based on this observation, we will next prove that $de/d\hat{\varepsilon} > 0$ by showing that $dD/d\hat{\varepsilon} > 0$.

$$\frac{dD}{d\hat{\varepsilon}} = \frac{\partial D}{\partial \hat{\varepsilon}} + \frac{\partial D}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}} = q[(1 - x^*)(F(x^*) - G(x^*)) - \int_{x^*}^{1} (F(x) - G(x))dx].$$

As $\hat{\varepsilon}$ goes to zero, x^* goes to q. Therefore:

$$\lim_{\hat{\varepsilon} \to 0} \frac{dD}{d\hat{\varepsilon}} = q \left[(1 - q)(F(q) - G(q)) - \int_{q}^{1} (F(x) - G(x))dx \right].$$

Lemma 7. *There exists a unique q^* in$(0.5, 1)$, such that:*

$$(1 - q)(F(q) - G(q)) - \int_{q}^{1} (F(x) - G(x))dx \begin{cases} < 0 & \text{if } q < q^*, \\ = 0 & \text{if } q = q^*, \\ > 0 & \text{if } q > q^*. \end{cases}$$

[6] When $r = 1$, the first-best surplus is attained.

Proof. Since $g(x)/f(x)$ increases on $(0.5, 1)$, there exists some point $c \in (0.5, 1)$, such that $g(x) \leq f(x)$ for $x \leq c$ and $g(x) > f(x)$ for $x > c$, implying that $F(x) - G(x)$ increases for $x \leq c$ and decreases for $x > c$. Figure 3 below depicts these relationships.

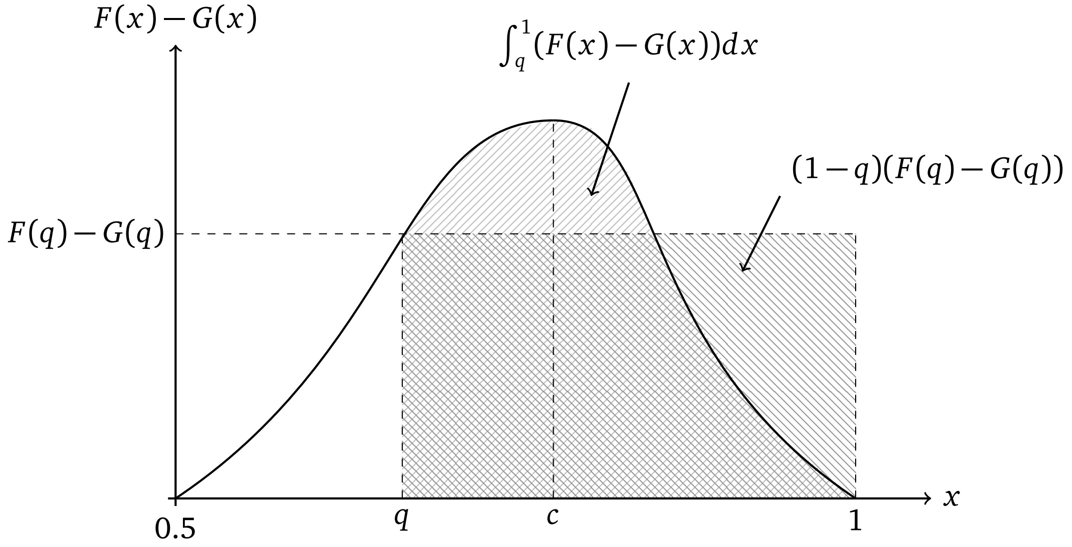

Figure 3: $F(x) - G(x)$, q, c

Let:

$$\begin{aligned}\lambda(q) &= (1-q)(F(q)-G(q)) - \int_q^1 (F(x)-G(x))dx \\ &= \int_q^1 (F(q)-G(q)-(F(x)-G(x)))dx.\end{aligned}$$

We have:

$$\begin{aligned}\lim_{q\to 0.5} \lambda(q) &= 0 - \int_{0.5}^1 (F(x)-G(x))dx < 0, \\ \lambda(c) &= \int_c^1 ((F(c)-G(c))-(F(x)-G(x)))dx > 0.\end{aligned}$$

For $q \in (0, c)$:

$$\lambda'(q) = (f(q) - g(q))(1-q) > 0.$$

So, there exists a unique $q^* \in (0.5, c)$, such that $\lambda(q^*) = 0$.

For $q \in (c,1)$, since $F(x)$—$G(x)$ decreases on $x \in (c,1)$, $F(q)$—$G(q) > F(x)$—$G(x)$ for all $x \in (q,1)$. Then, it follows immediately that:

$$\lambda(q) = \int_q^1 (F(q) - G(q) - (F(x) - G(x)))dx > 0.$$

This completes the proof of the lemma. □

Now, we return to the proof of the main result.

Case I: $q > q^*$ Since:

$$(1-q)(F(q)-G(q)) > \int_q^1 (F(x)-G(x))dx,$$

we have:

$$\lim_{\hat{\varepsilon}\to 0} \frac{dD}{d\hat{\varepsilon}} = q\left[(1-q)(F(q)-G(q)) - \int_q^1 (F(x)-G(x))dx\right] > 0.$$

This further implies:

$$\lim_{\hat{\varepsilon}\to 0} \frac{de}{d\hat{\varepsilon}} > 0.$$

Therefore, inducing *stubbornness* improves effort incentives given that the upper bound of the bonus schedule is fixed.

Next, we show that for a given effort level, shifting the cutoff x^* downward at $x^* = q$ only leads to a second-order loss. Total surplus under Contract S is given by:

$$S^S(\mathbf{b}) = \frac{\int_{0.5}^{x^*} q dH(x;e) + \int_{x^*}^1 x dH(x;e) - c(e)}{1-\delta}.$$

Then:

$$\begin{aligned} \frac{\partial S^S(\mathbf{b})}{\partial x^*} &= \frac{(q-x^*)[p(e)g(x^*) + (1-p(e))f(x^*)]}{1-\delta}, \\ \lim_{\hat{\varepsilon}\to 0} \frac{\partial S^S(\mathbf{b})}{\partial x^*} &= \frac{(q-q)[p(e)g(q) + (1-p(e))f(q)]}{1-\delta} = 0. \end{aligned} \tag{16}$$

The partial derivative of $S^S(\mathbf{b})$ with respect to e is:

$$\begin{aligned} \frac{\partial S^S(\mathbf{b})}{\partial e} &= \frac{p'(e)\left[(G(x^*)-F(x^*))(q-x^*) - \int_{x^*}^1 (G(x)-F(x))dx\right] - c'(e)}{1-\delta} \\ &= \frac{p'(e)}{1-\delta}\left\{(G(x^*)-F(x^*))(q-x^*) + (1-r+q\hat{\varepsilon})\int_{x^*}^1 (F(x)-G(x))dx\right\}, \\ \lim_{\hat{\varepsilon}\to 0} \frac{\partial S^S(\mathbf{b})}{\partial e} &= \frac{p'(e)}{1-\delta}\left\{(G(x^*)-F(x^*))\cdot 0 + (1-r+q\hat{\varepsilon})\int_{x^*}^1 (F(x)-G(x))dx\right\} > 0. \end{aligned}$$

The strict inequality holds because we look at the case where the first-best surplus is not achievable and thus $r < 1$.

Finally:

$$\begin{aligned}\lim_{\hat{\varepsilon}\to 0}\frac{dS^S(\mathbf{b})}{d\hat{\varepsilon}} &= \lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial e}\frac{de}{d\hat{\varepsilon}}+\lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}}\\ &= \lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial e}\lim_{\hat{\varepsilon}\to 0}\frac{de}{d\hat{\varepsilon}} > 0.\end{aligned}$$

Relaxing the constraint of the fixed upper bound on the bonus schedule, the equilibrium surplus will be further increased because an increase in $S^S(\mathbf{b})$ allows the payment of a larger bonus. Therefore, Contract S with some $x^* < q$ outperforms Contract S with no distortion when $q > q^*$.

Case II: $q \le q^*$ First, a marginal increase in $\hat{\varepsilon}$ induces a lower amount of effort:

$$\begin{aligned}&(1-q)(F(q)-G(q)) \le \int_q^1 (F(x)-G(x))dx\\ \Rightarrow\ & \lim_{\hat{\varepsilon}\to 0}\frac{dD}{d\hat{\varepsilon}} = q\left[(1-q)(F(q)-G(q))-\int_q^1 (F(x)-G(x))dx\right]\le 0\\ \Rightarrow\ & \lim_{\hat{\varepsilon}\to 0}\frac{de}{d\hat{\varepsilon}} \le 0.\end{aligned}$$

(16) still applies. Then, a marginal increase in $\hat{\varepsilon}$ induces a lower total surplus:

$$\begin{aligned}\lim_{\hat{\varepsilon}\to 0}\frac{dS^S(\mathbf{b})}{d\hat{\varepsilon}} &= \lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial e}\frac{de}{d\hat{\varepsilon}}+\lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}}\\ &= \lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial e}\lim_{\hat{\varepsilon}\to 0}\frac{de}{d\hat{\varepsilon}} \le 0.\end{aligned}$$

To obtain the global result, we analyze $\frac{dS^S(\mathbf{b})}{d\hat{\varepsilon}}$ for any $\hat{\varepsilon} > 0$. Since, by (x^*):

$$x^* = \frac{q(r-\hat{\varepsilon})}{r-q\hat{\varepsilon}} = q-\frac{q(1-q)\hat{\varepsilon}}{r-q\hat{\varepsilon}} < q \text{ when } \hat{\varepsilon} > 0,$$

we have $x^* < q \le q^*$. Then:

$$\frac{dD}{d\hat{\varepsilon}} = q\left[(1-x^*)(F(x^*)-G(x^*))-\int_{x^*}^1 (F(x)-G(x))dx\right] < 0 \Rightarrow \frac{de}{d\hat{\varepsilon}} < 0. \quad (17)$$

The partial derivative of $S^S(\mathbf{b})$ with respect to e could be written as:

$$\frac{\partial S^S(\mathbf{b})}{\partial e} = \frac{p'(e)\left[\int_{0.5}^{x^*} qd(G(x)-F(x))+\int_{x^*}^1 xd(G(x)-F(x))-D\right]}{1-\delta}.$$

Let:

$$\theta = \left[\int_{0.5}^{x^*} qd(G(x)-F(x)) + \int_{x^*}^{1} xd(G(x)-F(x)) - D\right].$$

Then, since $p'(e) > 0$ and $1-\delta > 0$,

$$\text{sgn}\left(\frac{\partial S^S(\mathbf{b})}{\partial e}\right) = \text{sgn}(\theta).$$

Differentiating θ with respect to $\hat{\varepsilon}$, we have:

$$\frac{d\theta}{d\hat{\varepsilon}} = -[(q-x^*)(g(x^*)-f(x^*))]\left[\frac{(1-q)qr}{(r-q\hat{\varepsilon})^2}\right] - \frac{dD}{d\hat{\varepsilon}} > 0.$$

The inequality holds because $q-x^* > 0$, $g(x^*)-f(x^*) < 0$ (since x^* is to the left of c), and $\frac{dD}{d\hat{\varepsilon}} < 0$.

We also have that $\lim_{\hat{\varepsilon}\to 0}\theta > 0$ since $\lim_{\hat{\varepsilon}\to 0}\frac{\partial S^S(\mathbf{b})}{\partial e} > 0$. So $\theta > 0$ for all $\hat{\varepsilon} > 0$. This implies:

$$\frac{\partial S^S(\mathbf{b})}{\partial e} > 0. \tag{18}$$

Besides, a marginal increase of x^* increases $S^S(\mathbf{b})$:

$$\frac{\partial S^S(\mathbf{b})}{\partial x^*} = \frac{(q-x^*)[p(e)g(x^*)+(1-p(e))f(x^*)]}{1-\delta} > 0. \tag{19}$$

Therefore, from (17), (18), (19) and (6):

$$\frac{dS^S(\mathbf{b})}{d\hat{\varepsilon}} = \frac{\partial S^S(\mathbf{b})}{\partial e}\frac{de}{d\hat{\varepsilon}} + \frac{\partial S^S(\mathbf{b})}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}} < 0 \text{ for } \hat{\varepsilon} > 0.$$

Relaxing the constraint of the fixed upper bound on the bonus schedule, the equilibrium surplus will further decrease because a decrease in $S^S(\mathbf{b})$ allows the payment of a smaller bonus only. This implies that it is optimal to set $\hat{\varepsilon} = 0$, i.e. Contract S with no distortion outperforms Contract S with distortion when $q \le q^*$.

Now, we compare the magnitude of q^* and q^{**} for $\delta \in \left(0.5, \delta^{FB}_{q\to 0.5}\right)$ in which range $q^{**} \in (0.5, 1)$ exists. On one hand, q^{**} decreases with δ. We also know that when $\delta = \frac{1}{2}$, $q^{**} = 1$, and when $\delta = \delta^{FB}_{q\to 0.5}$, $q^{**} = 0.5$. On the other hand, $q^* \in (0.5, c)$ is independent of δ. So, there exists δ^*, such that $q^{**} = q^*$, and when $\delta < \delta^*$, $q^{**} > q^*$; when $\delta \ge \delta^*$, $q^{**} \le q^*$.

When $q^{**} > q^*$, corresponding to $\delta \in (0.5, \delta^*)$, the optimal contract with the form of Contract S is the contract with no distortion if $q \le q^*$ (due to the dominance of Contract S with $\hat{\varepsilon} = 0$), or $q \ge q^{**}$ (due to the attainability of the first-best surplus), and the optimal contract with the form of Contract S is Contract S with distortion if $q \in (q^*, q^{**})$, corresponding to part (2) of Proposition 1.

When $q^{**} \leq q^*$, corresponding to $\delta \in \left[\delta^*, \delta^{FB}_{q \to 0.5}\right)$, since $q > q^{**}$ achieves the first-best surplus and when $q < q^*$, Contract S with no distortion outperforms Contract S with distortion, the optimal contract with the form of Contract S is always the one with no distortion. Together with the fact that the first-best surplus is always attainable when $\delta \in \left[\delta^{FB}_{q \to 0.5}, 1\right]$, for $\delta \in [\delta^*, 1]$, the optimal contract with the form of Contract S is the contract with no distortion, corresponding to part (3) of Proposition 1.

Part (1) of Proposition 1 comes directly from the fact that the first-best surplus is never attainable when $\delta \in (0, 0.5]$. The cutoff q^* governs the dominance of Contract S with and without distortion when the first-best surplus is not attainable.

For Contract F:

1. When the first-best surplus is attainable, Contract F with no distortion outperforms Contract F with distortion.
2. When the first-best surplus is not attainable, we will show below that Contract F with $\hat{\varepsilon} = 0$ also outperforms Contract F with $\hat{\varepsilon} > 0$.

We first prove that the bonus schedule in Contract F with distortion makes the principal to induce less effort from the expert. To show this, we adopt a similar strategy as we do for Contract S and first consider a fixed upper bound of a bonus schedule r. From (IC), the expert's effort choice is given by:

$$(r - \hat{\varepsilon} + q\hat{\varepsilon}) \int_{x^*}^{1} (F(x) - G(x))dx = \frac{c'(e)}{p'(e)}.$$

Similar to D, we define D' as follows:

$$D' := (r - \hat{\varepsilon} + q\hat{\varepsilon}) \int_{x^*}^{1} (F(x) - G(x))dx.$$

Proving that e increases in $\hat{\varepsilon}$ is equivalent to proving that D' increases in $\hat{\varepsilon}$.

$$\frac{dD'}{d\hat{\varepsilon}} = \frac{\partial D'}{\partial \hat{\varepsilon}} + \frac{\partial D'}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}} = (1-q)\left[x^*(G(x^*) - F(x^*)) + \int_{x^*}^{1} (G(x) - F(x))dx\right] < 0.$$

Therefore, inducing under-utilization of the private signal reduces effort incentives given that the upper bound of the bonus schedule is fixed, i.e.,

$$\frac{de}{d\hat{\varepsilon}} < 0. \tag{20}$$

Next, we show that for a given effort level, shifting the cutoff x^* upward leads to a loss. The total surplus is given by:

$$S^U(\mathbf{b}) := \frac{\int_{0.5}^{x^*} q\, dH(x;e) + \int_{x^*}^{1} x\, dH(x;e) - c(e)}{1-\delta}.$$

Distortion of x^* leads to loss:

$$\frac{\partial S^U(\mathbf{b})}{\partial x^*} = \frac{(q-x^*)[p(e)g(x^*) + (1-p(e))f(x^*)]}{1-\delta} \leq 0. \tag{21}$$

An increase in effort induces a higher surplus:

$$\begin{aligned}\frac{\partial S^U(\mathbf{b})}{\partial e} &= \frac{p'(e)\left[\int_{0.5}^{x^*} q d(G(x)-F(x)) + \int_{x^*}^{1} x d(G(x)-F(x)) - D'\right]}{1-\delta} \\ &= \frac{p'(e)}{1-\delta}\left[(x^*-q)(F(x^*)-G(x^*)) + (1-r+\hat{\varepsilon}(1-q))\int_{x^*}^{1}(F(x)-G(x))dx\right] \\ &> 0. \end{aligned} \tag{22}$$

The inequality follows from $q \leq x^*$ and $r < 1$.

Finally, from (20), (21), (22) and (7):

$$\frac{dS^U(\mathbf{b})}{d\hat{\varepsilon}} = \frac{\partial S^U(\mathbf{b})}{\partial e}\frac{de}{d\hat{\varepsilon}} + \frac{\partial S^U(\mathbf{b})}{\partial x^*}\frac{dx^*}{d\hat{\varepsilon}} < 0.$$

Relaxing the constraint of the fixed upper bound on the bonus schedule, the equilibrium surplus will be further decreased because a decrease in $S^U(\mathbf{b})$ constrains the payment of a bonus to a lower level. Therefore, the optimal contract with the form of Contract F is the contract with no distortion.

Since Contract F with distortion is always dominated by a contract with no distortion, Contract S weakly dominates Contract F. So, the optimal contract must be of the form of Contract S, and Proposition 1 follows.