UNIVERSITY OF COPENHAGEN

How Ego-threats Facilitate Contracts Based on Subjective Evaluations

Sebald, Alexander; Walzl, Markus

UNIVERSITY OF COPENHAGEN

Discussion Papers
Department of Economics
University of Copenhagen

No. 08-19

How Ego-threats Facilitate Contracts Based on
Subjective Evaluations

Alexander Sebald
Markus Walzl

How ego-threats facilitate contracts based on subjective evaluations *

Alexander Sebald[†]   Markus Walzl[‡]

September 17, 2008

Abstract

We show that individuals' desire to protect their self-esteem against ego-threatening feedback can mitigate moral hazard in environments with purely subjective performance evaluations. In line with evidence from social psychology we assume that agents' react aggressively to evaluations by the principal which do not coincide with their own positive self-perceptions and thereby generate costs of conflict for the principal. We identify conditions for a positive welfare effect of increasing costs of conflict or increasing sensitivity to ego-threats, and a negative welfare effect of a more informative information technology. As a consequence, principals may choose imperfect information technologies in equilibrium even if the signal quality is costless.

Keywords: Contracts, Subjective Evaluations, Self-Esteem, Ego-Threats.

JEL classification: D01; D02; D82; D86; J41.

1 Introduction

Since the 1890s self-esteem is one of the most intensively studied concepts in social psychology [see *e.g.* James (1890)]. It refers to people's self-evaluation or, in other words, the belief they hold about their self-worth. Everywhere people seem to care about it, try to enhance, maintain and protect it [see *e.g.* Greenwald (1980)]. Anything that gives a boost in self-esteem is almost universally welcome. People feel good when their self-perception is high and rising, and people feel bad when it is low or dropping. Hardly anyone enjoys events that constitute a blow or a loss to their self-esteem [Baumeister (2005)].

In recent years also economists have started to acknowledge the importance of self-esteem in decision making and strategic interactions [*e.g.* Köszegi (2006), Bénabou & Tirole (2002), Compte & Postlewaite (2004), Ellingsen & Johannesson (2008)]. It is argued that people strive for positive self-perceptions because it entails a consumption, signaling and motivational

value. Köszegi (2006), for example, endows individuals with 'ego-utility' and demonstrates the effects on choice between more or less ambitious tasks. In particular, this model explains the phenomenon of overconfidence by individuals who update believes according to Bayes' rule. Bénabou & Tirole (2002) and Compte & Postlewaite (2004), on the other hand, center on the motivational value of self-confidence. It is argued that confidence in one's ability and efficacy can help individuals to undertake more ambitious goals. When people have imperfect knowledge about their own ability and/or when effort and ability are complements, then more self-confidence enhances peoples' motivation to act [Bénabou & Tirole (2002): 873].

Psychologists, however, have not only identified the implicit impact of self-esteem on information processing and motivation, but also stress the individual's eagerness to actively maintain and protect positive self-perceptions [Greenwald (1980), Bushman & Baumeister (1998), Baumeister (2005)]. First, people protect their self-esteem by systematically taking credit for success and denying blame for failure. Second, people have a tendency to uncritically accept positive feedback and eagerly search for flaws/faults in other's criticism [*e.g.* Baumeister (2005), Greenwald (1980)]. Third and most importantly for our investigation, psychologists have found that conflicts and aggression tend to result from positive self-images that are challenged or threatened [*e.g.* Baird (1977), Raskin et al (1991), Bushman & Baumeister (1998)]. It is argued that hostile aggression is an expression of the self's rejection of ego-threatening evaluations received from other people [*e.g.* Baumeister et al (1996)]. People with high self-esteem usually hold confident and highly favorable ideas about themselves, *i.e.* they exhibit ego-involvement, and react belligerently to ego-threatening feedback from others [Baird (1977), Shrauger & Lund (1975) and Korman (1969)]. Furthermore, these behavioral reactions have been found to be the stronger the lower the perceived quality of the feedback source [*e.g.* Albright and Levy (1995), Steelmann and Rutkowski (2004), Roberson and Stewart (2006)]. The perceived accuracy of feedback information and the perceived competence of the feedback source (*i.e.* the appraiser's 'knowledge' of the employee's job and performance) are two important elements of evaluation processes determining fairness perceptions of employees [see *e.g.* Landy et al (1978), Greenberg (1986), Evans & McShane (1988), Fedor et al (1989), Shapiro et al (1994), Taylor et al (1995), Leung & Morris (2001), Roberson & Stewart (2006)]. It has consistently been found that the higher the perceived accuracy of feedback information and the higher the perceived competence/'knowledge' of the feedback source, the higher the employees' acceptance of negative feedback information and the lower the level of conflict in the relation between feedback source and feedback recipient.

In this paper we formalize these findings and analyze the impact of aggressive reactions to ego-threatening feedback on principal-agent relationships. More specifically, we show how the individuals' desire to protect their self-esteem can explain the existence of short-term or one-shot contractual relationships in environments with unobservable effort and subjective

performance measures. We concentrate on situations in which neither effort nor output can be measured objectively as these constitute exactly the settings in which disagreements about effort and performance (and corresponding ego-threats) can arise.

In reality, it is very often impossible to objectively measure workers' and especially managers' individual contributions to the success of projects. Therefore it is widely prevalent to (also) take into account subjective evaluations in performance pay. Already in 1981 the Bureau of National Affairs reports, for example, that pay for performance systems involving subjective measures are more common than those involving only objective performance signals. Furthermore, Milkovich and Wigdor (1991) and Levine (2003) cite more recent evidence on the wide usage of subjective performance appraisal systems in performance pay in *e.g.* investment banks, law firms and consultancies.

Our paper considers the following set-up. A principal wants to motivate an agent to spend effort on a complex good or service. Neither the agent's effort nor the outcome of the project (the quality of the good or service) is publicly observable. However, the principal and the agent receive private, *i.e.* subjective, signals about the effort of the agent. These signals are imperfectly correlated with each other and to the actual effort level. To motivate the agent to spend positive effort, a contract has to specify payments which increase in the subjective signal of the principal (an increase in the reported signal of the agent would just motivate him to misrepresent his information). However, due to the imperfect signal technology the principal can credibly report that he has received a signal of low effort regardless of his actual private information. As payments increase in the subjective signal of the principal, he is always better off by misrepresenting positive information and paying the agent the minimum wage. This will be anticipated by the agent and subgame - perfect equilibrium efforts are zero, *i.e.* no principal agent relationship can be established.

In a recent paper [MacLeod (2003)], it has been assumed that the principal can credibly promise to make payments to a third party (contingent on the signal configuration). In the simplest case of two different performance signals, the optimal contract fixes a payment from the principal to a third party if she pays the agent according to a bad signal and the agent reports a good signal which satisfies the principal's truthtelling constraint. The complete flexibility of third-party payments thereby ensures that a relationship (*i.e.* a positive effort level) can be established regardless of the parameters of the model (*e.g.* the correlation between the principal's and the agent's signal, the size of the project etc.). Of course this result crucially depends on the credibility of payments to the third party. In particular, while the principal cannot credibly promise the agent to report his signal truthfully, it is assumed that he can make such a promise to the third party. To explain the widespread use of subjective information in particular in labor market relations, MacLeod (2003) refers to the third party payments as anticipations of future conflict in an un-modelled dynamic game.

In this paper, we explicitly model the conflict discussed in Macleod (2003) and show that a principal-agent relationship can be established on the basis of subjective performance evaluations, if the agent tries to defend his self-esteem through the creation of conflict or aggressive actions. In line with the aforementioned psychological evidence, we assume that the agent perceives a negative psychological payoff from ego-threatening performance evaluations by the principal. He suffers from bad performance evaluations by the principal, whenever she does not share his opinion based on his own subjective signal. Furthermore, we assume that he suffers the more the lower the accuracy of the information technology used by the principal. The agent can reduce his negative psychological payoff through conflict/trouble imposed on the principal, *e.g.* the agent goes to court in order to enforce the bonus payment, steals, or refuses to cooperate on other tasks.[1] If the agent creates trouble, the principal will face costs of conflict.[2] The costs of conflict play the very same role as MacLeod (2003)'s third-party payments - they enforce truth-telling by the principal. In our setting, however, costs of conflict are not at the principal's disposal but rather depend on the agent's sensitivity to ego-threats, the quality of the information technology *etc.*. Our analysis identifies conditions on conflict levels, project returns, the quality of information, and the sensitivity to ego-threats which promote or rule-out the implementation of positive equilibrium effort levels. In particular, we identify conditions for a positive welfare effect of increasing costs of conflict or increasing sensitivity to ego-threats, and a negative welfare effect of a more informative information technology. As a consequence, it can be shown that principals may choose imperfect information technologies in equilibrium even if the signal quality is costless.

Our model is related but conceptually different from Ellingsen & Johannesson (2008)'s model of self-esteem. They model a situation in which agents sense a psychological payoff from being esteemed by others (and thereby refer to the motivational value of self-esteem – see above). Agents in their setting take pride in what others think about them, *i.e.*,, agents would derive utility from their belief about the principal's evaluation of their performance in our setting. Wage payments contingent on the principal's subjective evaluation would then up-date the agents belief about the principal's appraisal. But as the principal has an incentive to misrepresent positive information as long as he does not expect any conflict (see above), such a psychological payoff structure would not establish positive equilibrium efforts in our setting.

The organization of the paper is as follows: In Section 2 we present the principal-agent relation and the psychological payoff structure. As a benchmark, Section 3 analyzes the situation of pure moral hazard and determines the optimal effort choice and comparative

---

[1]Note, all that counts is that these conflicts are anticipated as costs by the principal.

[2]This mechanism could be interpreted as negative reciprocity. Unlike the existing models of reciprocity [*e.g.,* Rabin (1993), Dufwenberg & Kirchsteiger (2004) and Falk & Fischbacher (2006)], however, what is considered psychologically costly in our model does not depend on beliefs about strategies and their associated outcomes, but rather on (reported) signal constellations.

statics of social welfare in the absence of binding truth-telling constraints. Section 4 continues with an analysis of the impact of binding truth-telling constraints on optimal effort choice and social welfare. While Sections 3 and 4 consider an exogenously given information technology, Section 5 will investigate the principal's optimal choice of an information technology. Section 6 concludes with some remarks on the practical implications of our model and its robustness.

2   The model

In this section we introduce the principal-agent relationship and present a psychological payoff structure which captures the empirical evidence on self-esteem and ego-threats from social psychology. Furthermore we characterize the first best solution and present auxiliary results on the agent's decision on conflict creation and the optimality of simple bonus contracts.

Production Technology   Assume there is a risk-neutral principal, P, who decides upon undertaking a project which generates a value of $\phi > 0$ if successful. The project requires effort of an agent, A. Assume that if the agent spends effort $p \in [0, 1]$, the project will be successful (create value ϕ) with probability p. The project is a complex good or service and its success is not verifiable, *i.e.* contracts contingent on the generation of ϕ are not feasible.

Information Technology   Neither principal nor agent can observe whether the project is successful or not. Rather, both form an opinion about the agent's performance during the production process. Formally, they receive private signals about the agent's performance. The principal receives $s_P \in S_P$, where $S_P = \{L, H\}$, *i.e.* the principal's opinion can be such that he regards the agent's performance as either high (H) or low (L). Analogously, the agent receives $s_A \in S_A$ with $S_A = \{L, H\}$. The signals s_P and s_A are non-verifiable private pieces of information of the principal and the agent, respectively.

The signals are informative with respect to the success of the project. If the project is not successful (which happens with probability $(1-p)$), principal and agent receive the signal $s_P = s_A = L$. If the project is successful, the principal receives the signal $s_P = H$ with probability g, the agent receives the same signal with probability ρ and receives $s_A = H$ as an independent signal with probability x. Hence, g measures the quality of the principal's signal, ρ indicates the correlation between the agent's and the principal's signal - or the counter-probability of an independent judgment - and x quantifies the quality of the agent's signal if he forms an independent judgment (*i.e.,* we adopt the specification of the information technology in Mcleod (2003), p.228).

Assumption 1. *Information Technology*

We assume that the principal's and the agent's signal are imperfect, i.e., $g \in (0,1)$ and $x \in (0,1)$, and positively but imperfectly correlated, i.e., $\rho \in (0,1)$.

We denote by γ_{kl} the conditional probability that $s_P = k$ and $s_A = l$ given that the project is a success. Then, the ex-ante probability for the signal pair $s_P = L$ and $s_A = H$, for instance, will be $p\gamma_{LH} = p(1-g)(1-\rho)x$.[3] Note that by Assumption 1, $\gamma_{HH}\gamma_{LL} > \gamma_{HL}\gamma_{LH}$.

The Game The timing of the game is as follows:

1. The principal offers a contract to the agent and the agent decides upon acceptance.[4] Upfront payments are arranged.

2. The agent decides upon effort p.

3. The project generates value ϕ with probability p.

4. The principal receives s_P and the agent receives s_A. The principal and the agent report (not necessarily truth-fully) on s_P and s_A. Denote the reports by t_P and t_A, respectively. t_P and t_A are verifiable.

5. The payments contingent on t_P and t_A are arranged.

6. Contingent on s_A and received payments, the agent decides upon retaliation (with effort q).

Agent For an effort of p the agent incurs costs $v(p)$ with $v \in C^2$, $v(0) = 0$, $v'(0) = 0$, $v''(p) > 0$ and $\lim_{p \to 1} v(p) = \infty$.

First Best Effort Level Had the principal access to the agent's production technology, his effort choice would solve $v'(p) = \phi$. For further reference, we will denote the first best effort level by p_{FB} and the respective surplus by Π_{FB}. Our assumptions on $v(p)$ ensure that $p_{FB} \in (0, 1)$.

Psychological Payoffs The agent is risk-neutral and senses a psychological payoff that depends on his opinion about his own performance, s_A, and the reported opinion of the principal, t_P. More specifically, the agent's utility function reads:

$$U = w - v(p) - Y(t_P, s_A, g)(1-q) - c(q) \tag{1}$$

Thereby, w denotes the wage payment, $Y(t_P, s_A, g)$ represents the agent's psychological payoff for a given configuration of (reported) signals and a given quality of the principal's signal, q is the level of conflict (or retaliation) created by the agent and $c(q)$ is the agent's cost for the level of conflict q with $c \in C^2$, $c(0) = 0$, $c'(0) = 0$, $c''(q) > 0$ and $\lim_{q \to 1} c(q) = \infty$.

---

[3] All γ_{kl} as functions of g, ρ, and x can be found in Appendix 8.1.

[4] In section 5, the principal will in addition also choose the quality of his own signal (g).

We continue with a specification of $Y(t_P, s_A, g)$ which tries to capture the empirical evidence from social psychology on self-esteem, ego-threats, and retaliatory behavior.

Assumption 2. *Psychological Costs*

(i) $Y(t_P, s_A = L, g) = 0$ *for all* t_P *and* g.

(ii) $Y(t_P = H, s_A, g) = 0$ *for all* s_A *and* g.

(iii) $Y(L, H, g) \in C^1$ *and* $Y(L, H, g) > 0$, $\frac{dY(L,H,g)}{dg} \leq 0$ *for all* g.

Part (i) captures that individuals with low self-esteem (represented by $s_A = L$) do not exhibit ego-involvement and show less reaction to feedback (be it confirming or threatening) [see *e.g.* Baumeister, Smart & Boden (1996)]. Parts (ii) and (iii) respectively formalize the finding that individuals who hold a high opinion about themselves and are 'ego-involved' ($s_A = H$) uncritically accept positive or confirming feedback [see *e.g.* Baumeister (2005)] - formalized by zero psychological costs - and suffer from negative or threatening assessments [see *e.g.* Bushman and Baumeister (1998)] – represented by non-zero psychological cost in our model. Furthermore, we assume that psychological costs (weakly) decrease in the quality of the feedback source (*i.e.*, the quality of the principal's signal parameterized by g) which captures the observation that individuals are the more willing to accept negative feedback the higher the perceived accuracy of the feedback information [*e.g.* Albright and Levy (1995), Steelmann and Rutkowski (2004), Roberson and Stewart (2006)].

In response to an ego-threat the agent can reduce his psychological costs that arise from the deviant (reported) opinions about his performance by creating conflict/trouble (as observed by [Baird (1977), Shrauger & Lund (1975) and Korman (1969)]). For further reference, we summarize some results concerning the agent's optimal conflict level.

Lemma 1. *Conflict Creation*

Suppose $Y(t_P, s_A, g)$ satisfies Assumption 2.

(i) *Then, the agent chooses $q = argmax(Y(t_P, s_A, g)(1 - q) - c(q))$.*

(ii) *Suppose $s_A = L$ and/or $t_P = H$. Then, $Y(t_P, s_A, g) = 0$ and the agent chooses $q = 0$.*

(ii) *Suppose $s_A = H$ and $t_P = L$. Then, the agent chooses $q \in (0, 1)$.*

Proof. Follows from Eqn. 1 and Assumption 2. $\qquad\square$

According to Lemma 1, the agent retaliates (*i.e.*, $q > 0$) if and only if $s_A = H$ and $t_P = L$. The agent retaliates if he has a high opinion of himself and his ego / self-perception is threatened. For further reference we abbreviate $Y(L, H, g) \equiv Y$. [5] Moreover, $q^* > 0$ will

---

[5]In section 5, g is endogenized and we will refer to $Y(L, H, g)$ as $Y(g)$.

henceforth denote the conflict level for the configuration $t_P = L$ and $s_A = H$. As the agent chooses $q = 0$ for all other configurations, no confusion should arise. Note that the higher the psychological costs created by the difference in the principal's and agent's evaluation (Y), the higher the level of conflict q^*. In particular, the poorer the information technology of the principal – the less he is regarded "competent" by the agent–, the more belligerently the agent will react to ego-threats (*i.e.*, $\frac{dq^*}{dg} \le 0$).[6] We assume throughout this paper that Assumptions 1 and 2 are satisfied.

Principal In contrast to the agent, the principal only cares about his profit

$$\Pi = p\phi - E\{w\} - E\{q\}\psi, \tag{2}$$

where $p\phi$ is the expected benefit generated by the agent, $E\{w\}$ are the expected wage cost of employing the agent, and $E\{q\}\psi$ are the expected costs of conflict due to retaliation. As our assumptions on $c(q)$ ensure that $q \in [0, 1]$, we can interpret q as the probability with which the agent creates costs of $\psi > 0$ for the principal. First best profits are given by $\Pi_{FB} = p_{FB}\phi - v(p_{FB})$.

Contracts In our setting with unobservable effort and subjective measures of performance, a contract Γ can only be contingent on the reported subjective opinions of the principal and the agent. Hence, a contract fixes payments for all configurations of reports t_P and t_A and reads $\Gamma = \{w_{kl} \mid k \in S_P, l \in S_A\}$. The agent accepts a contract if he expects a (weakly) positive utility from it (individual rationality) and chooses p as to maximize his utility (incentive compatibility). If a contract Γ is individually rational and the agent chooses effort p, we say that Γ *implements* p. Principal and agent report their opinions, *i.e.* signals, truthfully if and only if they weakly benefit from doing so.

Cost Minimizing Contracts How do optimal contracts look like given that effort is unobservable, performance measures are subjective and agents try to protect a positive self-image through the creation of conflict? A standard application of the revelation principle implies that we can restrict ourselves to simple bonus contracts without any loss of generality.

Lemma 2. *Reduced Form Contracts*

Suppose there exists a contract Γ which implements $p > 0$. Then, there always exists a contract $\hat{\Gamma}$ which implements p at weakly lower costs and

(i) Principal and agent tell the truth.

(ii) $w_{kl} = w_{km} \equiv w_k$ for all $k \in S_P$ and $l, m \in S_A$.

---

[6]The inequality is strict if and only if $\frac{dY(L,H,g)}{dg} < 0$.

(iii) $w_H > w_L$.

Proof. See Appendix 8.2. □

For convenience, we define $w_H = f + b$, $w_L = f$ and $\Gamma = (b, f)$. By Lemma 2(iii), $b > 0$.[7]

The principal's objective to offer a profit maximizing contract – *i.e.,* an optimal combination of a fixed payment and a bonus – is burdened with (i) moral hazard as the agent's effort is unobservable and (ii) a truth-telling problem as the principal has to credibly commit herself to a truthful revelation of his own signal.[8] In the next sections we will first analyze the pure moral hazard problem (*i.e.,* the case of non-binding truth-telling constraints) and then proceed with an analysis of the truthtelling problem.

3 Pure Moral Hazard Problem

In this section we abstract from the truthtelling problem inherent to the principal-agent relationship in order to analyze the isolated impact of moral hazard on the optimal effort level chosen by the principal and social welfare. Hence, we assume throughout this section that the contract $\Gamma = (f, b)$ guarantees truth-telling (*i.e.,* truth-telling constraints are non-binding).

Incentive Compatibility For a given contract $\Gamma = (f, b)$, the agent chooses effort p as to maximize his utility (see Eqn. 1) while anticipating the generation of ex-post conflict at level q^* as depicted in Lemma 1. This means, he maximizes

$$U(p) = p(\gamma_{HH} + \gamma_{HL})b + f - v(p) - p\gamma_{LH}(Y(1 - q^*) + c(q^*))$$

which induces the first order condition[9]

$$
\begin{aligned}
b(p) &= \frac{v'(p) + \gamma_{LH}(Y(1 - q^*) + c(q^*))}{\gamma_{HH} + \gamma_{HL}} \\
&= \frac{1}{g}(v'(p) + (1 - g)(1 - \rho)x(Y(1 - q^*) + c(q^*))).
\end{aligned}
\tag{3}
$$

Note that $\frac{d^2U(p)}{dp^2} = v''(p) > 0$ such that the agent's optimization problem is well-behaved. Eqn. (3) shows that the incentive compatible bonus that the principal pays to the agent in case he beliefs that the agent did a good job has to overcome marginal effort costs and marginal psychological costs. If the principal wants to induce a positive effort level, he has

---

[7] f can be interpreted as an up-front payment or a franchise fee with a payment of zero at stage 5 if the principal reports $t_P = L$ and a payment of b (a bonus) if she reports $t_P = H$.

[8] According to Lemma 2(ii), the agent's report does not affect his payoff such that we are save to assume a truthful revelation of his signal.

[9] We denote a bonus which implements an effort level of p by $b(p)$.

to offer a positive bonus. Note, however, that the required bonus does not vanish in the limit of small efforts, because marginal psychological costs do not vanish for $p = 0$. Finally, observe that the incentive compatible bonus increases in target effort p, psychological costs Y, and the conditional probability of conflict (γ_{LH}). In particular, a higher quality of the principal's signal g reduces the incentive compatible bonus because the agent expects higher returns to effort, the probability of conflict decreases, and the agent's psychological costs in case of conflict diminish. Likewise, a lower correlation of the signals or a higher probability of a positive independent evaluation by the agent enhances the compensation requested by the agent for a given effort level.

Individual Rationality   The agent accepts a contract $\Gamma = (f, b)$ whenever his expected utility from it is weakly positive, *i.e.*

$$p(\gamma_{HH} + \gamma_{HL})b + f - v(p) - p\gamma_{LH}(Y(1 - q^*) + c(q^*)) \geq 0.$$

To maximize her profits, the principal sets the upfront payment for a given bonus b to

$$f(b) = -p(\gamma_{HH} + \gamma_{HL})b + v(p) + p\gamma_{LH}(Y(1 - q^*) + c(q^*)).$$

Observe that the upfront-payment can well be negative (*i.e.*, a franchise fee) as the agent is not protected by limited liability. Note in particular that $f(b)$ can always be fixed such that the agent does not receive any rents from the relationship.

What are the principal's costs to implement an effort level $p > 0$ on the basis of these incentive compatibility and individual rationality constraints?

Effort Costs   To implement effort $p > 0$ the principal's costs are $C(p) = f + p(\gamma_{HH} + \gamma_{HL})b(p) = v(p) + p\gamma_{LH}((1 - q^*)Y + c(q^*))$. Note that $C(p)$ is convex and that $C(0) = 0$. We adopt the convention that an effort $p > 0$ which is not implementable requires infinite costs.

Optimal Effort   The principal's profit now reads

$$\Pi(p) = p\phi - p\gamma_{LH}q^*\psi - C(p)$$

which is zero for $p = 0$ and concave for $p > 0$. We denote the maximum of $\Pi(p)$ on $[0, 1]$ by $\widetilde{p}$ and the corresponding profit for the principal by $\tilde{\Pi}$[10] and derive the following set of results.

Proposition 1. *Pure Moral Hazard*

(i) $\widetilde{p} > 0$ *if and only if* $\phi > \underline{\phi} \equiv \gamma_{LH}(q^*\Psi + ((1 - q^*)Y + c(q^*)))$.

---

[10]$\widetilde{p}$ and $\tilde{\Pi}$ are equilibrium effort and profit whenever the truth-telling constraints are non-binding.

(ii) Suppose $\phi > \underline{\phi}$. Then, $\frac{d\tilde{p}}{d\phi} > 0$, $\frac{d\tilde{p}}{d\psi} < 0$, $\frac{d\tilde{p}}{dg} > 0$, $\frac{d\tilde{p}}{d\rho} > 0$, and $\frac{d\tilde{p}}{dx} < 0$.

(iii) Suppose $\phi > \underline{\phi}$. Then, $\frac{d\tilde{\Pi}}{d\phi} > 0$, $\frac{d\tilde{\Pi}}{d\psi} < 0$, $\frac{d\tilde{\Pi}}{dg} > 0$, $\frac{d\tilde{\Pi}}{d\rho} > 0$, and $\frac{d\tilde{\Pi}}{dx} < 0$.

Proof. See Appendix 8.4 □

Part (i) indicates that a relationship with positive effort level (*i.e.*, $\tilde{p} > 0$) can only be established if the value of the project exceeds the expected costs of retaliation for the principal and the expected compensation for the psychological costs of the agent. If a relationship is established because the value of the project is above this threshold, the comparative statics of the optimal effort level $\tilde{p}$ are straightforward. As indicated in Part (ii), an increase in the value of the project certainly enhances marginal benefits and thereby $\tilde{p}$. Likewise, higher costs of conflict for the principal enhance marginal costs and lower the optimal effort level. A higher quality of the principal's signal reduces the probability of conflict and psychological costs for the agent which reduces marginal costs and leads to higher optimal effort levels. A higher correlation of signals or a lower quality of an independent judgement have a similar effect as they also result in lower expected conflict levels and a lower compensation of psychological costs.

As indicated in Part (iii), these intuitive effects also carry over to the comparative statics of the principal's profit. The higher the value of the project and the lower expected costs associated with the retaliation of the agent, the more profit is awarded to the principal. In particular, the principal gains from a decrease in retaliation costs ψ, an increase in the principal's signal quality g (which reduces the probability of conflict and the agent's psychological costs), an increase in the signal correlation ρ and a decrease in the probability that the agent receives an independent signal x.

As the agent does not receive any rents in the optimal contract, the principal's profit also measures the surplus of the relationship. Hence, in the case of non-binding truth-telling constraints, conflicts (*i.e.* their likelihood γ_{LH} and size $q^*\Psi$) as well as the agent's psychological sensitivity Y only have a welfare detrimental effect. Therefore, any property of the information technology which reduces conflict (*i.e.* an increase in g or ρ) is welfare-enhancing, while an increase in the quality of the agent's independent judgment x induces the adverse effect.

In this section we have abstracted from the truthtelling problem, *i.e.* we have concentrated on the case of non-binding truthtelling constraints, to isolate the impact of moral hazard. In the following section, we analyze the robustness of these findings in the presence of truthtelling constraints.

4   Truth-Telling Problem

With a contract as characterized in Lemma 2(ii), the agent is indifferent between all possible reports as his payment (and also his psychological payoff) will be unaffected by his own

reporting decision. Hence, we can savely adopt the convention that the agent always tells the truth. This given, the principal's profit contingent on the agent's and her own report can be represented in the following table (with the principal's report depicted in the rows and the agent's report (and signal) depicted in the columns).

	H	L
H	$p\phi - f - b$	$p\phi - f - b$
L	$p\phi - f - q^*\psi$	$p\phi - f$

Suppose $s_P = H$. Then, the principal tells the truth, whenever his payoff from doing so (which reads $p\phi - f - b$) is larger than his payoff from reporting $t_P = L$ (which reads $p\phi - f - Pr(s_A = H \mid s_P = H)q^*\psi$). This means the principal reports $t_P = H$ if

$$b \leq \frac{\gamma_{HH}}{(\gamma_{HH} + \gamma_{HL})}q^*\psi = (\rho + (1-\rho)x)q^*\psi \equiv b^{max}. \tag{4}$$

The principal can only credibly promise a bonus b below b^{max}. Note that this upper bound to credible bonuses increases in the signal correlation ρ and in the quality of an independent judgment x. An increase in each of these parameters lowers the probability of the configuration $s_P = H$ and $s_A = L$ in which case the principal could cheat without facing retaliation and therefore reduces the incentive to save the bonus payment. Moreover, b^{max} certainly increases in the level of conflict $q^*\psi$ and thereby decreases in the quality of the principal's signal g. Intuitively, the agent anticipates that there is the less potential for conflict in the relation, the more competent the principal is in evaluating his job and performance. The less potential for conflict, however, the lower the maximum bonus the principal can offer without having an incentive to cheat ex-post.

If $s_P = L$, the principal tells the truth, whenever his payoff from doing so (which reads $p\phi - f - Pr(s_A = H \mid s_P = L)q^*\psi$) is larger than his payoff from reporting $t_P = H$ (which reads $p\phi - f - b$). Hence, the principal reports $t_P = L$ if

$$b \geq \frac{\gamma_{LH}}{(\gamma_{LH} + \gamma_{LL})}q^*\psi = \frac{(1-\rho)x}{(1-\rho x)}q^*\psi \equiv b^{min}. \tag{5}$$

The principal can also not promise to pay arbitrarily low bonuses as he has an incentive to evade conflict through 'unconditional bonuses'. By paying the bonus independently of his signal, the principal avoids any conflict with an agent who is prepared to protect his positive self-image. The minimal credible bonus is thereby decreasing in the signal correlation ρ and increasing in the quality of an independent judgment x because the larger ρ and the smaller x the smaller is the probability of the configuration $s_A = H$ and $s_P = L$ in which case the principal would benefit from conflict evasion. Similarly to b^{max}, an increase in g lowers the retaliation probability q^* and thereby b^{min}.

12

Note in particular that $b^{max} > b^{min} > 0$ and that the difference between b^{max} and b^{min} gets larger and the respective interval is shifted towards larger bonuses as q^* or ψ increases. Hence, the larger the potential conflict level, the higher are the bonuses that can be implemented. In fact, for every bonus b there is a conflict level ψ such that b is credible.[11] While elevated levels of conflict were only welfare detrimental in the pure moral hazard case, *i.e.* the case of non-binding truthtelling constraints (see Proposition 1), they relax the upper- and tighten the lower threshold of credible bonuses.

Implementable Efforts We call a certain effort level $p > 0$ *implementable* if $b(p) \in [b^{min}, b^{max}]$. Furthermore, we define the minimum implementable effort p^{min} and the maximum implementable effort p^{max} implicitly by $b^{min} = b(p^{min})$ and $b^{max} = b(p^{max})$.

Optimal Effort Level We denote the maximum of $\Pi(p) = p\phi - p\gamma_{LH}q^*\psi - C(p)$ on $\{0\} \cup [p^{min}, p^{max}]$ by p^*. p^* will be referred to as the optimal effort level (p^* is the optimal effort level for the principal given that only effort levels between p^{min} and p^{max} are feasible) and $\Pi^* = \Pi(p^*)$ will be the corresponding profit for the principal.

Proposition 2. *Optimal Effort Level*

$p^* > 0$ *if and only if* $\phi > \overline{\phi} > \underline{\phi}$ *with* $\Pi(p^{min})|_{\phi=\overline{\phi}} = 0$.
Now suppose that $\phi > \overline{\phi}$.

(i) *Binding Lower Truth-Telling Constraint: If* $0 < \tilde{p} < p^{min}$, *then the principal implements* $p^* = p^{min}$ *with bonus* b^{min} *[Figure 1].*

(ii) *Binding Upper Truth-Telling Constraint: If* $\tilde{p} > p^{max}$, *then the principal implements* $p^* = p^{max}$ *with bonus* b^{max} *[Figure 2].*

(iii) *Non-Binding Truth-Telling Constraint: If* $\tilde{p} \in [p^{min}, p^{max}]$, *then the principal implements* $p^* = \tilde{p}$ *by paying* $b(\tilde{p})$ *[Figure 3].*

Proof. See Appendix 8.5 □

According to Proposition 2, there will be no principal-agent relationship (*i.e.* $p^* = 0$) whenever the returns to the project are below a certain threshold. Note in particular, that the presence of a truth-telling problem increases the corresponding threshold value compared to the pure moral hazard case $(\overline{\phi} > \underline{\phi})$ which already indicates potential welfare losses due to truth-telling constraints. Finally, observe that in the absence of conflict (*i.e.*, $q^*\Psi = 0$) it holds that $(p^{min} = 0)$ such that profits for the principal are zero at p^{min} for any ϕ. This

---

[11]A comprehensive discussion of the comparative statics of b^{max} and b^{min} can be found in Appendix 8.3.

establishes the familiar result that no positive effort can be implemented in the absence of conflict if performance evaluations are subjective.

[Figures 1-3 here]

As already discussed $b^{max} > b^{min} > 0$. This implies that $p^{max} > p^{min} > 0$ (see Figures 1 to 3). If the value of the project is sufficiently large to establish a relationship one can distinguish between three cases (see Proposition 2): i) the case of a binding lower truth-telling constraint, ii) the case of a binding upper truth-telling constraint (see Figure 1 and 2, respectively) and iii) the case of a non-binding truth-telling constraint (see Figure 3). The comparative statics in the latter case have already been analyzed in Proposition 1 (as the principal simply implements $p^* = \tilde{p}$). The analysis of cases i) and ii) deserves some more attention. To this end, the following Lemma captures the comparative statics of p^{min} and p^{max} with respect to the level of conflict and the parameters of the information technology.

Lemma 3. *Truth-Telling Constraints*

(i) $\frac{dp^{min}}{d\Psi} > 0$ *and* $\frac{dp^{max}}{d\Psi} > 0$. *(ii)* $\frac{dp^{max}}{dg} > 0$ *and* $\frac{dp^{min}}{dg} > 0$ *for a given* Ψ *if* $\frac{dY}{dg}$ *is sufficiently small and* $\frac{dp^{max}}{dg} < 0$ *and* $\frac{dp^{min}}{dg} < 0$ *for a given* Y *if* Ψ *is sufficiently large.* *(iii)* $\frac{dp^{max}}{d\rho} > 0$ *and* $\frac{dp^{min}}{d\rho} < 0$ *if* ψ *is sufficiently large.* *(iv)* $\frac{dp^{max}}{dx} > 0$ *and* $\frac{dp^{min}}{dx} > 0$ *if* ψ *is sufficiently large.*

Proof. See Appendix 8.6 □

As the level of conflict ψ lifts the minimal credible bonus b^{min} and the maximal credible bonus b^{max}, while leaving the incentive compatible bonus $b(p)$ unaltered, p^{min} and p^{max} increase in ψ (Part (i)). Intuitively, the more conflict, the less tempting it is to cheat on the agent (upper truth-telling constraint) and the more tempting it is to evade conflict through unconditional bonus-payments (lower truth-telling constraint).

In contrast to this, the impact of the information technology on p^{min} and p^{max} is more subtle (see Parts (ii)-(iv)). All parameters of the information technology influence the minimal, the maximal and the incentive compatible bonus. For instance, a higher quality of the principal's signal g reduces the retaliation probability (recall that $\frac{dq^*}{dg} \leq 0$). This, in turn, lowers b^{min} and b^{max} and ceteris paribus lowers p^{min} and p^{max}. However, as we have seen in the previous section, a more accurate signal also reduces the incentive compatible bonus to implement a certain effort level as the agent faces a lower probability of conflict and lower psychological costs if the principal's signal is more precise. Ceteris paribus this effect increases p^{min} and p^{max}. Now, if the agent's psychological costs are sufficiently insensitive to a change in g (*i.e.* $\frac{dY}{dg}$ is sufficiently small), the retaliation probability in the case of conflicting signals does not change significantly while the probability of conflicting signals does such that the latter effect dominates the former. Likewise, for a given sensitivity to the signal quality
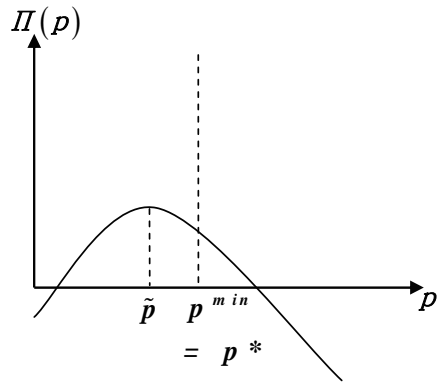
14

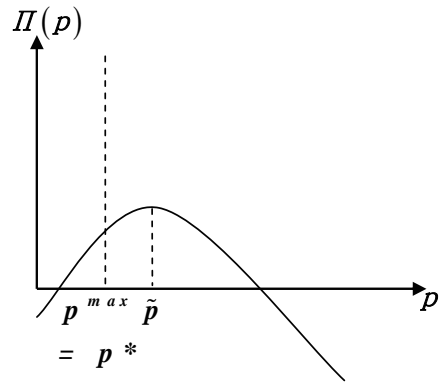Figure 1: Binding Lower Truth-Telling Constraint

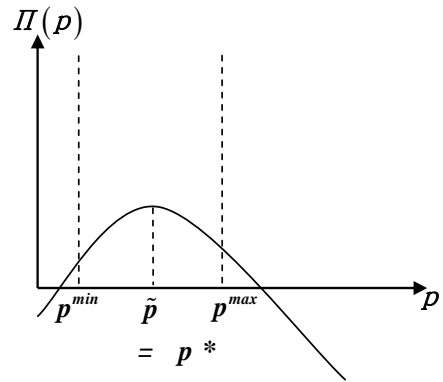Figure 2: Binding Upper Truth-Telling Constraint

Figure 3: Non-Binding Truth-Telling Constraint

($\frac{dY}{dg} < 0$), the former effect dominates the latter if Ψ is sufficiently large. Similarly to the impact of g, also ρ and x have a direct and an indirect effect on b^{min} and b^{max}. Both parameters modify the probability with which the principal could gain from a lie, but also change the expected psychological costs which have to be compensated by the incentive compatible bonus. Part (iii) and (iv) of Lemma 3 show that, if ψ is sufficiently large such that the gains from a lie are sufficiently pronounced, the former effect dominates the latter.

Welfare Analysis To analyze the comparative statics of the surplus (which is identical to the principal's profits in our set-up), observe that the impact of a parameter y on profits $\Pi(p)$ can be written as $\frac{d\Pi(p)}{dy} = \frac{\partial\Pi(p)}{\partial y} + \frac{\partial\Pi}{\partial p}\frac{dp}{dy}$. We will refer to the first term as the direct effect and the second as the indirect effect. The direct effect captures the impact of the parameter on profits *for a given effort level*. By the envelope theorem, this fully determines the comparative statics of equilibrium profits in the pure moral hazard case (as $\frac{\partial\Pi}{\partial p} = 0$ for $p = \tilde{p}$). The indirect effect captures the impact of the parameter on the *chosen effort level* and the resulting change in profit.

In the case of non-binding truthtelling constraints, *i.e.* case iii) in Proposition 2, the indirect effect vanishes and comparative statics are as depicted in Proposition 1(iii). For cases i) and ii) in Proposition 2, on the other hand, the indirect effect can no longer be neglected and may well dominate and reverse the comparative statics of the pure moral hazard case as demonstrated in the following result.

Proposition 3. *Comparative Statics of Welfare*

(i) There exists $\tilde{\phi}$ such that $\frac{d\Pi^*}{d\psi} > 0$ for all $\phi > \tilde{\phi}$. (ii) There exists ϕ, Y and $v(p)$ such that $\frac{d\Pi^*}{dg} < 0$ and there exists Ψ and $\tilde{\phi}$ such that $\frac{d\Pi^*}{dg} < 0$ for all $\phi > \tilde{\phi}$. (iii) There exists $\tilde{\phi}$ and $\tilde{\psi}$ such that $\frac{d\Pi^*}{dx} > 0$ if $\psi > \tilde{\psi}$ and $\phi > \tilde{\phi}$.

Proof. See Appendix 8.7. □

Proposition 3 indicates two different effects which may reverse the comparative statics of the pure moral hazard case. First, the upper truthtelling constraint may be binding. This is in particular the case for large project values ϕ which induce large marginal benefits and therefore require optimal effort levels beyond p^{max}. An increase in ψ or x is welfare detrimental for a given effort level (*i.e.* ($\frac{\partial\Pi}{\partial\psi}$) < 0) but also pushes p^{max} (as indicated in Lemma 3(iv), p^{max} is increasing in Ψ and increasing in x if Ψ is sufficiently large) and thereby relaxes the upper truth-telling constraint. In contrast, an increase in g is welfare enhancing for a given effort level (*i.e.* ($\frac{\partial\Pi}{\partial g}$) > 0) but also reduces p^{max} (as indicated in Lemma 3(ii), p^{max} is decreasing in g if Ψ is sufficiently large) and thereby tightens the upper truth-telling constraint. As indicated by Proposition 3(i), (ii), and (iii), the latter (indirect) effects indeed dominate the former (direct) effects if project values are sufficiently large. Hence, higher probabilities or

levels of conflict are welfare enhancing while better signals are welfare detrimental in the case of valuable projects for which the upper truth-telling constraint is tight.

Second, the lower truthtelling constraint may be binding. This is in particular the case for small project values which are sufficiently attractive to sign contracts on small positive effort levels but operate with bonus payments which tempt the principal to evade conflict by paying the bonus unconditional on the signal. In this case, the principal suffers from parameter changes which tighten the lower truthtelling constraint. For instance, the higher the quality of the principal's signal g, the larger p^{min} (if $\frac{dY}{dg}$ is sufficiently small) and the more tight the lower truthtelling constraint. In contrast, an increase in g enhances the principal's profit *for a given effort level*. According to Proposition 3(ii) the latter (direct) effect may well be dominated by the former (indirect) effect. As a consequence, a better signal for the principal may be welfare detrimental in the case of small projects for which the lower truth-telling constraint is tight.

Note that similar detrimental effects cannot be derived for the correlation of signals ρ, as a higher correlation directly enhances the principal's profit and relaxes the lower and the upper truthtelling constraint as long as ψ is sufficiently large (see Lemma 3(iii)).

Finally, we compare equilibrium profits with the first best solution and discuss the limit of a perfect signal to the principal, perfectly correlated signals, and no correct independent judgment of the agent.

Proposition 4. *First Best Comparison*

(i) Suppose $g < 1$, $\rho < 1$, and $x > 0$. Then, $\Pi(p^*) < \Pi(p_{FB})$. (ii) Let $\rho = 1$ and/or $x = 0$. Then, $p^* = p_{FB}$ and $\Pi(p^*) = \Pi(p_{FB})$ if and only if $\frac{\phi}{g} \leq \rho q^* \psi$. (iii) Let $g = 1$. Then, $p^* = p_{FB}$ and $\Pi(p^*) = \Pi(p_{FB})$ if and only if $\frac{(1-\rho)x}{1-\rho x} q^* \psi \leq \phi \leq (\rho + (1-\rho)x) q^* \psi$

Proof. See Appendix 8.8. □

Part i) indicates that an imperfect information technology of the principal together with an imperfect correlation of the principal's and the agent's signals and at least some correct independent judgment of the agent induces a welfare loss.

In Part ii) it can be seen that, if signals are perfectly correlated ($\rho = 1$) or the agent does not receive a correct signal, if he has to form an independent judgment ($x = 0$), a first best will be reached whenever the respective incentive compatible bonus $b(p_{FB}) = \frac{v'(p_{FB})}{g} = \frac{\phi}{g}$ is credible, *i.e.*, $b(p_{FB}) \leq b^{max}$. As the minimal credible bonus b^{min} is zero for $\rho = 1$ or $x = 0$, only the upper truth-telling constraint matters in this case and a first best will be established, if the project value is not too small compared to the expected costs of retaliation.

This changes if we consider the limit $g = 1$. Again, a first best is reached whenever the incentive compatible bonus $b(p_{FB}) = \phi$ is credible, *i.e.* $b(p_{FB})$ is between b^{min} and b^{max}.

However, as $b^{min} = \frac{(1-\rho)x}{1-\rho x}q^*\psi$ does not vanish as long as $\rho < 1$ and $x > 0$, the first best effort can be too large (as in Part (ii)) or too small to be implementable (see Part (iii)). Hence, it requires a 'fine-tuning' of ϕ (relative to expected costs of conflict) to guarantee a first best solution in this case.

5 Endogenous Evaluation Process

Until now, we have investigated optimal contract design and welfare implications of an *exogenously* given information technology. In reality, however, the principal often does not only decide upon the contractual arrangements such as bonuses or fixed payments. He may also decide upon the acquisition of information on the agent's performance. The principal can, for example, decide how much time he spends on supervising the agent in the accomplishment of the project. He could *(i)* sit next to the agent during the whole project, or *(ii)* close the door to his office and only have a glance at the result. Arguably, the quality of the signal g is expected to be better under the first evaluation procedure.[12] Note that we retain the assumption that ρ and x are exogenous as it seems unrealistic to assume that the agent's information acquisition about his own performance is at the discretion of the principal. We assume that the quality of the signal is costless. This assumption is made i) to simplify the analysis and ii) to show that even with costless monitoring the principal might not choose a perfect evaluation procedure.

Implementable Bonuses Recall from the previous sections that a bonus $b(p)$ which makes the effort choice of p incentive compatible only satisfies the upper and lower truthtelling constraint of the principal if $b(p) \in [b^{min}, b^{max}]$. As displayed in Eqn. 3, the incentive compatible bonus $b(p)$, is monotone decreasing in g with $\lim_{g\to 0} b(p) = \infty$ and $\lim_{g\to 1} = v'(p)$. Likewise, it follows from Eqn. 4 and 5 that b^{max} and b^{min} are (weakly) monotone decreasing in g (because $\frac{dq^*}{dg} \leq 0$) with $\lim_{g\to 0} b^{min} < \infty$ and $\lim_{g\to 1} b^{min} = \frac{(1-\rho)x}{1-\rho x}q^*(g=1)\psi > 0$, and $\lim_{g\to 0} b^{max} < \infty$ and $\lim_{g\to 1} b^{max} = (\rho + (1-\rho)x)q^*(g=1)\psi > 0$.

This allows to distinguish the following cases for the implementability of an effort level $p > 0$ (see also Figure 4).

Lemma 4. *Let $p > 0$. Then, one of the following cases holds:*

 (i) *Case 1. $b(p) > b^{max}$ for all g.*
   Then, p can not be implemented.

---

[12]Note that we explicitly avoid terms like *control* and *(dis)trust* here (as *e.g.* used in Falk & Kosfeld (2006) and Ellingsen & Johannesson (2008)). The choice of the quality of the evaluation procedure has an influence on how well the principal can observe an acceptable effort given that the project is a success. Therefore, the higher the quality of the principal's evaluation process, the higher the probability that the agent is rewarded in case of success. A higher quality is, hence, not regarded as negative by the agent.

(ii) *Case 2. $b(p) \leq b^{max}$ for some $g < 1$ but $b(p) > b^{max}$ for $g = 1$.*
   Then, $p > 0$ is implemented with the maximal g for which $b(p) = b^{max}$.

(iii) *Case 3. $b(p) \leq b^{max}$ for some $g < 1$ but $b(p) < b^{min}$ for $g = 1$.*
   Then, $p > 0$ is implemented with the maximal g for which $b(p) = b^{min}$.

(iv) *Case 4. $b(p) \in [b^{min}, b^{max}]$ for $g = 1$.*
   Then, $p > 0$ is implemented with $b(p) = v'(p)$ (at $g = 1$).

Proof. Obvious. □

[Figures 4 and 5 here]

Case 1 simply captures a configuration of the model where effort p cannot be implemented with any signal quality g because incentive compatible bonuses are too large to be credible. A special case of this situation is the absence of psychological payoffs ($Y = 0$) or the lack of retaliation opportunities ($\Psi = 0$). *Case 4* depicts the situation that the incentive compatible bonus is credible for signal quality $g = 1$. As effort costs $C(p)$ are decreasing in g (a better signal reduces the probability of conflict and expected psychological costs), the principal will always implement p with the largest possible signal quality.[13]

Figure 4 shows *Case 2* in which $b(p) \leq b^{max}$ for some $g < 1$ but not at $g = 1$. The optimal bonus and signal quality is denoted $\underline{b}^{max}$ and $\overline{g}$. Figure 5, on the other hand, shows *Case 3* in which $b(p) \leq b^{max}$ for some $g < 1$ but $b(p) < b^{min}$ for $g = 1$. In this case the optimal bonus and signal quality is respectively denoted by $\underline{b}^{min}$ and $\overline{g}$.

The following example shows that all these cases can occur in our model.

Example 1. *Let $c(q) = \frac{1}{1-q} - q - 1$, $Y = (2 - 1.9g)$, $\rho = 1/2$, $x = 1/2$, $\phi = 1/2$, and suppose that the principal wants to implement p_{FB}, i.e., $v'(p) = \phi$.*

- *Let $\psi = 1$. Then, $b(p) > b^{max}$ for all g (Case 1).*

- *Let $\psi = 10$. Then, $b(p) \leq b^{max}$ for some $g < 1$ and $b(p) > b^{max}$ for $g = 1$ (Case 2).*

- *Let $\psi = 100$. Then, $b(p) \leq b^{max}$ for some $g < 1$ and $b(p) < b^{min}$ for $g = 1$ (Case 3).*

- *Let $\psi = 20$. Then, $b(p) \in [b^{min}, b^{max}]$ for $g = 1$ (Case 4).*

As suggested by the example, *Case 1* (*Case 3*) will be the relevant description of implementability if the level of conflict ψ is sufficiently small (large) as the following result indicates.

---

[13]Recall that signal quality was assumed to be costless. Whenever costs of information acquisition are increasing in g there is an obvious tradeoff between decreasing effort costs $C(p)$ and increasing costs of quality.
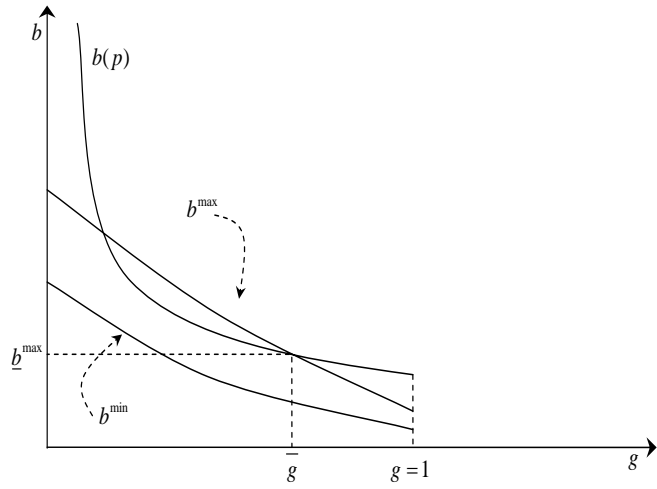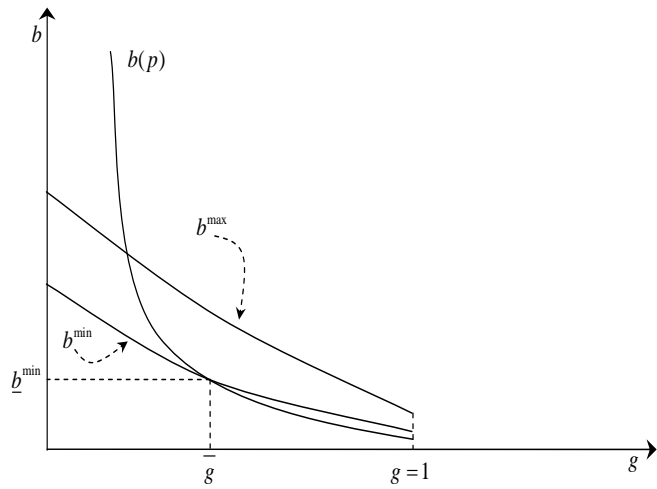
Figure 4: The Quality of the Evaluation Process: *Case 2.*

Figure 5: The Quality of the Evaluation Process: *Case 3.*

Lemma 5. *Suppose $p > 0$. Then, (i) there exists $\underline{\Psi} > 0$ such that Case 1 holds whenever $\psi < \underline{\Psi}$ and (iii) there exists $\overline{\Psi} > 0$ such that Case 3 holds whenever $\psi > \overline{\Psi}$ and $Y(g = 1) \neq 0$.*

Proof. The results follow directly from $\frac{db(p)}{d\psi} = 0$, $\lim_{g \to 0} b(p) = \infty$, and the fact that b^{min} and b^{max} are linear functions in Ψ with strictly positive slope ($\frac{db^{min}}{d\psi} = \frac{(1-\rho)x}{1-\rho x}q^* > 0$ and $\frac{db^{max}}{d\psi} = (\rho + (1-\rho)x)q^* > 0$) for $g \in (0,1)$. Moreover, b^{min} and b^{max} are strictly positive for $g \in (0,1)$ and are strictly positive for $g = 1$ whenever $Y(g = 1) > 0$. $\qquad\square$

Lemma 5 implies that the principal will not choose $g = 1$ to implement a certain effort level whenever ψ is too small or too large. If the conflict level is too low, the incentive compatible bonus may be too large to be credible. Perhaps a little bit more surprisingly, conflict levels can also be too large to implement a certain effort level with a perfect information technology even if the quality of the signal is costless. Intuitively, the principal has to maintain signal imperfections because incentive compatible bonuses decrease in signal quality (the more precise the signal, the smaller the necessary bonus for the implementation of a certain effort level) and these bonuses can be too small to be credible (the principal prefers to always pay the bonus because of conflict evasion).

Welfare Implications The previous paragraph demonstrated that additional discretion about the quality of her signal clearly increases the set of implementable efforts at the principal's disposal. However, certain effort levels still do not have to be implementable at any g (Case 1), or are not implementable at $g = 1$ (Cases 2 and 3). This holds in particular for p_{FB} which leads to the following welfare implications.

Proposition 5. *Suppose Case 1, 2 or 3 describes implementability of p_{FB}.[14] Then, $\Pi(p^*) < \Pi_{FB}$.*

Proof. Consider Case 1. As p_{FB} is not implementable, $\Pi(p^*) < \Pi_{FB}$ due to the unique optimality of p_{FB}.

Consider Case 2 and 3. Then p_{FB} can not be implemented with $g = 1$. Hence, for marginal costs of effort implementation it follows that $C'(p_{FB}) > v'(p_{FB})$ which implies $\tilde{p} < p_{FB}$ as $C(p)$ is convex and results in $\Pi^* < \Pi_{FB}$. $\qquad\square$

According to Proposition 5, the first best will not necessarily be implemented by the principal even if he can choose any signal quality at zero costs. As indicated by Lemma 5 this will be in particular the case if conflict level ψ is below a certain threshold such that the corresponding incentive compatible bonus is too large to be credible or above a certain threshold such that first best bonuses at $g = 1$ are too small to be credible. Hence, high

---

[14]Example 1 and Lemma 5 show that this holds true for an appropriate choice of ψ.

levels of conflict are suggested to be responsible for the endogenous choice of imperfect information technologies or the persistence of considerably subjective judgments in performance evaluation.

6   Concluding Remarks

The analysis of our model revealed that self-esteem and the individual's eagerness to protect it may facilitate principal-agent relationships even if performance signals are subjective and no third-party can enforce truth-telling. In particular, we analyzed the impact of the conflict level, the psychological sensitivity to ego-threats, and the quality of the information technology on optimal effort levels and social welfare.

Conflict Level   Conflict as modelled in this paper unambiguously reduces optimal effort levels and social welfare in the absence of truth-telling constraints. In the presence of truthtelling constraints, however, we demonstrate that some conflict potential is needed to establish a positive effort by the agent and that enhanced conflict levels have a positive effect on social welfare in the case of valuable projects which require substantial bonus payments to the agent. Hence, a well-functioning (internal or external) processing of appeals against managerial decision making is not only providing a more peaceful workforce, it may also implement the conflict level needed to make bonus payments credible and thereby raise firm profits.

Sensitivity to Ego-Threats   Higher levels of conflict unambiguously raise the maximum credible bonus and thereby relax the upper truthtelling constraint in a potentially welfare enhancing way. In contrast, the impact of psychological sensitivity to ego threats is more subtle. First of all, some sensitivity is needed to establish the prospect of conflict for the principal and thereby ensure truthtelling. The more aggressive the agent reacts to ego-threats, the higher the anticipated level of conflict and the less restrictive the upper truthtelling constraint. Hence, a more aggressive agent will induce a welfare improvement in case of valuable projects with associated high bonus payments as discussed above. However, the higher the sensitivity of the agent, the larger the required compensation for anticipated psychological costs. This ceteris paribus enhances necessary bonus payments for a given effort level and thereby reduces the principal's profit and social welfare. The ideal agent from the point of view of a principal who wishes to conduct a very valuable project is therefore someone who reacts very aggressively to ego-threats (*i.e.,* who has low costs of retaliation) but does not suffer to much from an ego-threat and the corresponding retaliation (*e.g.,* because q^* is large). This reinforces our above-made appraisal of appeal systems and suggests to ensure low costs of conflict creation for the employee (*e.g.,* low costs of law suits *etc.*). Note, however, that these recommendations only hold for very valuable projects which make the upper truthtelling

constraint binding. For non-binding truthtelling constraints, psychological sensitivity and the corresponding conflict remains detrimental to the principal's profits and welfare.

Information Technology   Moreover, we analyzed the impact of the information technology on optimal efforts and welfare. First of all, the principal is advised to use a signal technology which displays a perfectly correlated signal to her and the agent. With perfectly correlated signals the probability of conflicting signals is zero such that the agent does not expect any psychological costs. Moreover, the lower (upper) truthtelling constraint is decreasing (increasing) in the signal correlation such that the interval of credible bonuses is maximized for a given conflict level. Whenever the first best bonus is credible, perfectly correlated signals will allow the agent to implement a first best. This lends support to the practice of using information for performance evaluation which is not necessarily highly correlated with actual performance but ensures a high correlation with the agent's self-assessment. Similarly, the probability of conflict will be zero if the agent does not observe good performance independent of the principal. Hence, a first best can also be achieved with agents who lack an informative independent judgement (*i.e.,* $x = 0$). However, both truthtelling constraints are decreasing in x, such that implementability of the first best is less straightforward for $x = 0$ than for perfectly correlated signals.

The impact of the quality of the principal's signal has shown to be twofold. On the one hand, a better signal reduces necessary bonus payments (due to higher expected returns and lower psychological costs for the agent), on the other hand, a better signal reduces the psychological sensitivity and thereby yields a decrease in the level of conflict. If the agent's perception of an ego-threat is sufficiently sensitive to the signal quality of the principal or conflict levels are sufficiently large, a better signal will tighten the upper truthtelling constraint and therefore yield a welfare loss in the case of very valuable projects. If, however, the impact of the signal quality on psychological costs is rather weak, a better signal will mainly reduce agency costs and yield a welfare improvement – unless the lower truthtelling constraint binds, which may be the case for less valuable projects. Hence, the principal can only savely expect higher profits from employing a better information technology if project values are not too small or too large. As a consequence he will not always choose a perfect information technology even if this is costless. The optimal choice of an information technology rather deals with a tradeoff between agency costs (which are decreasing in the signal quality) and truthtelling constraints (which may well be tightened by a better information technology). Hence, imperfect information technologies as observed in reality may not only be optimal due to cost considerations but also due to the strategic aspects as discussed in this paper.

Discussion   We have decided to address the impact of ego-threats on principal-agent relationships in a rather simple model. The information technology is binary and never misiden-

tifies a bad outcome, agents are risk-neutral and not protected by limited liability, psychological costs are represented by an ad-hoc function of signal configurations, and team-effects are ignored. We have opted for these simplifications in order to provide a framework which allows for an easy identification of the relevant effects as discussed in the previous paragraphs. However, several extensions of our basic model may deserve some attention.

First of all, a certain robustness of our results can be expected for more general information technologies. The general impact of conflict and psychological sensitivity in the absence and presence of binding truthtelling constraints does not depend on the exact parametrization of the information technology but rather on the assumption that a tension between the principal's and the agent's signal creates conflict which induces truthtelling by the principal.

Second, we have chosen to model the agent as risk-neutral and with unlimited liability. While this obviously promoted expositional ease, it focuses on the special case of a principal-agent relationship which never leaves a rent to the agent. The presence of these rents clearly affects a welfare analysis. While different conflict levels only influence the principal's profit and truthtelling constraints (as analyzed in this paper), a change in the agent's sensitivity to ego-threats or the quality of the principal's signal does no longer leave the agent's profits unchanged. A proper welfare analysis under these circumstances, however, would ask for a comparison of (anticipated) psychological costs and material wage benefits. In contrast to our results which are not sensitive to assumptions in this respect (as long as the principal somehow manages to set the fix payment in such a way that the agent does not receive any rents from the relationship), a welfare analysis of rents and profits is. Meaningful results in this respect would require an empirical assessment of the anticipation of ego-threats and their present value which is not offered by the existing literature in social psychology and beyond the scope of the present paper.

Third, we opted for a simple and functional representation of self-esteem and ego-threats. In particular, we assumed that an agent retaliates for a certain signal configuration even though he anticipates that the principal truthfully reveals her private signal. A coherent justification of this assumption would either have to rely on a distinction between a cold-blood participation and effort selection decision versus emotional retaliation behavior or requires a modelling of a larger signal space (as *e.g.,* in [McLeod (2003)]) where the principal has an incentive to pool several different signals into one and the same bonus payment. Then, receiving a bonus payment which is also assigned to agents with lower performance again creates the tension analyzed in this paper. As a consequence, the wage compression effects as identified in [McLeod(2003)] would be mitigated by the principal's incentive to evade conflict. To go beyond the current model of signal dependent utility and to aim for a modelling of ego-threats with belief-dependent utility in a psychological game (comparable to recent models of reciprocity such as [Dufwenberg and Kirchsteiger (2003) or Falk and Fischbacher (2006)]

or procedural fairness [Sebald (2007)]) is certainly regarded as a valuable question for future theoretical and empirical research.

Finally, it is known since long [see Malcomson (1984)] that the problem of non-enforceable contracts in the presence of subjective performance measures is easily solved if the principal has to deal with a team of agents and can pay them according to a ranking with pre-committed payments for each rank. If agents do not suffer from psychological costs in these kind of tournaments, a first best can be achieved and performance pay as characterized in this paper is never superior. However, it is an empirical question whether tournaments actually lead to lower psychological costs. If self-esteem is threatened fiercely by the explicit announcement that someone-else is better, the principal may well face more conflict ex-post. This can lead to an inferiority of such a scheme and promote performance pay as discussed in our paper, where self-esteem is not threatened by a relative performance measure but by an absolute evaluation.

7 References

1. Albright, M. and Levy, P. (1995) *The effects of source credibility and performance rating discrepancy on reactions to multiple raters*, Journal of Applied Social Psychology, 25, 577600.

2. Baumeister, R. (2005), *Self-Concept, Self-esteem and Identity*, in Derlega, J., Jones, W. & Winstead, B. (2005), *Personality: Contemporary Theory and Research*, Wadsworth Publishing Company, 3rd Edition.

3. Baumeister, R., Smart, L. & Boden, J. (1996), *Relation of threatened egotism to violence and aggression: The dark side of high self-esteem*, Psychological Review, 103, 5-33.

4. Baird, L. (1977), *Self and Superior Ratings of Performance: As Related to Self-Esteem and Satisfaction with Supervision*, The Academy of Management Journal, 20(2), 291-300.

5. Bénabou, R. & Tirole, J. (2002), *Self-Confidence And Personal Motivation*, The Quarterly Journal of Economics, 117(3), 871-915.

6. Bureau of National Affairs (1981), *Wage and Salary Administration*, PPF Survey No. 131, Washington, D.C.

7. Bushman, B. & Baumeister, R. (1998), *Threatened Egotism, Narcissism, Self-Esteem, and Direct and Displaced Aggression: Does Self-Love or Self-Hate Lead to Violence?*, Journal of Personality and Social Psychology, 75(1), 219-229.

8. Compte, O. & Postlewaite, A. (2004), *Confidence-Enhanced Performance*, American Economic Review, 94(5), 1536-1557.

9. Dufwenberg, M. & Kirchsteiger, G. (2004), *A theory of sequential reciprocity*, Games and Economic Behavior, 47, 268-298.

10. Ellingsen, T. & Johannesson, M. (2008), *Pride and Prejudice: The Human Side of Incentive Theory*, American Economic Review, 98(3), 990-1008.

11. Falk, A. & Fischbacher, U. (2006), *A theory of reciprocity*, Games and Economic Behavior, 54(2), 293-315.

12. Greenwald, A. (1980), *The totalitarian ego: Fabrication and revision of personal history*, American Psychologist, 35, 603-618.

13. James, W. (1890), *The Principles of Psychology, Volumes 1 and 2*, Harvard University Press (1983).

14. Korman, A. (1969), *Toward a Hypothesis of Work Behavior*, Journal of Applied Psychology, 54, 31-41.

15. Köszegi, B. (2006), *Ego Utility, Overconfidence, and Task Choice*, Journal of the European Economic Association, 4(4), 673-707.

16. Levine, J. (2003), *Relational Incentive Contracts*, American Economic Review, 93(3), 835-857.

17. MacLeod, B. (2003) *Optimal Contracting with Subjective Evaluation*, American Economic Review, 93(1), 216-240.

18. Malcomson, J. (1984), *Work Incentives, Hierarchy, and Internal Labor Markets*, Journal of Political Economy, 92(3), 486-507.

19. Milkovich, G. & Wigdor, A. (1991), *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*, eds, National Academy Press, Washington, D.C.

20. Raskin, R., Novacek, J. & Hogan, R. (1991), *Narcissism, self-esteem, and defensive self-enhancement*, Journal of Personality, 59(1) 19-38.

21. Rabin, M. (1993), *Incorporating fairness into game theory and economics*, American Economic Review, 83, 12811302.

22. Roberson, Q. and Stewart, M. (2006), *Understanding the motivational effects of procedural and informational justice in feedback processes*, British Journal of Psychology, 97, 281298.

23. Sebald, A. (2007), *Procedural Concerns in Psychological Games*, ECORE Discussion Paper, 2007/62, Brussels, Belgium.

24. Shrauger, J. & Lund, A. (1975), *Self-evaluation and reactions to evaluations from others*, Journal of Personality. 43, 94-108.

25. Steelman, L. & Rutkowski, K. (2004), *Moderators of employee reactions to negative feedback*, Journal of Managerial Psychology, 19(1), 6-18.

8 Appendix

8.1 Information Technology

The conditional probabilities $\gamma_{k,l}$ for signal configuration $(s_P = k, s_A = l)$ read

$$\gamma_{HH} = g\left(\rho + (1-\rho)\,x\right) \text{ and } \gamma_{HL} = g(1-\rho)(1-x),$$

$$\gamma_{LL} = (1-g)\left(\rho + (1-\rho)\,(1-x)\right) \text{ and } \gamma_{LH} = (1-g)\,(1-\rho)\,x.$$

8.2 Proof of Lemma 2

To save on notation, we denote $Y(t_P = l, s_A = k, g)(1-q^*) - c(q^*) \equiv Y_{kl}$ throughout this proof.

Part(i). For a given contract Γ and signals s_P and s_A, the principal and the agent decide upon their report. Let $\sigma_P : S_P \to \Delta(S_P)$ and $\sigma_A : S_A \to \Delta(S_A)$ be the principal's and agent's reporting strategies (*i.e.*, mappings from the set of signals S_P and S_A to the set of probability distributions over S_P and S_A, respectively). Suppose that (σ_P^*, σ_A^*) is the pair of optimal reporting strategies for contract Γ. Then, the revelation principle implies that there exists a contract $\hat{\Gamma}$ which implements the same effort at the same costs and induces truthful reports by principal and agent. We will, henceforth, restrict our analysis to this type of (revelation) contracts.

Part (ii). Suppose that $\Gamma = \{w_{kl}\}$ is a revelation contract, *i.e.*, the principal and the agent tell the truth under contract Γ. As Γ implements $p > 0$, the incentive compatibility constraint

$$\Sigma_{k \in S_P, l \in S_A}(w_{kl} - Y_{kl})\frac{dPr\{s_P = k, s_A = l\}}{dp} = v'(p)$$

is satisfied. Consider a contract $\hat{\Gamma}$ which fixes payments of $\hat{w}_k = \sum_{l \in S_A} w_{kl} Pr\{s_P = k, s_A = l\}$ if the principal receives signal $s_P = k$, *i.e.*, payments are independent of s_A. These payments also satisfy the incentive compatibility constraint (see above).[15] Moreover, the agent weakly benefits from telling the truth. Finally, the principal's truth-telling constraint is also satisfied

---

[15]Individual rationality is trivially fulfilled as expected payments for the agent are the same under Γ and $\hat{\Gamma}$ and Γ is individually rational by assumption.

under $\hat{\Gamma}$. To see this observe that the principal reports k given that he has received k under contract Γ if

$$Pr\{s_A = H|s_P = k\}(w_{oH} - w_{kH}) + Pr\{s_A = L|s_P = k\}(w_{oL} - w_{kL}) \tag{6}$$
$$\geq Pr\{s_A = H|s_P = k\}((q^*\psi)_{kH} - (q^*\psi)_{oH}) + Pr\{s_A = L|s_P = k\}((q^*\psi)_{kL} - (q^*\psi)_{oL})$$

for all $o \in S_P$ (where $(q^*\psi)_{t_A,t_P}$ denotes the anticipated conflict costs for a reported configuration (t_A, t_P)). This set of inequalities holds because Γ implements truth-telling by assumption. $\hat{\Gamma}$ implements truth-telling if

$$\hat{w}_o - \hat{w}_k \geq Pr\{s_A = H|s_P = k\}((q^*\psi)_{kH} - (q^*\psi)_{oH}) + Pr\{s_A = L|s_P = k\}((q^*\psi)_{kL} - (q^*\psi)_{oL}).$$

holds for all $o, k \in S_P$. Inserting $\hat{w}_k$ and $\hat{w}_o$ yields

$$Pr\{s_A = H|s_P = k\}(w_{oH} - c_{kH}) + Pr\{s_A = L|s_P = k\}(w_{oL} - w_{kL})$$
$$\geq Pr\{s_A = H|s_P = k\}((q^*\psi)_{kH} - (q^*\psi)_{oH}) + Pr\{s_A = L|s_P = k\}((q^*\psi)_{kL} - (q^*\psi)_{oL}).$$

which coincides with Eqs. 6 and therefore shows that for $\hat{\Gamma}$ the principal's truthtelling constraint is satisfied as well. Hence, any revelation contract Γ can be substituted by a revelation contract $\hat{\Gamma}$ with w_{kl} independent of l which also implements $p > 0$ and leaves the principal weakly better off.

Part (iii). Suppose by contradiction that Γ implements $p > 0$ with $w_H = g$ and $w_L = g + \epsilon$ with $\epsilon \geq 0$. Then, the incentive compatibility constraint of the agent can be written as

$$\epsilon = \frac{v'(p) + \gamma_{LH}Y_{LH}}{(\gamma_{LH} + \gamma_{LL} - 1)}.$$

Observe that the numerator of the *rhs* is strictly positive and the denominator is strictly negative. Hence, the *rhs* is strictly negative and the incentive compatibility constraint is not satisfied for any $\epsilon \geq 0$. A contradiction.

8.3 Comparative Statics of Bonuses

$$b(p) = \frac{v'(p) + \gamma_{LH}(Y(1 - q^*) + c(q^*))}{\gamma_{HH} + \gamma_{HL}} = \frac{1}{g}(v'(p) + (1 - g)(1 - \rho)x(Y(1 - q^*) + c(q^*)))$$

and $\frac{dY}{dg} \leq 0$ imply the following results.

Lemma 6. *Comparative Statics of* $b(p)$
(i) Suppose $p > 0$. Then, $b(p) > 0$. (ii) $\lim_{p \to 0} b(p) > 0$. (iii) $\frac{db(p)}{dp} > 0$. (iv) $\frac{db(p)}{dg} < 0$. (v) $\frac{db(p)}{d\rho} < 0$. (vi) $\frac{db(p)}{dx} > 0$.

$$b^{max} = \frac{\gamma_{HH}}{(\gamma_{HH} + \gamma_{HL})} q^* \psi = (\rho + (1 - \rho)x)q^* \psi$$

$$b^{min} = \frac{\gamma_{LH}}{(\gamma_{LH} + \gamma_{LL})} q^* \psi = \frac{(1 - \rho)x}{(1 - \rho x)} q^* \psi$$

together with $\frac{dY}{dg} \leq 0$ imply the following results.

Lemma 7. *Comparative Statics of* b^{max} *and* b^{min}

(i) $b^{min} > 0$. (ii) $b^{max} > b^{min}$. (iii) $\Delta b \equiv b^{max} - b^{min}$ *is monotone increasing in* q^* *and* ψ. (iv) b^{min} *is monotone increasing in* q^*, ψ, *and* x *and monotone decreasing in* ρ *and* g. (v) b^{max} *is monotone increasing in* q^*, ψ, ρ *and* x *and monotone decreasing in* g.

Proof. (i) and (ii) follow from the positive correlation of signals, *i.e.,* $\rho > 0$ or $\gamma_{HH}\gamma_{LL} > \gamma_{HL}\gamma_{LH}$.

(iii) Follows from $\Delta b = \frac{\gamma_{HH}\gamma_{LL} - \gamma_{HL}\gamma_{LH}}{(\gamma_{HH}+\gamma_{HL})+(\gamma_{LH}+\gamma_{LL})} q^* \psi$.

(iv) and (v) follow directly from Eqs 4 and 5. $\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

8.4 Proof of Proposition 1

Part(i). Consider

$$\Pi(p) = p\phi - p\gamma_{LH}q^*\psi - C(p)$$

with $C(p) = v(p) + p\gamma_{LH}((1 - q^*)Y + c(q^*))$. Observe that $\Pi = ap - v(p)$ with $a = \phi - \gamma_{LH}(q^*\Psi + ((1 - q^*)Y + c(q^*)))$. Recall that $v(0) = 0$, $v'(0) = 0$, and $v''(p) > 0$. Then, $\widetilde{p} > 0$ if and only if $a > 0$.

Part (ii). We use the first order condition

$$\frac{d\Pi}{dp} = \phi - \gamma_{LH}q^*\psi - v'(p) - \gamma_{LH}(Y(1 - q^*) + c(q^*)) = 0. \tag{7}$$

as an implicit function of $\tilde{p}$. With we get

$$\frac{d\tilde{p}}{d\phi} = -\frac{1}{-v''(\tilde{p})} > 0,$$

$$\frac{d\tilde{p}}{d\psi} = -\frac{-\gamma_{LH}q^*}{-v''(\tilde{p})} < 0,$$

$$\frac{d\tilde{p}}{d\gamma_{LH}} = -\frac{-q^*\psi - (Y(1 - q^*) + c(q^*))}{-v''(\tilde{p})} < 0,$$

$$\frac{d\tilde{p}}{dY} = -\frac{-\gamma_{LH}\psi\frac{dq^*}{dY} - \gamma_{LH}(1 - q^*)}{-v''(\tilde{p})} < 0.$$

which implies Part (ii) (recall that $\frac{d\gamma_{LH}}{dg} < 0$, $\frac{d\gamma_{LH}}{d\rho} < 0$, $\frac{d\gamma_{LH}}{dx} > 0$, and $\frac{dY}{dg} < 0$).

Part (iii). Follows directly from

$$\frac{\partial \Pi(p)}{\partial \phi} = p > 0,$$

$$\frac{\partial \Pi(p)}{\partial \Psi} = -p\gamma_{LH}q^* < 0,$$

$$\frac{\partial \Pi(p)}{\partial g} = -p\frac{d\gamma_{LH}}{dg}(q^*\Psi + (Y(1-q^*) - c(q^*))) - p\gamma_{LH}\frac{dq^*}{dg}\Psi - p\gamma_{LH}(1-q^*)\frac{dY}{dg} > 0,$$

$$\frac{\partial \Pi(p)}{\partial \rho} = -p\frac{d\gamma_{LH}}{d\rho}(q^*\Psi + (Y(1-q^*) - c(q^*))) > 0,$$

$$\frac{\partial \Pi(p)}{\partial x} = -p\frac{d\gamma_{LH}}{dx}(q^*\Psi + (Y(1-q^*) - c(q^*))) < 0$$

for any $p > 0$ and the envelope theorem $\frac{d\tilde{\Pi}}{dy} = \frac{\partial \Pi}{\partial y}|_{p=\tilde{p}}$ for a parameter y.

8.5 Proof of Proposition 2

$''\Leftarrow''$ Suppose $\phi > \overline{\phi}$. As $\frac{\partial \Pi}{\partial \phi} > 0$, $\Pi(p^{min})|_\phi > 0 = \Pi(p = 0)$ and therefore $p^* > 0$. By Proposition 1, this implies that $\underline{\phi} \leq \overline{\phi}$.

$''\Rightarrow''$ Suppose $p^* > 0$. Then, $\phi > \underline{\phi}$ (see Proposition 1). Hence, $\Pi(p)$ is continuous in $p \geq 0$ and concave with a unique maximum at $\tilde{p} > 0$. Now suppose that $\phi < \overline{\phi}$ such that $\Pi(p^{min})|_\phi < 0$. Then, $\tilde{p} < p^{min}$ and $\Pi(p) < 0$ for all $p \in [p^{min}, p^{max}]$. A contradiction.

To see that $\overline{\phi} \neq \underline{\phi}$, recall from Lemma 7 that $b^{min} > 0$ which implies $p^{min} > 0$. Now suppose that $\phi = \overline{\phi} = \underline{\phi}$. Then, $\Pi(p^{min}) = 0$ by definition of $\overline{\Phi}$. Then, continuity and concavity of $\Pi(p)$ imply $0 < \tilde{p} < p^{min}$ where the first inequality contradicts Proposition 1(i).

Items (i) to (iii) are a direct implication of the fact that $\Pi(p)$ is continuous in $p \geq 0$ and concave with a unique maximum at $\tilde{p} > 0$ whenever $\phi > \overline{\phi}$, and the observation that $p^{max} > p^{min} > 0$.

8.6 Proof of Lemma 3

p^{min} is implicitly given by

$$b^{min} = \frac{(1-\rho)x}{1-\rho x}q^*\psi = \frac{1}{g}(v'(p^{min}) + (1-g)(1-\rho)x(Y(1-q^*) + c(q^*))) = b(p^{min})$$

and p^{max} is implicitly given by

$$b^{max} = (\rho + (1-\rho)x)q^*\psi = \frac{1}{g}(v'(p^{max}) + (1-g)(1-\rho)x(Y(1-q^*) + c(q^*))) = b(p^{max}).$$

We use these equations to compute the comparative statics of p^{min} and p^{max}. To be specific, let $F^{min} = b^{min} - b(p)$ and $F^{max} = b^{max} - b(p)$. Then, for a parameter y, $\frac{dp^{min/max}}{dy} = -\frac{\partial F^{min/max}/\partial y}{\partial F^{min/max}/\partial p}$. Note that $\frac{\partial F^{min/max}}{\partial p} = -\frac{v''(p)}{g} < 0$.

Part (i). Follows from $\frac{\partial b^{min}-b(p)}{\partial \psi} = \frac{(1-\rho)x}{1-\rho x}q^* > 0$ and $\frac{\partial b^{max}-b(p)}{\partial \psi} = (\rho + (1-\rho)x)q^* > 0$.

Part (ii). Observe that $\frac{\partial b^{min}-b(p)}{\partial g} = \frac{(1-\rho)x}{1-\rho x}\frac{dq^*}{dg}\psi + \frac{b(p)}{g} + \frac{1}{g}((1-\rho)x(Y(1-q^*)+c(q^*))) - (1-g)(1-\rho)(1-q^*)x\frac{dY}{dg}$ and $\frac{\partial b^{max}-b(p)}{\partial g} = (\rho + (1-\rho)x)\frac{dq^*}{dg}\psi + \frac{b(p)}{g} + \frac{1}{g}((1-\rho)x(Y(1-q^*)+c(q^*))) - (1-g)(1-\rho)(1-q^*)x\frac{dY}{dg}$. Both expressions are positive whenever $\frac{dq^*}{dg} \leq 0$ is sufficiently small (*e.g.*, if $\frac{dY}{dg} = 0$). For a given $\frac{dY}{dg} < 0$, however, there is always a conflict level Ψ such that both expressions are negative.

Part (iii). $\frac{\partial b^{min}-b(p)}{\partial \rho} = \frac{-x(1-x)}{(1-\rho x)^2}q^*\psi + \frac{1}{g}((1-g)x(Y(1-q^*)+c(q^*)))$ is negative if ψ is sufficiently large (for a given $0 < x < 1$) and positive if $x = 1$. Moreover, $\frac{\partial b^{max}-b(p)}{\partial \rho} = (1-x)q^*\psi + \frac{1}{g}((1-g)x(Y(1-q^*)+c(q^*))) > 0$.

Part (iv). $\frac{\partial b^{min}-b(p)}{\partial x} = \frac{1-\rho}{(1-\rho x)^2}q^*\psi - \frac{1}{g}((1-g)(1-\rho)(Y(1-q^*)+c(q^*)))$ is positive if ψ is sufficiently large or $g = 1$. $\frac{\partial b^{max}-b(p)}{\partial x} = (1-\rho)q^*\psi - \frac{1}{g}((1-g)(1-\rho)(Y(1-q^*)+c(q^*)))$ is positive if ψ is sufficiently large or $g = 1$.

8.7 Proof of Proposition 3

The impact of a parameter y on equilibrium profits $\Pi(p^*)$ can be denoted by $\frac{d\Pi(p^*)}{dy} = \frac{\partial \Pi(p^*)}{\partial y} + \frac{\partial \Pi(p^*)}{\partial p}\frac{dp^*}{dy}$. For $\frac{\partial \Pi(p)}{\partial y}$ see the proof of Proposition 1(iii). $\frac{\partial \Pi(p)}{\partial p} = \Phi - \gamma_{LH}q^*\psi - v'(p) - \gamma_{LH}(Y(1-q^*)+c(q^*))$. Note that for a fixed p, $\frac{\partial \Pi(p)}{\partial p}$ is a linear increasing function of ϕ and for a fixed ϕ it is a decreasing function of p with slope $-v''(p)$.

Part (i). Recall from Lemma 3(i) that $\frac{dp^{max}}{d\psi} > 0$. Fix any $p^{max} \in (0,1)$. Then, there exists a project value ϕ' such that $\frac{\partial \Pi(p)}{\partial p}|_{p=p^{max}} > 0$ and $p^* = p^{max}$ for all $\phi > \phi'$. In particular, $\frac{\partial \Pi(p)}{\partial p}|_{p=p^*}\frac{dp^*}{d\psi} > 0$. As $\frac{dp^{max}}{d\psi}$ and $\frac{\partial \Pi(p)}{\partial \psi}$ are independent of ϕ and $\frac{\partial \Pi(p)}{\partial p}$ is a linear increasing function of ϕ, there exists a ϕ'' such that $\frac{d\Pi(p^*)}{d\psi} > 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

Part (ii). Fix any $p^{min} \in (0,1)$ and a positive real number z. Then, there exists an effort cost function $v(p)$ such that $\frac{v'(p^{min})}{v''(p^{min})} > z$ and there exists a project value ϕ such that $0 < \tilde{p} < p^{min}$ and $\Pi(p^{min}) > 0$ (and therefore $p^* = p^{min}$). Now suppose that $\frac{dY}{dg} = 0$. Then, by Lemma 3(ii), $\frac{dp^{min}}{dg} > 0$ and $\frac{\partial \Pi(p)}{\partial g} = -p\frac{d\gamma_{LH}}{dg}(q^*\Psi + (Y(1-q^*)+c(q^*))) > 0$ (see the proof of Proposition 1). Now observe that $\frac{\partial \Pi(p)}{\partial g}$ is independent of $v(p)$ and its derivatives while $\frac{\partial \Pi}{\partial p}\frac{\partial p^{min}}{\partial g}$ is increasing in $\frac{v'(p)}{v''(p)}$. Hence, $\frac{d\Pi(p)}{dg} < 0$ if z is sufficiently large.

For the second part recall from Lemma 3(i) that $\frac{dp^{max}}{dg} < 0$ if Ψ is sufficiently large. Fix any $p^{max} \in (0,1)$. Then, there exists a project value ϕ' such that $\frac{\partial \Pi(p)}{\partial p}|_{p=p^{max}} > 0$ and $p^* = p^{max}$ for all $\phi > \phi'$. In particular, $\frac{\partial \Pi(p)}{\partial p}|_{p=p^*}\frac{dp^*}{dg} < 0$. As $\frac{dp^{max}}{dg}$ and $\frac{\partial \Pi(p)}{\partial g}$ are independent of ϕ and $\frac{\partial \Pi(p)}{\partial p}$ is a linear increasing function of ϕ, there exists a ϕ'' such that $\frac{d\Pi(p^*)}{dg} < 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

Part (iii). Recall from Lemma 3(iv) that there exists a $\tilde{\psi}$ such that $\frac{dp^{max}}{dx} > 0$ for all $\psi > \tilde{\psi}$. Fix any $p^{max} \in (0,1)$ with such a ψ. Then, there exists a project value ϕ' such that $\frac{\partial \Pi(p)}{\partial p}|_{p=p^{max}} > 0$ and $p^* = p^{max}$ for all $\phi > \phi'$. In particular, $\frac{\partial \Pi(p)}{\partial p}|_{p=p^*}\frac{dp^*}{dx} > 0$. As $\frac{dp^{max}}{dx}$ and $\frac{\partial \Pi(p)}{\partial x}$ are independent of ϕ and $\frac{\partial \Pi(p)}{\partial p}$ is a linear increasing function of ϕ, there exists a

ϕ'' such that $\frac{d\Pi(p^*)}{dx} > 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

8.8   Proof of Proposition 4

Part (i). Follows from $C(p) > v(p)$ for every $p > 0$.

Part (ii) and (iii). $g = 1$, $\rho = 1$, or $x = 0$ implies that $\gamma_{LH} = 0$ and therefore $\Pi(p) = p\phi - v(p)$ such that $\tilde{p} = p_{FB}$. However, $b(p_{FB})$ has to be in the interval $[b^{min}, b^{max}]$ which results in the condition displayed in the proposition (recall that for $x = 0$ or $\rho = 1$, $b^{min} = 0$).