Personality and conflict in principal-agent relations based on subjective performance evaluations*

Mie la Cour Sonne† Alexander Sebald‡

This Version: November 8, 2012

Abstract

We analyze the role of conflict in principal-agent environments with subjective performance evaluations, reciprocal agents and endogenous feelings of entitlements. By explicitly modeling conflict as the reciprocal reaction of agents that feel unkindly treated, we reveal intriguing welfare effects associated with the agents' personality and provide a rational for the widespread use of personality tests in recruitment and promotion processes. Finally, we extend our framework to allow principals to choose their evaluation procedure and show, for example, that even if it is costless to choose a high quality evaluation procedure, principals might not find it optimal to do so.

Keywords: Subjective performance evaluations; Reciprocity; Procedural Concerns.

JEL-Classifications: D01; D02; D82; D86; J41.

*We would like to thank Bentley MacLeod, Christian Schultz and seminar/workshop participants at the University of Copenhagen, Columbia University as well as Copenhagen Business School for helpful comments and suggestions. All errors are our own.

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353, Copenhagen K, Denmark. Phone: (+45) 3532-3993. Fax: (+45) 3532-3064. E-mail: mie.la.cour.sonne@econ.ku.dk.

‡Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353, Copenhagen K, Denmark. Phone: (+45) 3532-4418. Fax: (+45) 3532-3064. E-mail: alexander.sebald@econ.ku.dk.

1 Introduction

Evaluating performance and linking rewards such as bonuses and promotions to subjective performance appraisals is an integral and important part of many of today's work relations [see e.g. Bushman *et al.* (1996), Ittner *et al.* (1997), Ittner *et al.* (2003), Murphy and Oyer (2001), Gibbs *et al.* (2004)]. To capture performance in a purely objective way is very costly and often hard to accomplish, since a lot of valuable information about performance is captured by subjective impressions rather than objective measures. As a result, it is often preferred to leave (at least part of the) performance feedback to more holistic subjective appraisals.

However, principal-agent relations involving ex-post asymmetric information in the form of non-verifiable subjective performance evaluations are fragile and prone to conflict. If labor contracts specify rewards on the basis of subjective appraisals, principals have an incentive to claim that performance was poor according to their perception to establish low wages (i.e. ex-post hold-up [see e.g. Macleod (2000)]). In addition, agents might feel shortchanged and create conflict when they receive a performance appraisal and reward from their principal which is lower than what they feel entitled to on the basis of their own subjective performance assessment [see e.g. Sebald and Walzl (2012b)].

In this paper we theoretically analyze the impact and importance of conflict created by ex-post asymmetric information and hold-up in principal-agent environments based on non-verifiable subjective performance evaluations. We investigate factors that mitigate this conflict and describe implications for optimal recruitment policies and the principal's choice of evaluation procedure.

The existing literature analyzing the problem of ex-post asymmetric information and hold-up in principal-agent relations has already highlighted the need to break budget balance through 'money burning' or 'third-party payments' to establish mutual beneficial relations [see e.g. Levin (2003), MacLeod (2003) and Fuchs (2007)]. As shown in this literature: letting principals choose the optimal degree of money burning or third party payments mitigates the potential truthtelling problems inherent in these strategic environments.

Different to this, we characterize and analyze a principal-agent model with non-verifiable subjective performance evaluations in which we do not model conflict as e.g. third-party payments optimally chosen to ensure truthtelling. Instead, we explicitly formalize conflict as the reciprocal reaction of agents that feel shortchanged and unkindly treated by their principal. In our setting a principal decides upon undertaking a project. The project requires effort of an agent. However, as the project is a complex good or service and its success is non-verifiable, incentive contracts contingent on the successful completion of the project are not feasible. Contracts can only be based on non-verifiable subjective performance evaluations.

We assume that whenever the principal and agent voluntarily agree on a contract before the agent invests effort into the project, the contract shapes the agent's feeling of entitlement which defines the wage she feels entitled to ex-post (see Hart and Moore (2008)). This feeling of entitlement constitutes a benchmark or reference point against which she judges the kindness of the principal's action, i.e. his performance feedback. Whenever the agent receives a performance feedback and, hence, an associated payment that lies below her feeling of entitlement she reacts reciprocally by creating costly conflict.[1]

By now it is a well established theoretical and empirical finding that reciprocity is an important motivational driving force mitigating moral hazard in principal-agent relations [see e.g. Fehr *et al.* (1993), Fehr *et al.* (1997), Charness (2004), Kube *et al.* (2011)]. However, the existing literature has abstracted from employment relations containing ex-post asymmetric information and hold-up. Analyzing the role of reciprocity in such strategic environments thus requires us to go beyond existing conceptualizations [e.g. Rabin (1993), Dufwenberg and Kirchsteiger (2004) and Hart and Moore (2008)].

Two important dimensions in which our model differs from the existing

[1] Note that this type of reciprocity is different from the payoff independent form of reciprocity analyzed in Sebald and Walzl (2012a) in which it is assumed that the agent's feeling of entitlement is solely shaped by his own performance evaluation independent of the payments specified in the contract.

literature on reciprocity are:

First, the agent's feeling of entitlement in our setting is not exogenous, but is endogenously shaped by the agent's own performance assessment. If the agent believes she did a good job, she feels entitled to a higher wage, than if she believes she did a bad job. Quite intuitively, the more effort she puts into the project, the more likely it is that she receives a good performance signal and, hence, the more likely it is that she feels entitled to a higher wage. As a consequence, the reciprocal reaction of the agent depends on the agent's own subjective evaluation of her performance.[2]

Second, in line with findings in the psychological literature we assume that the extent to which the agent feels entitled to a reward also depends on the principal's expertise and familiarity with the agent's task [see e.g. Landy and Murphy (1978), Ilgen and Taylor (1979), Greenberg (1986a) and Greenberg (1986b)]. Specifically, we assume that in case the agent is uncertain about the success of the project, she feels less shortchanged by a low performance evaluation and reward by the principal the greater his familiarity with her task and the greater his expertise in evaluating the success of the project.

Explicitly modeling conflict as originating from the reciprocal reaction of an agent that feels shortchanged and unkindly treated uncovers intriguing welfare effects.

First, we demonstrate that an increase in the principal's cost of conflict can actually enhance welfare if the project is sufficiently valuable. The intuitive explanation is that a higher level of conflict helps the principal to truthfully commit to a higher wage. This, in turn, helps the principal to achieve a higher effort level from the agent and a higher expected profit.

Second, we find that it might be optimal for the principal to hire an agent

[2] Note, this concept is closely related to an idea put forward in Carmichael and MacLeod (2003) in which it is analyzed how caring about sunk costs can help agents achieve efficient investments in a team production environment in which agents bargain about the division of the surplus only after they have made their investment decisions. Also in their setting with symmetric information agents' feelings of entitlement in the ex-post bargaining stage depend on their ex-ante investments, i.e. feelings of entitlement are endogenous.

with a high emotional sensitivity to reciprocity. A high emotional sensitivity to reciprocity on the side of the agent expands the range of effort level the principal can truthfully commit to, which implies a higher expected profit.

Third, it may be optimal for the principal to hire an agent for whom the likelihood of having an own opinion is minimized and, lastly, the principal might find it optimal to hire an agent who has a high probability of identifying a successful project in case she forms an independent judgement.

Clearly, these findings relate to and complement Prendergast (1993)'s theory of 'Yes Men', i.e. agents that never form an own judgement concerning their performance and always agree with their principals' opinions. Prendergast (1993) analyzes the incentive that agents have to conform to their principals' opinions and the inefficiencies that this behavior creates. He concludes by mentioning that an important incompleteness of his analysis lies in the fact that it does not 'address why managers may wish to have cronies who agree with them' Prendergast (1993, p. 770). Our analysis addresses this issue by clearly characterizing the circumstances under which principals have an incentive to hire 'Yes Men'.

Interestingly, the last three results regarding the agent's 'characteristics' closely link to and extend a fairly recent discussion in the economics literature and a long standing debate in the human resource/organizational behavior literature concerning the importance and effectiveness of 'applicant screenings' and 'personality tests' in recruitment and promotion processes.

The recent economics literature highlights the importance of screening to identify applicants with e.g. high 'work ethics' [see Bartling *et al.* (2012) and Huang and Cappelli (2010)] which is shown to be associated with e.g. a lower need to control and higher employee productivity.

The human resource and organizational behavior literature, on the other hand, stresses that personality tests are used by firms to identify applicants whose personal traits (e.g. the applicant's openness, determination and ability to cope with hierarchies and feedback) fit best to the 'culture' of the organization and the 'character' of the vacancy [see e.g. Raymark *et al.* (1997), Kristof

(1996), Cable and Judge (1994), Judge and Cable (1997) and Li *et al.* (2008)]. According to this literature, the 'person-organization' and 'person-job' fit are very important for the performance of employees and success of companies [e.g. Barrick and Mount (1991), Tett *et al.* (1999), Chatman *et al.* (1999) and Tett and Christiansen (2007)]. The culture of an organization and the character of a vacancy are e.g. determined by the nature of the industry, the character of the projects the organization is involved in and the technologies it uses [e.g. Schein (2004)].

In line with this, our results also highlight that it is vital for the performance and success of firms operating in complex environments preventing the specification of complete contracts to employ agents whose personal traits are in line with the culture of the organization they work for and the character of the job they perform.

Finally, we extend our framework to allow the principal to choose the evaluation procedure. More precisely, we allow the principal to choose the quality of the process used to evaluate the performance of the agent. In reality, the principal often does not only decide upon the contractual arrangements such as bonuses or fixed payments. He also decides upon the acquisition of information on the agent's performance.

It has been suggested in recent experimental and theoretical works that such procedural choices are important in strategic interactions with reciprocal agents [see e.g. Blount (1995), Sebald (2010), Aldashev *et al.* (2010)]. According to this literature procedural choices are important as they influence agents' kindness perceptions. In our setting, procedural choices influence the agent's feeling of entitlement and the a priori probability of conflict which, in turn, influence the agent's reaction to a particular feedback and the 'price' that the principal has to pay to implement a specific effort level.

Interestingly, we show that even if it is costless for the principal to choose a high quality evaluation procedure, he might not always find it optimal to do so. Signal imperfections and, thus, potential conflict might be necessary to credibly implement the principal's preferred effort level. This highlights that the choice of the evaluation procedure constitutes an integral and important part of strategic environments in which agents are motivated by reciprocity.

The organization of our analysis is as follows. In section 2, we present our principal-agent environment in which the agent behaves reciprocal and performance can only be measured subjectively. In section 3, we characterize the agent's optimal effort and conflict level, the principal's truthtelling limits, his optimal choice of effort and the associated implications for welfare. The impact of the principal's procedural choice is analyzed in section 4, followed by a conclusion.

2 The model

Consider a principal P who decides upon undertaking a project which might generate a profit ϕ if successful. The project requires effort of an agent A. If the agent invests effort $\tau \in [0,1]$ the expected profit of the project is $\tau\phi$. The project is a complex good or service and its success is non verifiable.

The Information Technology. The agent's effort is unobservable and, as a result, the principal and the agent are left to subjectively judge the success of the project. That is, the principal and the agent receive private non-verifiable performance signals $s_P \in S_P$ and $s_A \in S_A$ with $S_A = S_P = \{H, L\}$ respectively. These signals are informative with respect to the success of the project. If the project is not successful, the principal and the agent receive the signal $s_P = s_A = L$. On the other hand, if the project is successful, the principal receives the signal $s_P = H$ with probability g, the agent receives the same evaluation as the principal with probability ρ and receives $s_A = H$ as an independent signal with probability x. Hence, g indicates the quality of the principal's signal, $(1-\rho)$ measures the likelihood with which the agent has an own independent opinion and x quantifies the quality of the agent's signal if he forms an independent judgment [Note, this specification of the information technology coincides with

(MacLeod, 2003, p. 228)].[3]

We denote the probability that the principal receives signal k while the agent receives signal l given that the project is a success by γ_{kl}. More specifically, $\gamma_{HH} = g(\rho + (1-\rho)x)$, $\gamma_{HL} = g(1-\rho)(1-x)$, $\gamma_{LL} = (1-g)(\rho + (1-\rho)(1-x))$ and $\gamma_{LH} = (1-g)(1-\rho)x$.

In this principal-agent environment contracts contingent on the generation of ϕ are not feasible. Instead, a contract Γ specifies payments ω contingent on verifiable events, i.e. $\Gamma = \{\omega_{kl} | k \in S_P, l \in S_A\}$ where k and l respectively are the principal's and the agent's report concerning their subjective performance evaluations. The agent accepts a contract if it is individually rational. We normalize the agent's outside option to zero and as a consequence, the agent accepts a contract if whenever expected utility is weakly positive. The agent then chooses τ so as to maximize her utility. In this case we say that Γ *implements* τ. The principal and the agent report their signal truthfully if and only if they weakly benefit from doing so.

The Agent. We assume that the agent is risk neutral and not only motivated by her material payoffs, but also by reciprocity. More specifically, the agent's utility function is

$$U = \omega - v(\tau) - \theta \cdot \max\{\tilde{\omega} - \omega, 0\} \cdot (1-q) - c(q). \tag{1}$$

where ω is the agent's wage, $v(\tau)$ is the effort cost at effort level τ with $v'(0) = 0$, $v'' > 0$ and $\lim_{\tau \to 1} v(\tau) = \infty$ and $\theta > 0$ is the agent's sensitivity to reciprocity. The agent acts reciprocally whenever her wage ω is below her feeling of entitlement $\tilde{\omega}$, i.e. whenever $(\tilde{\omega} - \omega) > 0$. The reciprocal action consists of creating a conflict q costly to the principal. This conflict could be interpreted as a law suit, stealing from the work place, creating rumors which could hurt the firm's reputation etc. As with effort, also conflict is costly to the agent. A conflict level of q incurs a cost $c(q) \geq 0$ with $c(0) = 0$, $c'(0) = 0$,

[3] We restrict ourself to a binary signal for expositional ease. The extension to a finer signal structure can be done at a notational cost.

$c'' > 0$ and $\lim_{q \to 1} c(q) = \infty$.

The Principal. In contrast to the agent, we assume that the principal only cares about his profit. His expected profit is given by:

$$\Pi = \tau\phi - E\{\omega\} - E\{q\}\psi \tag{2}$$

where $E\{\omega\}$ and $E\{q\}\psi$ are the principal's expected wage costs and expected costs of conflict respectively. The parameter ψ captures the principal's 'sensitivity' to conflict or the agent's ability to impose costs on the principal by causing conflict. Alternatively, as our assumptions on $c(q)$ ensure that $q \in [0, 1]$, one can also interpret q as the probability with which the agent creates costs of ψ for the principal.

Contracts. Quite naturally cost-minimizing revelation contracts in our environment have the following basic characteristics

Lemma 1. Suppose there exists a contract Γ which implements $\tau > 0$. Then, there always exists a contract $\hat{\Gamma}$ implementing τ at weakly lower costs which has the following characteristics:

(i) the principal and agent tell the truth,

(ii) wage payments only depend on the principal's report, i.e. $\omega_{kl} = \omega_{km} \equiv \omega_k$ for all $k \in S_P$ and $l, m \in S_A$ and

(iii) wage payments are higher in case the principal reports H than if he reports L, i.e. $\omega_H > \omega_L$.

Proof: Appendix A.1

Since signals are private and non-verifiable, the contract cannot be made contingent on the principal's signal. Instead, the optimal contract depends on the principal's report of his signal, i.e. it depends on the subjective performance evaluations of the principal. Furthermore, the wage payment associated with a good report H has to be strictly higher than the wage payment following a bad review L. We say that the agent is being paid a wage of ω_L if the principal reports L and is being paid $\omega_H > \omega_L$ if the principal reports H.

Feelings of Entitlement. Different feelings of entitlement have been suggested in the literature conceptualizing and analyzing the influence of reciprocity in strategic environments. Rabin (1993) and Dufwenberg and Kirchsteiger (2004), for example, assume that people feel entitled to the average of what they could receive. Translating this into our context means that the agent would feel entitled to the average she could have received irrespective of her own signal and irrespective of the contract that the principal and agent agreed upon before she invested effort into the project. On the other hand, Hart and Moore (2008) assume that contracts constitute reference points and that ex-post people feel entitled to the maximum as specified by the contract that all parties had voluntarily agreed upon ex-ante. In our setting this means $\tilde{\omega}(\Gamma) = \omega_H$ independent of the agent's own performance signal.[4]

As in Hart and Moore (2008), also in our context it is natural to assume that a contract that the principal and agent voluntarily agree upon before the agent invests effort into the project shapes the parties feelings of entitlements ex-post. In case there exists a mutual agreement on the terms of the contract, feelings of entitlement arise related to the possible payments agreed upon ex-ante.

Interestingly, the existing literature on reciprocity abstracts from the question in what situation/how feelings of entitlement arise. Agents either feel entitled to the high wage as in Hart and Moore (2008) or to the average of what they could have received as e.g. in Dufwenberg and Kirchsteiger (2004). In contrast, we assume that feelings of entitlement are not exogenous, but endogenously shaped by the agent's own performance assessment which is influenced by the agent's own work effort. As mentioned above, this concept is closely related to Carmichael and MacLeod (2003) in which it is analyzed

[4]Note that this seems to imply a strong self-serving bias (the so-called 'Lake Wobegon effect'; see Hoorens (1993)) as the agent feels entitled to the highest possible wage independent of her own perception concerning her performance.

how caring about sunk costs can help agents achieve efficient investments in a team production environment in which agents bargain about the division of the surplus only after they have made their investment decisions. Also in their setting with symmetric information agents' feelings of entitlement in the ex-post bargaining stage depend on their ex-ante investments, i.e. feelings of entitlement are endogenous.

More specifically, we model feelings of entitlement in the following way

$$\tilde{\omega} = \lambda_{s_A}\omega_H + (1 - \lambda_{s_A})\,\omega_L,$$

where the parameter $\lambda_{s_A} \in [0, 1]$ indicates the degree to which the agent feels entitled to the high wage ω_H given her own performance signal s_A. Conflict arises whenever the agent feels somehow entitled to the high wage ω_H, i.e. $\lambda_{s_A} > 0$, but only receives the low wage due to a bad performance evaluation by the principal.

In our set-up, the agent knows that the project is a success in case $s_A = H$, but she is uncertain about it if she receives the performance signal $s_A = L$. Given this, we assume that the agent feels fully entitled to the high wage ω_H, when she receives the positive performance signal $s_A = H$ (i.e. $\lambda_H = 1$), but only feels entitled to it with $\lambda_L < 1$, if she receives the low performance signal $s_A = L$.[5] In other words, if the agent believes she did a good job she feels entitled to a higher wage than if she believes she did a bad job. Quite intuitively then, the more effort she puts into the project, the more likely it is that the project is a success and, hence, the more likely it is that she feels fully entitled to the higher wage.

Furthermore, in line with the psychological evidence concerning agents' feelings of entitlements and fairness perceptions [see e.g. Landy and Murphy (1978), Ilgen and Taylor (1979), Greenberg (1986a) and Greenberg (1986b)], we assume that $\lambda_L(\cdot)$ is a decreasing function of the principal's signal quality

[5] Note that $\lambda_H = 1$ is assumed for presentational ease. All our results hold as long as the agent feels entitled to a higher wage in case she believes she did a good job than if she believes she did a bad job, i.e. $\lambda_H > \lambda_L$.

g (i.e. $\frac{\partial \lambda_L(g)}{\partial g} \leq 0$). Intuitively, the more knowledgable the principal is, the less the agent feels shortchanged when she does not get the high wage ω_H in case she is uncertain about the project's success (i.e. in case $s_A = L$).[6] This is, we assume

$$\tilde{\omega} = \begin{cases} \omega_H & \text{if } s_A = H \\ \lambda_L(g)\omega_H + (1 - \lambda_L(g))\omega_L & \text{if } s_A = L \end{cases} \tag{3}$$

with $\lambda_L(g) \in [0, 1)$ and $\frac{\partial \lambda_L(g)}{\partial g} \leq 0$.

3 Conflict, truthtelling and welfare

The agent's ability to create conflict has a negative as well as positive effect within our model. On the one hand, the principal's ability to incentivize the agent is burdened by potential future conflict. On the other hand, the risk of conflict enables the principal to truthfully reveal his signal. In absence of conflict, the principal never finds it optimal to pay out the bonus to the agent. As a result, the agent never finds it optimal to work. Conflict creates room for mutual beneficial relations. Furthermore, this dichotomy also shapes the principal's incentive to hire agents that are not likely to have an own opinion concerning their performance (i.e. 'Yes Men') and agents that are particularly good at evaluating themselves independently.

Optimal Conflict. The principal offers a contract where the agent is paid a higher wage when the principal reports H than when he reports L. We interpret this a a flat wage with a bonus payment following a report of H. That is, $\omega_H = f + b$ and $\omega_L = f$.[7] Given this and given the agent's feeling of entitlement, conflict arises when the principal reports L. However the size of this conflict depends on the agent's own performance signal. Following a

[6]For notational simplicity we write λ_L instead of $\lambda_L(g)$ whenever no confusion might arise.

[7]Notice that in principle f can be negative as long as the agent's participation constraints is not violated.

report L, the agent is paid the fixed wage f and $\tilde{\omega} - \omega = \lambda_{s_A} b$. Depending on her own signal s_A the agent chooses the level of conflict q_{s_A} to minimize her psychological cost of conflict, i.e.

$$\min_{q_{s_A}} \theta \cdot \lambda_{s_A} \cdot b \cdot (1 - q_{s_A}) + c(q_{s_A})$$

where the optimal level of conflict $q^*_{s_A}$ given the agent's own signal s_A is implicitly given by

$$c'\left(q^*_{s_A}\right) = \theta \cdot \lambda_{s_A} \cdot b.$$

That is, the optimal level of conflict is a function of the bonus, i.e. $q^*_{s_A}(b)$, with

$$\frac{dq^*_{s_A}}{db} > 0 \tag{4}$$

and it is increasing in the agent's sensitivity to reciprocity θ and in the degree to which she feels entitled to the high wage λ_{s_A}.

The higher the bonus agreed upon in the contract, the stronger the reciprocal agent's reaction in case of conflict. Intuitively, the higher the bonus that the agent could have earned in case the principal had reported a high signal, the stronger the agent's reaction when she does not receive this bonus. Thus, the higher the bonus b, the higher is the potential conflict level $q^*_{s_A}$.

Furthermore, as $\lambda_H = 1 > \lambda_L$, it follows that the optimal level of conflict is higher in case the agent's own evaluation is H compared to the case in which it is L, i.e. $q^*_H > q^*_L$.

The Agent's Choice of Effort. With knowledge of the potential future conflict level, the agent's optimal effort choice $\hat{\tau}$ can be derived. The agent maximizes utility

$$\begin{aligned} U = {} & f + \tau(\gamma_{HH} + \gamma_{HL})b - v(\tau) - \tau\gamma_{LH}[\theta b(1 - q^*_H(b)) + c(q^*_H(b))] \\ & - \tau\gamma_{LL}[\theta\lambda_L b(1 - q^*_L(b)) + c(q^*_L(b))], \end{aligned}$$

with respect to τ which yields the following implicit relationship between the bonus offered by the principal and the effort level optimally chosen by the agent

$$b(\tau) = \frac{v'(\tau) + \gamma_{LH}\left[\theta b\left(1 - q_H^*(b)\right) + c\left(q_H^*(b)\right)\right]}{\gamma_{HH} + \gamma_{HL}} + \frac{\gamma_{LL}\left[\theta \lambda_L b\left(1 - q_L^*(b)\right) + c\left(q_L^*(b)\right)\right]}{\gamma_{HH} + \gamma_{HL}}, \tag{5}$$

From equation 5 it can be seen that the incentive compatible bonus simultaneously has to overcome effort costs and expected costs of conflict. Thus performance pay creates an endogenous source of conflict if agents behave reciprocal. The principal would like to incentivize the agent to perform high effort, but by doing so he generates potential conflict.

Truthtelling. Since the principal can report either H or L irrespective of his actual signal s_P, he will only choose to report his true signal if his expected profit from doing so is higher than his expected profit from doing otherwise.

Suppose $s_P = H$. Then, the principal tells the truth whenever his expected payoff from doing so (which is given by $\tau\phi - f - b$) exceeds his expected payoff from reporting L (which is given by $\tau\phi - f - pr(s_A = H|s_P = H)\psi q_H^* - pr(s_A = L|s_P = H)\psi q_L^*$). Consequently, the principal reports H if

$$\begin{aligned} b &\leq \frac{\gamma_{HH}}{\gamma_{HH} + \gamma_{HL}}\psi q_H^* + \frac{\gamma_{HL}}{\gamma_{HH} + \gamma_{HL}}\psi q_L^* \\ &= \left(\rho + (1-\rho)x\right)\psi q_H^* + (1-\rho)(1-x)\psi q_L^* \equiv b^{max}. \end{aligned} \tag{6}$$

The principal cannot credibly commit to bonuses above b^{max}. The reason is that for very high bonuses the principal has an incentive to report L irrespective of his true signal s_P. In other words, for sufficiently high bonus levels he prefers to face possible costs of conflict rather than paying the bonus. The value of the maximal credible bonus b^{max} is increasing in the quality of the agent's independent signal x and the correlation between the principal's and the agent's signal ρ. Furthermore, b^{max} is increasing in the levels of conflict

q_H^* and q_L^*.

If instead the principal receives signal $s_P = L$ he tells the truth whenever the payoff from doing so (which is given by $\tau\phi - f - pr(s_A = H|s_P = L)\psi q_H^* - pr(s_A = L|s_P = L)\psi q_L^*$) exceeds his payoff from reporting H (which is given by $\tau\phi - f - b$). Hence, the principal reports L if

$$b \geq \frac{\gamma_{LH}}{\gamma_{LH} + \gamma_{LL}}\psi q_H^* + \frac{\gamma_{LL}}{\gamma_{LH} + \gamma_{LL}}\psi q_L^*$$
$$= (1-\rho)\, x\psi q_H^* + (\rho + (1-\rho)(1-x))\, \psi q_L^* \equiv b^{min}. \qquad (7)$$

From this expression it is clear that the principal can also not credibly commit to very low bonuses. The reason is that for such low bonuses the principal has an incentive to evade conflict by always paying out the bonus regardless of his signal. This, in turn, would be anticipated by the agent who would simply not provide any (costly) effort and still get the bonus. The value of the lowest credible bonus is decreasing in ρ, increasing in x, and increasing in the levels of conflict.

Importantly, the principal has to offer a bonus $b \in [b^{min}, b^{max}]$ to incentivize the agent. Furthermore, equations 6 and 7 reveal that without conflict, i.e. with $q_H^* = 0$ and $q_L^* = 0$, the principal cannot truthfully commit to any positive bonus as both $b^{min} = b^{max} = 0$. Hence, potential conflict is crucial for principal-agent environments based on non-verifiable subjective performance evaluations as only bonuses that the principal can truthfully commit to create the basis for any mutually beneficial relation.

As can be concluded from this section, in order to incentivize the agent the principal has to offer a bonus which is credible. In addition to being credible, the bonus also has to sufficiently compensate the agent for his cost of effort and potential cost of conflict [see Appendix A.2 for a complete presentation of the pure moral hazard effect in our principal-agent environment]. In particular, as the contract establishes a reference point, and incentive pay constitutes an endogenous source of conflict, there could be situations in which the agent's optimal choice of effort is unresponsive to increases in the bonus offered by the principal simply because the risk of future conflict outweighs the potential benefit from receiving a bonus. Given this, the question arises which effort levels can and will optimally be implemented by the principal.

Optimal Effort Level. What is the optimal effort level τ^* that the principal implements? Let τ^{min} and τ^{max} be the (incentive compatible) effort levels implemented by bonus b^{min} and b^{max} respectively. That is, τ^{min} is the effort level optimally chosen by the agent when she is offered the bonus level b^{min}. Furthermore, let $\tilde{\tau} = \arg\max \Pi(\tau)$ be the effort choice that the principal would choose in the absence of the truthtelling limits τ^{min} and τ^{max}.

Remember that the principal has to offer a bonus $b \in [b^{min}, b^{max}]$ to incentivize the agent to work. Whether the principal finds it worthwhile to offer the agent such a contract depends on whether his expected profit from doing so is positive. This depends, among other things, on the project value.

As it turns out, not all project values are large enough for the principal to find it profitable to induce the agent to work. The bonus required to incentivize the agent may be too large relative to the expected value of the project. The principal will find it profitable to incentivize the agent to work only if the value of the project is such that $\phi > \overline{\phi}$ where $\overline{\phi}$ is the value of the project at which the principal's expected profit is zero if τ^{min} is implemented, i.e $\Pi(\tau^{min})|_{\phi=\overline{\phi}} = 0$. [See Appendix A.3 for a complete characterization of the conditions under which a mutually beneficial relationship arises]. When the principal finds it profitable to offer the agent a contract that induces her to work, we say that the principal *implements* a positive effort level.

Suppose the value of the project is such that the principal finds it optimal to induce the agent to work. The following lemma characterizes the optimal effort level τ^* that the principal implements given the lower and upper truthtelling constraint b^{min} and b^{max}. Remember that the relationship between bonus and effort level optimally chosen by the agent, $b(\hat{\tau})$, is given by equation 5.

Lemma 2. The effort level implemented by the principal is described by the following three cases

(1) Binding lower truth-telling constraint: the principal implements $\tau^* = \tau^{min}$ with bonus b^{min} if $0 < \tilde{\tau} < \tau^{min}$.

(2) Binding upper truth-telling constraint: the principal implements $\tau^* = \tau^{max}$ with bonus b^{max} if $\tilde{\tau} > \tau^{max}$.

(3) Non-binding truth-telling constraint: the principal implements $\tau^* = \tilde{\tau}$ by paying $b(\tilde{\tau})$ if $\tilde{\tau} \in [\tau^{min}, \tau^{max}]$.

Proof: Follows directly from the shape of the profit function. See Appendix A.2.

The principal implements $\tilde{\tau}$ whenever possible (i.e. Case (3) of lemma 2). However, he is limited to τ^{max} and τ^{min} whenever the bonus associated with the effort level that he actually would like to implement in absence of the truthtelling limits lies above or below the thresholds that he can credibly commit to (i.e. Cases (1) and (2) of lemma 2).

Welfare. What are the welfare implications of conflict costs and agent characteristics such as the agent's sensitivity to reciprocity in our strategic environment?

Before getting to the results, note two things. First, it is useful to define a characteristic of the agent's effort cost function which proves important for some of our welfare results. Generally speaking, any parameter change in our setting has two effects on welfare: a direct and an indirect. The direct effect captures the change in the principal's profit due to a change in the price of effort as a result of the parameter change. The indirect effect, on the other hand, regards the agent's optimal choice of effort which might change in response to a change in parameter. The magnitude of the indirect effect will depend on the curvature of the agent's effort cost function $v(\tau)$. Specifically, it depends on the measure

$$\frac{v'(\tau)}{v''(\tau)}, \tag{8}$$

which captures the degree of 'convexity' of the agent's effort costs.

Second, note that total welfare is given by the principal's profit since the agent does not earn any rent.

Given this, the following results obtain:

Proposition 1. Welfare is increasing in the principal's costs of conflict ψ, if the value of the project ϕ is sufficiently high.

Proof: Appendix A.5

Proposition 1 shows that conflict can have a welfare enhancing impact in principal-agent environments based on subjective performance evaluations. When the upper truthtelling constraint is binding, an increase in the principal's sensitivity to conflict ψ can increase the principal's profit and thus increase welfare. As already hinted at in the beginning of this section, there are two effects of an increase in ψ. First, ignoring the truthtelling problem (i.e. the pure moral hazard case) the direct effect on welfare of an increase in ψ for a given effort level is negative. However, since $\frac{d\tau^{max}}{d\psi} > 0$, an increase in the principal's sensitivity to conflict ψ also relaxes the upper truthtelling constraint. Thus, when the upper truthtelling constraint is binding, an increase in ψ can enable the principal to credibly commit to, and hence implement, higher effort levels. This in turn increases the expected profit. When the potential value of the project ϕ is sufficiently high, the latter effect dominates and welfare is increasing in ψ.

It is not only the principal's cost of conflict that has an impact on welfare. Agent characteristics' such as the agent's sensitivity to reciprocity, the likelihood with which she forms an independent opinion concerning her performance as well as her ability to independently identify a successful project can also influence the principal's profit, and thus welfare, as the following results demonstrate.

First, we focus on the agent's sensitivity to reciprocity θ. Imagine that the principal can choose between two agents that are identical in all respects except for their value of θ.[8] One has a high value of θ and the other a low value. Which type will the principal prefer to hire?

[8]Notice that when we say the agents are *identical* in all respects except for their sensi-

Proposition 2. An increase in the agent's sensitivity to reciprocity can increase welfare if

(i) the expected value of the project is sufficiently high, and the principal is sufficiently sensitive to conflict.

(ii) the value of the project is small, the principal is not too sensitive to conflict and the agent's effort costs are not too 'convex' (i.e. the measure 8 is sufficiently large).

Proof: Appendix A.6

Interestingly, proposition 2 shows that it might be beneficial for the principal to hire an agent who has a high emotional sensitivity to reciprocity even if this agent will potentially impose high conflict costs on the principal. An increase in the agent's sensitivity to reciprocity increases the expected cost of conflict. This will make it less tempting for the principal to lie if he receives the high performance signal, which in turn relaxes the upper truthtelling constraint. However, an increase in the agent's sensitivity also makes a given effort level more expensive to implement since the agent must be compensated for potential conflict costs. If the principal is sufficiently sensitive to conflict, the first effect will dominate, and welfare is increasing in θ. Likewise, an increase in the agent's sensitivity makes it more tempting to lie if the principal receives the low performance signal. However, it also increases the price the principal has to pay for a specific effort level which makes it less tempting to lie. If the principal is not very sensitive to conflict this last effect will dominate.

Second, regarding the correlation between the principal's and agent's signal ρ:

Proposition 3. Welfare is decreasing in the correlation of signals ρ, if the value of the project ϕ is sufficiently low, the principal is not too sensitive to conflict and the agent's effort costs are not too 'convex' (i.e. the measure 8 is sufficiently large).

---

tivity θ that also includes their outside option, which is normalized to zero for both types of agents.

Proof: Appendix A.7

Third, regarding the agent's ability to independently evaluate the success of the project x:

Proposition 4. An increase in the agent's ability to independently identify a successful project can increase welfare in the following two ways:

(i) if the value of the project is small, the principal is not too sensitive to conflict and the agent's effort cost is not too 'convex' (i.e. the measure 8 is sufficiently large).

(ii) if the expected value of the project is sufficiently large, and the principal is sufficiently sensitive to conflict.

Proof: Appendix A.8

Imagine two agents, Agent 1 and Agent 2, who are identical except for their ability to independently evaluate the success of the project and the correlation between their own signal and the principal's signal. Assume Agent 2 has a lower value of ρ and a higher value of x compared to Agent 1. That is, Agent 2 will - compared to Agent 1 - more often create conflict. Hiring Agent 2 instead of agent 1 has two effects on welfare. First, it is more expensive to induce Agent 2 to work. Agent 2 requires a higher incentive compatible bonus for every given effort level compared to Agent 1. On the other hand, since conflict is more of a risk with Agent 2, it is possible that the range of effort levels the principal can credibly implement is larger for Agent 2 compared to Agent 1. Naturally, in some cases the first effect will dominate and the principal will find it welfare enhancing to hire Agent 1. In other cases, the principal might be able to implement more desired effort levels with Agent 2 which are unfeasible in case the principal decides to hire Agent 1.

Intuitively, propositions 3 and 4 show that if the project value is such that truthtelling constraints are not a concern, the principal should always hire an agent for whom the likelihood of having an own opinion is minimized (i.e. 'Yes Men'), and an agent that is not good in independently evaluating the success of the project. These two 'agent characteristics' or 'personality traits' minimize the potential for conflict and, hence, increase welfare. However, if the project value is sufficiently low, the principal might find it optimal to hire an agent for whom the likelihood of having an own opinion concerning her performance is high. In addition, for sufficiently high and low project values the principal might find it optimal to hire an agent who is very good in independently identify the success of the project.

These welfare effects highlight that personality tests that e.g. assess an applicant's sensitivity to reciprocity or ability to form an own opinion can play an important role in recruitment processes in work environments in which firms cannot write complete contracts that specify all aspects of the employment relation. To form a mutually beneficial and optimal relation the personality of an applicant should fit the character of the vacancy he or she applies to.

4 The choice of evaluation procedures

Until now, we have investigated optimal contract design and welfare implications of an *exogenously* given quality of the principal's signal g. 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.[9]

Of course, under classical assumptions about preferences the quality of the

[9] Note that we explicitly avoid terms like *control* and *(dis)trust* here (as e.g. used in Falk and Kosfeld (2006) and Ellingsen and 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.

evaluation procedure has no impact on the effort choice of the agent in our setting. The agent simply does not trust the principal to truthfully reveal his signal and hence provides no effort. In contrast to this, however, it has been suggested in recent experimental and theoretical works that procedural choices are important in strategic interactions with reciprocal agents [see e.g. Blount (1995), Sebald (2010), Aldashev *et al.* (2010)]. Procedural choices are important because reciprocal agents might exhibit procedural concerns and, hence, react differently in outcome-wise identical situation depending on the evaluation/decision-making procedure which led to the outcome. Translated into our setting, the agent's perception concerning the kindness of the principal towards her depends on the evaluation procedure chosen by the principal. The higher the quality of the evaluation process, the kinder the agent perceives the principal and, hence, the kinder the agent's response.

To formally analyze the impact of the quality of the evaluation procedure on the agent's effort choice and the principal's optimal choice of signal quality in our setting, 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 in our setting with subjective performance evaluations.

Implementable Bonuses Remember the following properties of the relation between $b(\tau)$, b^{min}, b^{max} and the quality of the principal's evaluation procedure g:

(i) a bonus b which makes the effort choice of τ incentive compatible only satisfies the upper and lower truthtelling constraint of the principal if $b \in [b^{min}, b^{max}]$,

(ii) the incentive compatible bonus $b(\tau)$ in our setting is monotonically decreasing in the principal's signal quality g with $\lim_{g\to 0} b(\tau) = \infty$ and $\lim_{g\to 1} b(\tau) = v'(\tau)$ [see also Appendix A.2] and

(iii) b^{max} and b^{min} are (weakly) monotonically decreasing in g (because $\frac{dq_L^*}{dg} \leq$

0) with $\lim_{g \to 0} b^{min} < \infty$,

$$\lim_{g \to 1} b^{min} = (1-\rho)\, x\psi q_H^* + (\rho + (1-\rho)(1-x))\, \psi q_L^*(g=1) > 0,$$

$\lim_{g \to 0} b^{max} < \infty$ and

$$\lim_{g \to 1} b^{max} = (\rho + (1-\rho)x)\, \psi q_H^* + (1-\rho)(1-x)\psi q_L^*(g=1) > 0.$$

Properties (i)-(iii) allow us to distinguish the following possible cases describing the optimal choice of g for the implementation of an effort level $\tau > 0$.

Lemma 3. Fix some effort level $\hat{\tau} > 0$. In order to implement that specific effort level, the principal has to offer a bonus $b(\hat{\tau})$. The implementability of that effort level depends on the signal quality. One of the following cases will hold:

(1) $\hat{\tau}$ cannot be implemented regardless of the choice of g if $b(\hat{\tau}) > b^{max}$ for all g.

(2) $\hat{\tau}$ is implemented with the maximal g for which $b(\hat{\tau}) = b^{max}$ if $b(\hat{\tau}) \leq b^{max}$ for some $g < 1$ but $b(\hat{\tau}) > b^{max}$ for $g = 1$. That is, the principal chooses a less than perfect signal quality.

(3) $\hat{\tau}$ is implemented with the maximal g for which $b(\hat{\tau}) = b^{min}$ if $b(\hat{\tau}) \leq b^{max}$ for some $g < 1$ but $b(\hat{\tau}) < b^{min}$ for $g = 1$. That is, the principal chooses a less than perfect signal quality.

(4) $\hat{\tau}$ is implemented with $b(\hat{\tau}) = v'(\hat{\tau})$ and the principal chooses perfect signal quality (i.e. $g = 1$) if $b(\hat{\tau}) \in [b^{min}, b^{max}]$ for $g = 1$.

Proof: Follows directly from the aforementioned properties (i)-(iii).

In Case (1) effort $\hat{\tau}$ cannot be implemented with any signal quality g because the incentive compatible bonuses are too large to be credible. This situation arises, for example, if the agent is insensitive to reciprocity (i.e. $\theta = 0$) or there are no retaliation opportunities (i.e. $\psi = 0$). Case (4), on the other hand, depicts the situation in which the incentive compatible bonus is credible for signal quality $g = 1$. Cases (2)-(4) of lemma 3 show that, as a better signal quality reduces the probability of conflict and expected psychological costs, the principal will always implement τ with the largest possible signal quality which still ensures truthtelling.[10]

To graphically exemplify Cases (2) and (3) of lemma 3 consider the following two possible scenarios

[Figures 1 and 2 here]

Figure 1 shows Case (2) in which the incentive compatible bonus $b(\tau)$ is lower than b^{max} for some $g < 1$, but higher at $g = 1$. In this case the optimal bonus and signal quality is denoted $\overline{b}^{max}$ and $\overline{g}$. Figure 2, on the other hand, shows Case (3) in which $b(\tau) \leq b^{max}$ for some $g < 1$ but $b(\tau) < b^{min}$ for $g = 1$. In this case the optimal bonus and signal quality is respectively denoted by $\overline{b}^{min}$ and $\overline{g}$.

Welfare Implications. From the above analysis it is clear that welfare is not always increasing in the quality of the principal's signal. What are the precise conditions under which welfare is increasing or decreasing in the quality of the evaluation procedure?

Proposition 5. The welfare effect of the choice of signal quality:

(i) If the principal's preferred choice of effort is unbounded by the truthtelling constraints, welfare is unambiguously increasing in the principal's signal quality g.

(ii) Welfare can be decreasing in the signal quality, if the principal's choice of effort is bounded by his truthtelling constraints:

[10] 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 $C(\tau)$ [see Appendix A.2] and increasing costs of quality.

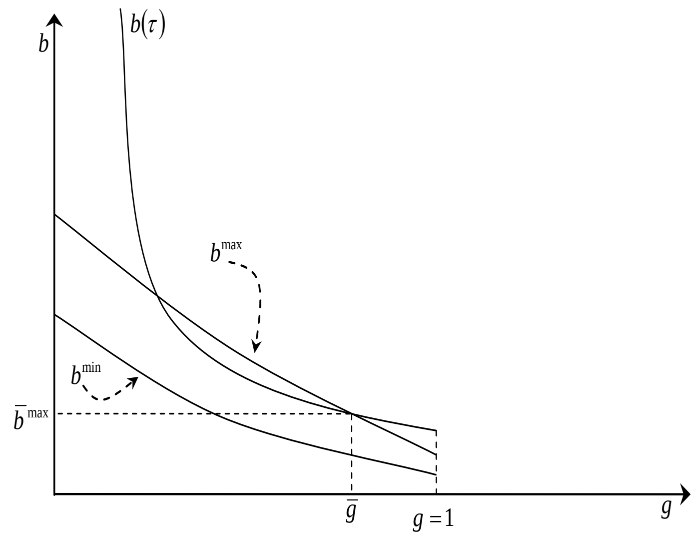

Figure 1: The Quality of the Evaluation Process: *Case (2)*

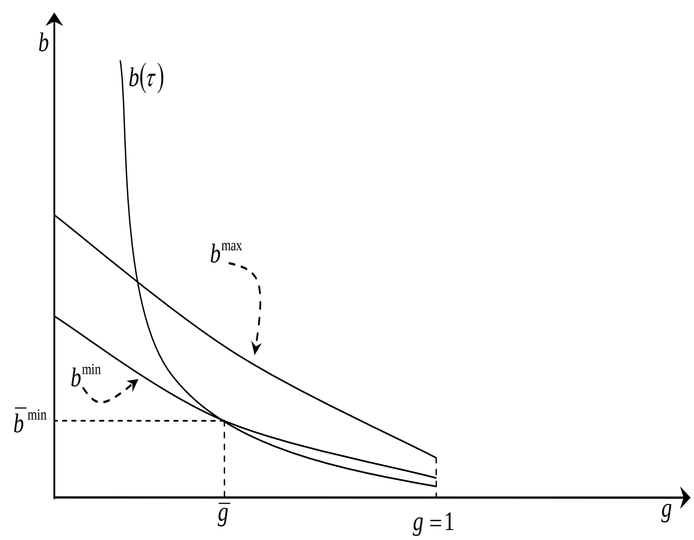

Figure 2: The Quality of the Evaluation Process: *Case (3)*

(a) Welfare is decreasing in the signal quality if the project value is sufficiently high such that the principal is bounded by the upper-truthtelling constraint and the principal's cost of conflict is sufficiently high $\psi > \tilde{\psi}$.

(b) Welfare is decreasing in the signal quality if the project value is sufficiently low such that the principal is bounded by the lower-truthtelling constraint, the principal's cost of conflict is not too high $\psi < \tilde{\psi}$, and the agent's effort cost is not too 'convex' (i.e. the measure 8 is sufficiently large).

Proof: Appendix A.9

Intuitively, if the principal is free to implement his most preferred effort level there is only the direct effect on profit from a change in g and this is positive since the 'price' that the principal has to pay to implement a certain effort level is decreasing in g.

However, when the lower or upper truthtelling constraint binds it is possible that welfare be decreasing in the quality of the principal's signal. This is because an increase in g affects the highest and lowest implementable effort levels by changing the range of credible bonuses.

As a first example, imagine that the principal is bounded by the upper truthtelling constraint. An increase in g has ambiguous effects on the maximum implementable effort level. First, effort is cheaper for higher levels of signal quality which has an increasing effect on τ^{max}. Second, since a higher value of g also makes conflict less likely, it is more tempting for the principal to lie when he receives signal H. This decreases the maximum credible bonus and pulls towards a lower value of τ^{max}. If the principal is sufficiently sensitive to conflict, the level of b^{max} will change so much that the positive effect on τ^{max} is outweighed by the negative effect. Thus the principal will find himself unable to commit to high bonus levels following an increase in g. If the project is sufficiently valuable and hence requires high effort levels, such a change can decrease welfare since it limits the principal to choose 'too low' effort levels.

As a second scenario, imagine that the principal is bounded by the lower truthtelling constraint. When the principal chooses a better signal quality the agent responds by creating lower potential conflict. This decreases the minimum credible bonus. However, since an increase in g also makes a given effort level cheaper it is possible that the minimum credible effort level is increasing in g because the principal will more often prefer to evade conflict by paying out the bonus unconditionally. Now, if the principal is bounded by the lower truthtelling constraint and an increase in g tightens this truthtelling constraint, it is possible that welfare decreases overall because the principal has to implement a 'too high' effort level which is costly - even if this effort level comes cheaper as a result of the higher value of g.

5 Conclusion

Our analysis focused on the role and importance of conflict in work environments based on non-verifiable subjective performance evaluations. Contrary to the existing literature we did not model conflict as e.g. third-party payments optimally chosen to ensure truthtelling, but explicitly formalized conflict as the reciprocal reaction of agents that feel shortchanged and unkindly treated by their principal.

In our setting, contracts constitute frames/reference points and performance pay creates an endogenous source of conflict since the agent's feelings of entitlement, and her potential reciprocal reaction, is intensifying in the bonus. In other words, by promising to pay a bonus the principal incentivizes the agent to perform effort while simultaneously generating potential conflict.

We showed that the principal's optimal choice of contract in such an environment is limited by a maximum and minimum bonus that he can credibly commit to. Bonuses above the upper threshold or below the lower threshold fail to fulfill the principal's truthtelling constraint and, hence, lead to an inefficiently low effort provision by the agent. These limits, in turn, influence the optimal effort levels that the principal can actually implement.

Explicitly modeling conflict as originating form the reciprocal reaction of an agent who feels shortchanged revealed interesting welfare effects. In particular, linking up with a fairly recent literature in economics and a long standing debate in the organizational behavior/human resource literature we showed that agent characteristics play a crucial role in principal-agent environments based on non-verifiable performance evaluations. In this way, our analysis provides one rational for the use of personality tests and applicant screenings in recruitment and promotion processes.

Furthermore, following the recent literature on procedural concerns, we extended our framework to allow the principal to choose between evaluation procedures that differ in terms of the quality of the principal's signal. In our setting the choice of evaluation procedure influences the agent's feeling of entitlement as well as the a priori probability of conflict. Interestingly, our analysis reveals that even if it is costless for the principal to choose a perfect evaluation procedure, he might not optimally choose a perfect evaluation process. The principal may benefit from some 'noise' in the evaluation procedure since this creates a risk of conflict making more desired effort levels implementable.

Finally, we feel that there are at least two important directions for future research. First, in the same way as Fuchs (2007) has extended MacLeod (2003), we feel that it is also important to take our ideas to a repeated setting and explore the interplay between personal traits, reciprocity and reputational effects. This seems particularly important in the light of empirical evidence showing the important connection between concerns for reciprocity and reputation [e.g. Fehr *et al.* (2009) and Gächter and Falk (2002)]. Second, experiments should be conducted that not only test the assumptions we make regarding the agent's reciprocal inclination, but also the theoretical implications of our theory.

References

Aldashev, G., Kirchsteiger, G. and Sebald, A. (2010). How (not) to decide: Procedural games. *Working Papers ECARES*.

Barrick, M. and Mount, M. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, **44** (1), 1–26.

Bartling, B., Fehr, E. and Schmidt, K. (2012). Screening, competition, and job design: Economic origins of good jobs. *The American Economic Review*, **102** (2), 834–864.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes*, **63** (2), 131–144.

Bushman, R., Indjejikian, R. and Smith, A. (1996). Ceo compensation: The role of individual performance evaluation. *Journal of Accounting and Economics*, **21** (2), 161–193.

Cable, D. and Judge, T. (1994). Pay preferences and job search decisions: A person-organization fit perspective. *Personnel psychology*, **47** (2), 317–348.

Carmichael, L. and MacLeod, W. (2003). Caring about sunk costs: a behavioral solution to holdup problems with small stakes. *Journal of law, economics, and organization*, **19** (1), 106–118.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics*, **22** (3), 665–688.

Chatman, J., Caldwell, D. and O'Reilly, C. (1999). Managerial personality and performance: A semi-idiographic approach. *Journal of Research in Personality*, **33** (4), 514–545.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, **47** (2), 268–298.

Ellingsen, T. and Johannesson, M. (2008). Pride and prejudice: The human side of incentive theory. *The American Economic Review*, **98** (3), 990–1008.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *The American economic review*, **96** (5), 1611–1630.

Fehr, E., Brown, M. and Zehnder, C. (2009). On reputation: A microfoundation of contract enforcement and price rigidity*. *The Economic Journal*, **119** (536), 333–353.

—, Gächter, S. and Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica: journal of the Econometric Society*, pp. 833–860.

—, Kirchsteiger, G. and Riedl, A. (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics*, **108** (2), 437.

Fuchs, W. (2007). Contracting with repeated moral hazard and private evaluations. *The American Economic Review*, **97** (4), 1432–1448.

Gächter, S. and Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *The Scandinavian Journal of Economics*, **104** (1), 1–26.

Gibbs, M., Merchant, K., Van der Stede, W. and Vargus, M. (2004). Determinants and effects of subjectivity in incentives. *Accounting Review*, **79** (2), 409–436.

Greenberg, J. (1986a). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology*, **71** (2), 340–342.

— (1986b). Organizational performance appraisal procedures: what makes them fair? In L. R. Bazerman, M.H. and B. Sheppard (eds.), *Research on Negotiation in Organizations*, JAI Press, Greenwich, CT.

Hart, O. and Moore, J. (2008). Contracts as reference points*. *Quarterly Journal of Economics*, **123** (1), 1–48.

Hoorens, V. (1993). Self-enhancement and superiority biases in social comparison. *European review of social psychology*, **4** (1), 113–139.

Huang, F. and Cappelli, P. (2010). Applicant screening and performance-related outcomes. *American Economic Review*, **100** (2), 214–218.

Ilgen, F. C. D., D. R. and Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, **64**, 349–371.

Ittner, C., Larcker, D. and Meyer, M. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *Accounting Review*, pp. 725–758.

—, — and Rajan, M. (1997). The choice of performance measures in annual bonus contracts. *Accounting Review*, pp. 231–255.

Judge, T. and Cable, D. (1997). Applicant personality, organizational culture, and organization attraction. *Personnel Psychology*, **50** (2), 359–394.

Kristof, A. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel psychology*, **49** (1), 1–49.

Kube, S., Maréchal, M. A. and Puppe, C. (2011). The currency of reciprocity - gift-exchange in the workplace. *The American Economic Review*, fortcoming.

Landy, B. J. L., E. J. and Murphy, K. R. (1978). Correlates of perceived fairness and accuracy of performance evaluation. *Journal of Applied Psychology*, **63**, 751–754.

Levin, J. (2003). Relational incentive contracts. *The American Economic Review*, **93** (3), 835–857.

Li, W., Wang, Y., Taylor, P., Shi, K. and He, D. (2008). The influence of organizational culture on work-related personality requirement ratings: A multilevel analysis. *International Journal of Selection and Assessment*, **16** (4), 366–384.

Macleod, W. (2000). Complexity and contract. *Revue d'Économie Industrielle*, **92** (1), 149–178.

MacLeod, W. (2003). Optimal contracting with subjective evaluation. *The American Economic Review*, **93** (1), 216–240.

Murphy, K. and Oyer, P. (2001). Discretion in executive incentive contracts: Theory and evidence. *Working paper*.

Prendergast, C. (1993). A theory of" yes men". *The American Economic Review*, pp. 757–770.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review*, **83** (5), 1281–1302.

Raymark, P., Schmit, M. and Guion, R. (1997). Identifying potentially useful personality constructs for employee selection. *Personnel Psychology*, **50** (3), 723–736.

Schein, E. (2004). *Organizational culture and leadership*, vol. 1. Jossey-Bass Inc Pub.

Sebald, A. (2010). Attribution and reciprocity. *Games and Economic Behavior*, **68** (1), 339–352.

— and Walzl, M. (2012a). Optimal contracts based on subjective evaluations and reciprocity. *Working Paper Series, University of Innsbruck, 2012–16.*

— and — (2012b). Subjective performance evaluations and reciprocity in principal-agent relations. *Working Paper Series, University of Innsbruck, 2012–15.*

Tett, R. and Christiansen, N. (2007). Personality tests at the crossroads: A response to morgeson, campion, dipboye, hollenbeck, murphy, and schmitt (2007). *Personnel Psychology*, **60** (4), 967–993.

—, Jackson, D., Rothstein, M. and Reddon, J. (1999). Meta-analysis of bidirectional relations in personality-job performance research. *Human Performance*, **12** (1), 1–29.

A Appendix

A.1 Proof of Lemma 1

To save on notation, we denote $\theta \cdot \max\{\tilde{\omega} - \omega, 0\} \cdot (1-q) - c(q) \equiv Y_{kl}(\tau)$ 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.

Suppose that $\Gamma = \{\omega_{kl}\}$ is a revelation contract, i.e. the principal and the agent tell the truth under contract Γ and Γ implements $\tau > 0$. Then the incentive compatibility constraint

$$\Sigma_{k \in S_P, l \in S_A} (\omega_{kl} - Y_{kl}(\tau)) \frac{dPr\{s_P = k, s_A = l\}}{d\tau} = v'(\tau)$$

is satisfied. Consider a contract $\hat{\Gamma}$ which fixes payments of $\hat{\omega}_k = \sum_{l \in S_A} \omega_{kl} Pr\{s_P = k, s_A = l\}$ 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).[11] Moreover, the agent weakly benefits from telling the truth. Finally, the principal's truth-telling constraint is also satisfied under $\hat{\Gamma}$. To see this observe that the principal reports k given that he has received k under contract Γ if

$$\begin{aligned} & Pr\{s_A = H | s_P = k\}(\omega_{oH} - \omega_{kH}) + Pr\{s_A = L | s_P = k\}(\omega_{oL} - \omega_{kL}) \\ \geq \; & Pr\{s_A = H | s_P = k\}((q^*\psi)_{kH} - (q^*\psi)_{oH}) \\ & + Pr\{s_A = L | s_P = k\}((q^*\psi)_{kL} - (q^*\psi)_{oL}) \end{aligned} \tag{9}$$

for all $o \in S_P$ (where $(q^*\psi)_{l,k}$ denotes the anticipated conflict costs for a reported configuration (l, k)). This set of inequalities holds because Γ imple-

[11] 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.

ments truth-telling by assumption. $\hat{\Gamma}$ implements truth-telling if

$$\hat{\omega}_o - \hat{\omega}_k \geq \quad Pr\{s_A = H|s_P = k\}((q^*\psi)_{kH} - (q^*\psi)_{oH}) + Pr\{s_A = L|s_P = k\}((q^*\psi)_{kL} - (q^*\psi)_{oL}). \tag{10}$$

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

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

which coincides with equation 9 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 ω_{kl} independent of l which also implements $\tau > 0$ and leaves the principal weakly better off.

Part (ii). Suppose by contradiction that Γ implements $\tau > 0$ with $\omega_H = z$ and $\omega_L = z + \epsilon$ 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.

Not for publication / Online material

A.2 Pure moral hazard

The principal's objective to offer a profit maximizing contract - i.e. an optimal combination of a fixed payment and a bonus - is burdened by (i) a moral hazard problem and (ii) a truthtelling problem as the agent's effort is unobservable, and the principal has to credibly commit himself to a truthful revelation of his own signal. This section of the appendix will analyze the pure moral hazard problem abstracting from the truthtelling problem. That is, we focus on the dynamics between the bonus offered and the effort level optimally chosen by the agent taking truthtelling as given.

The optimal level of conflict $q^*_{s_A}$ is implicitly given by the following expression

$$c'\left(q^*_{s_A}\right) = \theta\lambda_{s_A} b.$$

The Agent's Effort Choice. The following result characterizes the relationship between bonus and optimal effort as chosen by the agent.

Result 1. The endogeneity of the conflict creates two cases describing the relation between bonus and optimal effort:

(i) There is a positive relationship between the offered bonus and the optimally chosen effort level for all levels of bonuses if $g - (1-g)(1-\rho)x\theta - (1-g)\left(\rho + (1-\rho)(1-x)\right)\theta\lambda_L \geq 0$.

(ii) There is a positive relationship between the offered bonus and the optimally chosen effort level only for bonuses above $\underline{b} > 0$, where $\underline{b}$ is the bonus level that solves the following equation

$$\begin{aligned} gb = {} & (1-g)(1-\rho)x\left[\theta\underline{b}(1-q^*_H) + c(q^*_H)\right] \\ & + (1-g)\left(\rho + (1-\rho)(1-x)\right)\left[\theta\lambda_L\underline{b}(1-q^*_L) + c(q^*_L)\right], \end{aligned}$$

if $g - (1-g)(1-\rho)x\theta - (1-g)\left(\rho + (1-\rho)(1-x)\right)\theta\lambda_L < 0$. For bonuses below $\underline{b}$ the optimally chosen effort level is zero and thus the effort level is unresponsive to changes in the offered bonus.

Proof. Inserting for γ_{HH},γ_{HL} and γ_{LH} in equation 5 and rearranging leads to:

$$\begin{aligned} & gb - (1-g)(1-\rho)x\left[\theta b(1-q^*_H(b)) + c(q^*_H(b))\right] \\ & \quad - (1-g)\left(\rho + (1-\rho)(1-x)\right)\left[\theta\lambda_L b(1-q^*_L(b)) + c(q^*_L(b))\right] = v'(\tau). \end{aligned} \tag{11}$$

Note that the LHS of equation 11 depends on the bonus b (and is independent of the effort τ), whereas the RHS depends on the effort τ (and is independent of the bonus b). Furthermore, as $v'' > 0$, the RHS is monotonically increasing in τ and $RHS(\tau) : [0,1] \to [0,\infty]$.

With regard to the LHS note first that $\frac{\partial q^*_{s_A}}{\partial b} > 0$ (see equation 4), $lim_{b\to 0}$ $q^*_{s_A} = 0$ and $lim_{b\to\infty} q^*_{s_A} = 1$.

This implies that

$$\frac{\partial LHS}{\partial b}|_{b=0} = g - (1-g)(1-\rho)x\theta - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L$$

$\frac{\partial LHS}{\partial b}|_{b=0}$ is (i) weakly positive if $g \geq (1-g)(1-\rho)x\theta - (1-g)(\rho+(1-\rho)(1-x))\theta\lambda_L$ and (ii) negative if $g < (1-g)(1-\rho)x\theta - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L$. It is important to see that as the exogenous parameters $g, \rho, x, \lambda_L \in [0,1]$ and $\theta \in (0,\infty)$, both cases are possible.

Now, the second derivative is given by

$$\frac{\partial^2 LHS}{\partial b^2} = (1-g)(1-\rho)x\theta\frac{\partial q^*_H}{\partial b} + (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L\frac{\partial q_L*}{\partial b} > 0$$

If the exogenous parameters, g, ρ, x, λ and θ, are such that $\frac{\partial LHS}{\partial b}|_{b=0} \geq 0$, then the LHS of equation 11 is monotonically increasing in b and, hence, Case (i) of Result 1 obtains. In contrast, if the exogenous parameters, g, ρ, x, λ and θ, are such that $\frac{\partial LHS}{\partial b}|_{b=0} < 0$, then the LHS of equation 11 is decreasing until $g = (1-g)(1-\rho)x\theta(1-q^*_H(b)) + (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q^*_L(b))$ and increasing thereafter. Hence, Case (ii) of Result 1 obtains. □

The following result shows that the non-positive relation between effort and bonus described in Case (ii) can always be overcome. That is, there always exist bonus levels high enough such that the agent will find it optimal to provide effort

Result 2. Irrespective of the information technology (i.e. g, x and ρ), sensitivity to reciprocity (i.e. θ) or the agent's feeling of entitlement, there always exists a bonus level $\hat{b} > \underline{b}$ above which the optimal effort level is positive, $\tau > 0$.

Proof. The derivative of the LHS of equation 11 with respect to b is

$$g - (1-g)(1-\rho)x\theta(1-q^*_H(b)) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q^*_L(b)) \quad (12)$$

Expression 12 might be negative for some b but

$$\lim_{b\to\infty} g-(1-g)(1-\rho)x\theta(1-q_H^*(b))-(1-g)(\rho+(1-\rho)(1-x))\theta\lambda_L(1-q_L^*(b)) = g > 0.$$

That is, there always exists a b such that the LHS of equation 11 is positive. This implies that there always exists a $\hat{b} : \underline{b} < \hat{b} < \infty$ for which $\tau > 0$ at $\hat{b}$. □

Result 2 shows that the counterproductive relation between bonus and conflict can always be overcome by paying a sufficiently high bonus. In other words, sufficiently high bonuses always imply a positive relation between bonus and effort irrespective of the information technology, sensitivity to reciprocity or the agent's feeling of entitlement. Intuitively, for sufficiently large bonuses the monetary incentive associated with a bonus always outweighs the potential conflict that this bonus creates and consequently the agent will choose to work.

Principal's Choice of Contract. We now know that the principal can always pay a bonus high enough such that the agent finds it optimal to work. The next question is whether the principal will always find it profitable to offer the agent such a bonus. That is, given the agent's optimal choice of conflict and effort, we are now interested in the question whether the principal always wants to implement a positive effort level $\tilde{\tau}$ independent of the potential profitability of the project ϕ. For Case (i) in Result 1 we can state the following result:

Result 3. When parameter values are such that Case (i) in Result 1 obtains, the principal will implement a positive effort level $\tilde{\tau}$ for all project values $\phi > 0$.

Proof. The principal's expected profit is given by

$$\Pi = \tau\phi - \tau\psi\left[\gamma_{LH} q_H^* + \gamma_{LL} q_L^*\right] - C(\tau)$$

where $C(\tau)$ is the expected labor cost given by $f + \tau(\gamma_{HH} + \gamma_{HL}) b$. The agent accepts the contract if he receives a weakly positive payoff from doing so. His participation constraint is therefore given by

$$f + \tau(\gamma_{HH} + \gamma_{HL}) b - v(\tau) - \tau\gamma_{LH}(\theta b(1 - q_H^*) + c(q_H^*)) \\ - \tau\gamma_{LL}(\theta\lambda_L b(1 - q_L^*) + c(q_L^*)) \geq 0$$

The principal will always choose f such that this participation constraint binds. Use this to rewrite $C(\tau)$ such that

$$C(\tau) = v(\tau) + \tau\gamma_{LH}\left[\theta b(1 - q_H^*) + c(q_H^*)\right] + \tau\gamma_{LL}\left[\theta\lambda_L b(1 - q_L^*) + c(q_L^*)\right].$$

Inserting this into the principal's profit function yields

$$\Pi = \tau\phi - \left[v(\tau) + \tau\gamma_{LH}\left(\psi q_H^* + \theta b\left(1 - q_H^*\right) + c\left(q_H^*\right)\right) + \tau\gamma_{LL}\left(\psi q_L^* + \theta\lambda_L b\left(1 - q_L^*\right) + c\left(q_L^*\right)\right)\right]$$

We want to show that there always exists an effort level such that the expected profit is positive. Hence we look into the shape of the profit function. The first term of the profit function $\tau\phi$ is linearly increasing in τ as long as $\phi > 0$. From this we subtract a function of τ given by the term in the square brackets. Label this function $F(\tau)$, i.e.

$$F(\tau) \equiv v(\tau) + \tau\gamma_{LH}\left(\psi q_H^* + \theta b\left(1 - q_H^*\right) + c\left(q_H^*\right)\right) + \tau\gamma_{LL}\left(\psi q_L^* + \theta\lambda_L b\left(1 - q_L^*\right) + c\left(q_L^*\right)\right).$$

$F(\tau)$ is convex and has the following properties: $F(0) = 0$ and $F'(0) = 0$. As a consequence, it is true that $\pi = \tau\phi - F(\tau)$ is positive for some values of τ, and a positive optimal $\tilde{\tau}$ exists if $\phi > 0$. □

For Case (ii) in Result 1 matters are different, and the principal does not always want to implement a positive effort level.

Result 4. When parameter values are such that Case (ii) in Result 1 obtains, the principal will implement a positive effort level only if

$$\begin{aligned}\phi &> \underline{\phi} \\ &= \gamma_{LH}\left(\psi q_H^*(\underline{b}) + \theta\underline{b}(1 - q_H^*(\underline{b})) + c(q_H^*(\underline{b}))\right) + \gamma_{LL}\left(\psi q_L^*(\underline{b}) + \theta\lambda_L\underline{b}(1 - q_L^*(\underline{b})) + c(q_L^*(\underline{b}))\right)\end{aligned}$$

Proof. Again, the principal's expected profit is equal to

$$\Pi = \tau\phi - F(\tau)$$

with

$$F(\tau) \equiv v(\tau) + \tau\gamma_{LH}\left(\psi q_H^* + \theta b(1 - q_H^*) + c(q_H^*)\right) + \tau\gamma_{LL}\left(\psi q_L^* + \theta\lambda_L b(1 - q_L^*) + c(q_L^*)\right)$$

$F(\tau)$ is convex and $F(0) = 0$. It remains to check the derivative.

$$\begin{aligned}F'(\tau) = v'(\tau) \quad &+ \quad \gamma_{LH}\left(\psi q_H^* + \theta b(1 - q_H^*) + c(q_H^*)\right) \\ &+ \quad \tau\gamma_{LH}\left(\psi\frac{\partial q_H^*}{\partial b}\frac{\partial b}{\partial \tau} + \theta\frac{\partial b}{\partial \tau}(1 - q_H^*)\right) \\ &+ \quad \gamma_{LL}\left(\psi q_L^* + \theta\lambda_L b(1 - q_L^*) + c(q_L^*)\right) \\ &+ \quad \tau\gamma_{LL}\left(\psi\frac{\partial q_L^*}{\partial b}\frac{\partial b}{\partial \tau} + \theta\lambda_L\frac{\partial b}{\partial \tau}(1 - q_L^*)\right).\end{aligned}$$

Now, for $\tau = 0$ we have $b = \underline{b}$. From equation (4) we know that the conflict level is increasing in b. Thus, when $b = \underline{b}$ it is not the case that $q^* = 0$. It holds,

$$F'(0) = \gamma_{LH}\left(\psi q_H^*(\underline{b}) + \theta \underline{b}(1 - q_H^*(\underline{b})) + c(q_H^*(\underline{b}))\right) \\ + \gamma_{LL}\left(\psi q_L^*(\underline{b}) + \theta \lambda_L \underline{b}(1 - q_L^*(\underline{b})) + c(q_L^*(\underline{b}))\right) \geq 0.$$

The expected profit function has maximum for a positive value of τ as long as the slope of $\tau\phi$ is steeper than the slope of $F(\tau)$ in $\tau = 0$:

$$\phi > F'(0) \Leftrightarrow \\ \phi > \gamma_{LH}\left(\psi q_H^*(\underline{b}) + \theta \underline{b}(1 - q_H^*(\underline{b})) + c(q_H^*(\underline{b}))\right) \\ + \gamma_{LL}\left(\psi q_L^*(\underline{b}) + \theta \lambda_L \underline{b}(1 - q_L^*(\underline{b})) + c(q_L^*(\underline{b}))\right).$$

Thus, the minimum value of ϕ, which ensures a positive effort level, is

$$\underline{\phi} = \gamma_{LH}\left(\psi q_H^*(\underline{b}) + \theta \underline{b}(1 - q_H^*(\underline{b})) + c(q_H^*(\underline{b}))\right) \\ + \gamma_{LL}\left(\psi q_L^*(\underline{b}) + \theta \lambda_L \underline{b}(1 - q_L^*(\underline{b})) + c(q_L^*(\underline{b}))\right)$$

For project values below $\underline{\phi}$, the expected value of the project will not exceed the costs of providing the agent with incentives to work. As a result, the principal will not find it profitable to incentivize the agent to work. □

A.3 Mutually beneficial principal-agent relations

The following lemma characterizes the conditions under which a mutual beneficial principal-agent relationship arises:

Lemma 4. We distinguish between the two cases described in Result 1. For Case (i), the principal finds it profitable to implement a positive effort level, $\tau^* > 0$, if and only if the project is sufficiently valuable, i.e. $\phi > \overline{\phi} > 0$ with $\Pi(\tau^{min})|_{\phi=\overline{\phi}} = 0$. For Case (ii), the principal implements a positive effort level under the same condition except if $b^{max} < \underline{b}$ where $\underline{b}$ is the bonus level that solves the following equation

$$\begin{aligned} gb &= (1-g)(1-\rho)x\left[\theta b(1-q_H^*(b)) + c(q_H^*(b))\right] \\ &\quad + (1-g)\left(\rho + (1-\rho)(1-x)\right)\left[\theta\lambda_L b(1-q_L^*(b)) + c(q_L^*(b))\right] \end{aligned}$$

In this case, no credible effort level can induce the agent to work and consequently no principal-agent relationship will be established.

Proof. To establish sufficiency, pick some $\phi' > \overline{\phi}$. Since $\frac{\partial \Pi}{\partial \phi} > 0$ it holds that $\Pi(\tau^{min})|_{\phi'} > 0$. Now, $\Pi(\tau = 0) = 0 < \Pi(\tau^{min})|_{\phi'}$ and therefore $\tau^* > 0$.

To show necessity, suppose $\tau^* > 0$. Then it must be the case that $\phi > \overline{\phi}$. $\Pi(\tau)$ is continuous in $\tau \geq 0$ and concave with a unique maximum at $\tilde{\tau} > 0$. Now suppose that $\phi = \phi' < \overline{\phi}$. Then, as a consequence, $\Pi(\tau^{min})|_{\phi'} < 0$. From this we must conclude that $\tilde{\tau} < \tau^{min}$ and $\Pi(\tau) < 0$ for all $\tau \in [\tau^{min}, \tau^{max}]$, which contradicts $\tau^* > 0$.

For Case (ii) the proof is similar. However, if parameters are such that Case (ii) obtains *and* $b^{max} < \underline{b}$ no credible bonus is high enough to induce the agent to work. □

A.4 Comparative Statics of τ^{max} and τ^{min}

The bonuses b^{min} and b^{max} put limits on the effort levels that the principal can implement. Denote by τ^{min} and τ^{max} the optimal effort levels as chosen by the agent when presented with a bonus of b^{min} and b^{max} respectively. The following Result summarizes some comparative statics with regard to the lowest and highest possible effort levels, τ^{min} and τ^{max}, which the principal can implement.

Result 5. (i)$\frac{d\tau^{min}}{d\psi} > 0$ and $\frac{d\tau^{max}}{d\psi} > 0$, (ii) $\frac{d\tau^{min}}{dg} > 0$ and $\frac{d\tau^{max}}{dg} > 0$ if ψ is sufficiently small, (iii) $\frac{d\tau^{min}}{dx} > 0$ and $\frac{d\tau^{max}}{dx} > 0$ if ψ is sufficiently large, (iv) $\frac{d\tau^{max}}{d\theta} > 0$ and $\frac{d\tau^{min}}{d\theta} > 0$ if ψ is sufficiently large, (v) $\frac{d\tau^{max}}{d\rho} > 0$ regardless of the size of ψ and $\frac{d\tau^{min}}{d\rho} > 0$ if ψ is sufficiently small.

Proof. A change in parameters has two effects on τ^{min} and τ^{max}. First, there is an effect through the incentive compatible bonus $b(\tau)$. That is, the price of effort changes. Second, there is an effect through b^{min} and b^{max}. The overall effect will depend on the sign and magnitude of these two effects.

τ^{min} and τ^{max} are implicitly given by

$$b^{max} = (\rho + (1-\rho)x)\,\psi q_H^* + (1-\rho)(1-x)\psi q_L^* = b(\tau^{max})$$
$$b^{min} = (1-\rho)x\psi q_H^* + (\rho + (1-\rho)(1-x))\,\psi q_L^* = b(\tau^{min})$$

These equations will be used to compute the comparative statics of τ^{min} and τ^{max}. Let $F^{min} = b^{min} - b(\tau)$ and $F^{max} = b^{max} - b(\tau)$. Then, for some parameter κ, $\frac{d\tau^{min/max}}{d\kappa} = -\frac{\partial F^{min/max}/\kappa}{\partial F^{min/max}/\tau}$. Notice that $\frac{\partial F^{min/max}}{\partial \tau} < 0$ since $b(\tau)$ is increasing in τ in the relevant range and $b^{min/max}$ is independent of τ.

Part (i). The determining part is the sign of $\frac{\partial F^{min/max}}{\partial \psi}$. Since $b(\tau)$ does not depend on ψ and b^{min} is increasing in ψ it must be the case that $\frac{\partial F^{min}}{\partial \psi} > 0$. Then we can conclude that $\frac{d\tau^{min}}{d\psi} > 0$. The argumentation is the same for $\frac{\partial F^{max}}{\partial \psi} > 0$. Since $\frac{\partial b^{max}}{\partial \psi} > 0$ it holds that $\frac{\partial F^{max}}{\partial \psi} > 0$ and consequently $\frac{d\tau^{max}}{d\psi} > 0$. The parameter ψ reflects the principal's cost of conflict. A higher level of ψ increases both the minimum and the maximum credible bonuses, b^{min} and b^{max}, but it does not change the price of effort, $b(\tau)$. For this reason τ^{min} and τ^{max} are increasing in ψ. Intuitively, when ψ increases the principal will find it less tempting to cheat on the agent by lying since the expected cost of conflict is now higher. He can therefore credibly offer higher bonuses than for lower values of ψ. Likewise, for higher values of ψ the principal will more often prefer to pay out the bonus regardless of state and in this way avoid potential conflict costs. Consequently, there are now some bonuses too low to be credible compared to a situation with a lower value of ψ. Thus, τ^{min} increases.

Part (ii). Because λ_L is decreasing in g, the value of b^{min} and b^{max} will depend on g through the optimal conflict level. Focus first on b^{min} and differentiate this with respect to g

$$\frac{\partial b^{min}}{\partial g} = (\rho + (1-\rho)(1-x))\,\psi \frac{\partial q_L^*}{\partial \lambda_L}\frac{\partial \lambda_L}{\partial g} < 0$$

where the inequality holds since $\frac{\partial q_L^*}{\partial \lambda_L} > 0$ and $\frac{\partial \lambda_L}{\partial g} < 0$.

Now we need to sign the effect on the incentive compatible bonus $b(\tau)$ when g changes. Remember that $b(\tau)$ is implicitly given by

$$\begin{aligned} &b\,(g - (1-g)(1-\rho)x\theta(1-q_H^*) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)) \\ &\quad = v'(\tau) + (1-g)(1-\rho)xc(q_H^*) + (1-g)(\rho + (1-\rho)(1-x))c(q_L^*) \end{aligned} \tag{13}$$

Differentiating both sides with respect to g yields

$$\begin{aligned} &\frac{\partial b}{\partial g}\left[g - (1-g)(1-\rho)x\theta(1-q_H^*) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)\right] \\ &= -\,(\rho + (1-\rho)(1-x))c\,(q_L^*) - b\,(1 + (1-\rho)x\theta(1-q_H^*) - (1-\rho)xc\,(q_H^*)) \\ &\quad - b\,((\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)) + b(1-g)(\rho + (1-\rho)(1-x))\theta(1-q_L^*)\frac{\partial \lambda_L}{\partial g} \end{aligned}$$

Since $\frac{\partial \lambda_L}{\partial g} < 0$ the right hand side is clearly negative. The term in the square brackets on the left hand side is positive for the relevant range of bonuses (see Appendix A.2) and as a result we conclude that $\frac{\partial b}{\partial g}$ is negative for all the relevant bonus levels.

Having determined that both $\frac{\partial b^{min}}{\partial g}$ and $\frac{\partial b}{\partial g}$ are negative, we see that the effects pull in opposite directions (a lower b^{min} pulls in the direction of a lower τ^{min} while a lower $b(\tau)$ makes it more tempting for the principal to evade conflict by reporting H when his true signal is L. This pulls in the direction of a higher τ^{min}). However, $\frac{\partial b^{min}}{\partial g}$ is clearly less negative the smaller is the principal's cost of conflict ψ. Hence, there will exist a $\bar{\psi}$ such that for $\psi < \bar{\psi}$, $\frac{\partial F^{min}}{\partial g} > 0$ and consequently $\frac{d\tau^{min}}{dg} > 0$.

The argument is identical for $\frac{d\tau^{max}}{dg}$. $\frac{\partial b^{max}}{\partial g}$ is negative since $\frac{\partial \lambda_L}{\partial g}$ is negative.

Therefore, again the effects pull in opposite directions. However, there will exist a $\bar{\psi}$ such that for $\psi < \bar{\psi}$, $\frac{\partial F^{max}}{\partial g} > 0$ and consequently $\frac{d\tau^{max}}{dg} > 0$.

Part (iii). $\frac{\partial b^{min}}{\partial x} = (1-\rho)(q_H^* - q_L^*)\psi$ which is positive since $q_H^* > q_L^*$. For the effect on $b(\tau)$ we get

$$\frac{\partial b}{\partial x}\left[g - (1-g)(1-\rho)x\theta(1-q_H^*) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)\right]$$
$$= (1-g)(1-\rho)\left[\theta b(1-q_H^*) + c(q_H^*) - (\theta\lambda_L b(1-q_L^*) + c(q_L^*))\right]$$

The left hand side is positive for the relevant range of bonuses (see Appendix A.2). The right hand side is positive and hence we conclude that $\frac{\partial b}{\partial x} > 0$.

Again, the effects pull in opposite directions. Now, $\frac{\partial b^{min}}{\partial x}$ is clearly more positive the larger is ψ whereas $\frac{\partial b}{\partial x}$ does not depend on ψ. As a consequence there will exist a $\tilde{\psi}$ such that for all $\psi > \tilde{\psi}$ $\frac{\partial F^{min}}{\partial x} > 0$ and therefore $\frac{\partial \tau^{min}}{\partial x} > 0$.

The argument is identical for the case of τ^{max}.

Part (iv). For the case of θ we have

$$\frac{\partial b^{min}}{\partial \theta} = (1-\rho)x\psi\frac{\partial q_H^*}{\partial \theta} + (\rho + (1-\rho)(1-x)\psi\frac{\partial q_L^*}{\partial \theta} > 0$$
$$\frac{\partial b^{max}}{\partial \theta} = (\rho + (1-\rho)x)\psi\frac{\partial q_H^*}{\partial \theta} + (1-\rho)(1-x)\psi\frac{\partial q_L^*}{\partial \theta} > 0$$

Where the inequalities hold since $\frac{\partial q_i^*}{\partial \theta} > 0$.

For the effect through $b(\tau)$ we have

$$\frac{\partial b}{\partial \theta}\left[g - (1-g)(1-\rho)x\theta(1-q_H^*) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)\right]$$
$$= (1-g)(1-\rho)x(1-q_H^*) + (1-g)\left(\rho - (1-\rho)(1-x)\right)(1-q_L^*)$$

The term in the square brackets on the left hand side is positive for the relevant range of bonuses (see Appendix A.2). Furthermore, the right hand side is positive. As a result we can conclude that $\frac{\partial b}{\partial \theta} > 0$.

Again, the effects pull in opposite directions. Now, $\frac{\partial b^{min}}{\partial \theta}$ is clearly more positive the larger is ψ whereas $\frac{\partial b}{\partial \theta}$ does not depend on ψ. As a consequence there will exist a $\tilde{\psi}$ such that for all $\psi > \tilde{\psi}$ $\frac{\partial F^{min}}{\partial \theta} > 0$ and therefore $\frac{\partial \tau^{min}}{\partial \theta} > 0$.

The argument is identical for the case of τ^{max}.

Part (v). For the effect of a change in ρ we have

$$\frac{\partial b^{min}}{\partial \rho} = x\psi\left(q_L^* - q_H^*\right) < 0$$
$$\frac{\partial b^{max}}{\partial \rho} = (1-x)\psi\left(q_H^* - q_L^*\right) > 0$$

where the inequalities hold since $q_H^* > q_L^*$.

For the effect through the incentive compatible bonus we see that

$$\frac{\partial b}{\partial \rho}\left[g - (1-g)(1-\rho)x\theta(1-q_H^*) - (1-g)(\rho + (1-\rho)(1-x))\theta\lambda_L(1-q_L^*)\right]$$
$$= (1-g)xc\left[\theta\lambda_L b(1-q_L^*) + c(q_L^*) - (\theta b(1-q_H^*) + c(q_H^*))\right]$$

The right hand side is negative since $q_H^* > q_L^*$. The term in the square bracket on the left hand side is positive for the relevant range of bonus levels (see Appendix A.2). As a result, $\frac{\partial b}{\partial \rho}$ is negative.

For the case of τ^{max} the effects pull in the same direction. A higher level of ρ makes effort cheaper and also makes truthtelling more attractive in the case where the principal receives signal H. Therefore we conclude that $\frac{\partial F^{max}}{\partial \rho} > 0$ and as a result $\frac{\partial \tau^{max}}{\partial \rho} > 0$.

For τ^{min} the result is ambiguous since $\frac{\partial b^{min}}{\partial \rho} < 0$ but also $\frac{\partial b}{\partial \rho} < 0$. However, if $\frac{\partial b^{min}}{\partial \rho}$ is not too negative - which will hold if ψ is not too large - we will have $\frac{\partial F^{max}}{\partial \rho} > 0$ and therefore $\frac{\partial \tau^{min}}{\partial \rho} > 0$.

□

A.5 Proof of Proposition 1

The effect of a change in some parameter κ on equilibrium profits is given by

$$\frac{d\Pi(\tau^*)}{d\kappa} = \frac{\partial\Pi(\tau^*)}{\partial\kappa} + \frac{\partial\Pi(\tau^*)}{\partial\tau}\frac{d\tau^*}{d\kappa} \tag{14}$$

For the direct effect $\frac{\partial\Pi(\tau^*)}{\partial\psi}$ we know that

$$\frac{\partial\Pi}{\partial\psi} = -\tau\left[\gamma_{LH}q_H^* + \gamma_{LL}q_L^*\right] < 0$$

That is, the direct effect of an increase in ψ is negative.

For the effect through the effort level we now focus on the second term of equation 14, $\frac{\partial\Pi(\tau^*)}{\partial\tau}\frac{d\tau^*}{d\psi}$. This effect will be zero when the chosen effort level is optimal, i.e. $\tau^* = \tilde{\tau}$. However, if the principal is bounded by the upper or lower truthtelling constraint he implements τ^{max} or τ^{min} respectively and $\frac{\partial\Pi(\tau^*)}{\partial\tau}\frac{d\tau^*}{d\psi}$ will be different from zero. Investigating first $\frac{\partial\Pi(\tau)}{\partial\tau}$ we see that

$$\begin{aligned}
\frac{\partial\Pi(\tau)}{\partial\tau} \quad = \quad & \phi - v'(\tau) - \gamma_{LH}\left(\psi q_H^* + \theta b(1-q_H^*) + xc(q_H^*)\right) \\
& -\tau\gamma_{LH}\left[\psi\frac{\partial q_H^*}{\partial b}\frac{\partial b}{\partial\tau} + \theta(1-q_H^*)\frac{\partial b}{\partial\tau}\right] - \gamma_{LL}\left(\psi q_L^* + \theta\lambda_L b(1-q_L^*) + xc(q_L^*)\right) \\
& -\tau\gamma_{LL}\left[\psi\frac{\partial q_L^*}{\partial b}\frac{\partial b}{\partial\tau} + \theta\lambda_L(1-q_L^*)\frac{\partial b}{\partial\tau}\right].
\end{aligned}$$

Notice that for a fixed τ, $\frac{\partial\Pi(\tau)}{\partial\tau}$ is linearly increasing in ϕ. For a fixed value of ϕ the derivative is decreasing in τ.

Now recall from Result 5 that $\frac{d\tau^{max}}{d\psi} > 0$. Fix any $\tau^{max} \in (0,1)$. Then, there exists a ϕ' such that $\frac{\partial\Pi(\tau)}{\partial\tau}|_{\tau=\tau^{max}} > 0$ and $\tau^* = \tau^{max}$ for all $\phi > \phi'$. Then the indirect effect is negative, $\frac{\partial\Pi(\tau)}{\partial\tau}|_{\tau=\tau^*}\frac{d\tau^*}{d\psi} > 0$. Since $\frac{d\tau^{max}}{d\psi}$ and $\frac{\partial\Pi(\tau)}{\partial\psi}$ do not depend on ϕ and $\frac{\partial\Pi(\tau)}{\partial\tau}$ is linearly increasing in ϕ there exist a ϕ'' such that $\frac{d\Pi(\tau^*)}{d\psi} > 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$ meaning that the second effect through the implemented effort level dominates and welfare is increasing in the principal's cost of conflict.

A.6 Proof of Proposition 2

Similar to the proof of Proposition 1.

Again, we need to determine the signs of the direct and the indirect effects.

For the direct effect $\frac{\partial \Pi(\tau)}{\partial \theta}$ we have

$$\begin{aligned}\frac{\partial \Pi}{\partial \theta} =& -\tau \gamma_{LH} b(1-q_H^*(b)) - \tau\gamma_{LH}(\psi \tfrac{\partial q_H^*(b)}{\partial b}\tfrac{\partial b}{\partial \theta} + (1-q_H^*(b))\tfrac{\partial b}{\partial \theta}) \\ & - \tau\gamma_{LL}\lambda_L b(1-q_L^*(b)) - \tau\gamma_{LL}(\psi \tfrac{\partial q_L^*(b)}{\partial b}\tfrac{\partial b}{\partial \theta} + (1-q_L^*(b))\tfrac{\partial b}{\partial \theta}).\end{aligned}$$

From equation 4 we know that $\frac{\partial q_H^*(b)}{\partial b} > 0$ and from part (iv) of Appendix A.4 we know that $\frac{\partial b}{\partial \theta} > 0$. From this we conclude that $\frac{\partial \Pi}{\partial \theta} < 0$.

Part (i) For the sign of the indirect effect, fix any $\tau^{max} \in (0,1)$ with a ψ in accordance with Result 5 such that $\frac{d\tau^{max}}{d\theta} > 0$. There exist a project value ϕ' such that $\tau^* = \tau^{max}$. Since $\frac{d\tau^{max}}{d\theta}$ and $\frac{\partial \Pi(\tau)}{\partial \theta}$ do not depend on ϕ and $\frac{\partial \Pi(\tau)}{\partial \tau}$ is linearly increasing in ϕ there exist a ϕ'' such that $\frac{d\Pi(\tau^*)}{d\theta} > 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

Part (ii) If the principal is bounded by the lower truthtelling constraint we have $\frac{\partial \Pi(\tau^*)}{\partial \tau} < 0$. There exist a project value $\tilde{\phi}$ such that $0 < \tilde{\tau} < \tau^{min}$ and therefore $\tau^* = \tau^{min}$ for all $\phi < \tilde{\phi}$. Then fix a ψ in accordance with Result 5 such that $\frac{d\tau^{min}}{d\theta} < 0$.

Now, notice that $\frac{\partial \Pi(\tau^*)}{\partial \tau}$ is decreasing in τ for a fixed ϕ and is more negative the larger is $v'(\tau)$. Further, $\frac{d\tau^{min}}{d\theta}$ is larger, the smaller is $v''(\tau)$. Hence, we have that $\frac{\partial \Pi(\tau)}{\partial \theta}$ is independent of $v(\tau)$ and its derivatives whereas $\frac{\partial \Pi(\tau)}{\partial \tau}\frac{d\tau^{min}}{d\theta}$ is increasing in $\frac{v'(\tau)}{v''(\tau)}$. Fix a positive real number z. Then there exists an effort cost function $v(\tau)$ such that $\frac{v'(\tau)}{v''(\tau)} > z$. Hence, $\frac{d\Pi(\tau)}{d\theta} > 0$ if z is sufficiently large.

A.7 Proof of Proposition 3

Similar to the proof of Proposition 1. First we investigate the sign of $\frac{\partial \Pi(\tau)}{\partial \rho}$.

$$\frac{\partial \Pi}{\partial \rho} = -\tau \frac{d\gamma_{LH}}{d\rho}(\psi q^* + \theta\lambda b(1-q^*) + c(q^*)) - \tau\gamma_{LH}\left(\psi \frac{\partial q^*}{\partial b}\frac{\partial b}{\partial \rho} + \theta\lambda(1-q^*)\frac{\partial b}{\partial \rho}\right) > 0$$

where the inequality holds since $\frac{d\gamma_{LH}}{d\rho}$ is negative, we know from equation 4 that $\frac{\partial q^*}{\partial b}$ is positive and from Appendix A.4 that $\frac{\partial b}{\partial \rho}$ is negative.

Now, fix any $\tau^{min} \in (0,1)$ with a ψ small enough such that $\frac{d\tau^{min}}{d\rho} > 0$ and a positive real number z. There exist a project value $\tilde{\phi}$ such that $\tau^* = \tau^{min}$ and there exists an effort cost function such that $\frac{v'(\tau^{min})}{v''(\tau^{min})} > z$. Notice that $\frac{\partial \Pi(\tau)}{\partial \rho}$ is independent of $v(\tau)$ and its derivatives. Furthermore, $\frac{\partial \Pi(\tau)}{\partial \tau}|_{\tau=\tau^{min}} < 0$ and

$\frac{\partial \Pi(\tau)}{\partial \tau}\frac{d\tau^{min}}{d\rho}$ is increasingly negative in $\frac{v'(\tau)}{v''(\tau)}$. For this reason there exists a z large enough such that $\frac{d\Pi(\tau)}{d\rho}$ is negative.

A.8 Proof of Proposition 4

Similar to the proof of Proposition 1.

For the direct effect we have

$$\frac{\partial \Pi}{\partial x} = -\tau\frac{d\gamma_{LH}}{dx}\left(\psi q^* + \theta\lambda b(1-q^*) + c(q^*)\right) - \tau\gamma_{LH}\left(\psi\frac{\partial q^*}{\partial b}\frac{\partial b}{\partial x} + \theta\lambda(1-q^*)\frac{\partial b}{\partial x}\right) < 0$$

where the inequality holds since $\frac{d\gamma_{LH}}{dx} > 0$ is positive, $\frac{\partial q^*}{\partial b} > 0$ (equation 4) and $\frac{\partial b}{\partial x} > 0$ (Appendix A.4).

Part (i) For the indirect effect fix any $\tau^{min} \in (0,1)$ with a ψ in accordance with Result 5 such that $\frac{d\tau^{min}}{dx} < 0$ and a positive real number z. There exist a project value ϕ' such that $\tau^* = \tau^{min}$ and there exists an effort cost function such that $\frac{v'(\tau^{min})}{v''(\tau^{min})} > z$. Notice that $\frac{\partial \Pi(\tau)}{\partial x}$ is independent of $v(\tau)$ and its derivatives. Furthermore, $\frac{\partial \Pi(\tau)}{\partial \tau}|_{\tau=\tau^{min}} < 0$ and $\frac{\partial \Pi(\tau)}{\partial \tau}\frac{d\tau^{min}}{dx}$ is increasingly negative in $\frac{v'(\tau)}{v''(\tau)}$. For this reason there exists a z large enough such that $\frac{d\Pi(\tau)}{dx}$ is positive.

Part (ii) For the indirect effect fix any $\tau^{max} \in (0,1)$ with a ψ in accordance with Result 5 such that $\frac{d\tau^{max}}{dx} > 0$. There exist a project value ϕ' such that $\tau^* = \tau^{max}$. Since $\frac{d\tau^{max}}{dx}$ and $\frac{\partial \Pi(\tau)}{\partial x}$ do not depend on ϕ and $\frac{\partial \Pi(\tau)}{\partial \tau}$ is linearly increasing in ϕ there exist a ϕ'' such that $\frac{d\Pi(\tau^*)}{dx} > 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

A.9 Proof of Proposition 5

The effect of a change in g on equilibrium profits is given by

$$\frac{d\Pi(\tau^*)}{dg} = \frac{\partial \Pi(\tau^*)}{\partial g} + \frac{\partial \Pi(\tau^*)}{\partial \tau}\frac{d\tau^*}{dg}.$$

Part (i) When the principal is not bounded by either truthtelling constraint it holds that $\frac{\partial \Pi(\tau^*)}{\partial \tau} = 0$ and hence the total welfare effect from a change in g will be given by the first term $\frac{\partial \Pi(\tau^*)}{\partial g}$. The sign of this term is investigated below.

$$\frac{\partial \Pi}{\partial g} = -\tau \frac{d\gamma_{LH}}{dg}\left(\psi q_H^* + \theta b(1-q_H^*) + c(q_H^*)\right) - \tau\gamma_{LH}\left(\psi \frac{\partial q_H^*}{\partial b}\frac{\partial b}{\partial g} + \theta(1-q_H^*)\frac{\partial b}{\partial g}\right)$$
$$- \tau \frac{d\gamma_{LL}}{dg}\left(\psi q_L^* + \theta\lambda_L b(1-q_L^*) + c(q_L^*)\right)$$
$$- \tau\gamma_{LL}\left(\psi \frac{\partial q_L^*}{\partial b}\frac{\partial b}{\partial g} + \theta\lambda_L(1-q_L^*)\frac{\partial b}{\partial g} + \theta b(1-q_L^*)\frac{\partial \lambda_L}{\partial g}\right) \quad (15)$$

We know that $\frac{\partial\gamma_{LH}}{\partial g} < 0$, $\frac{\partial\gamma_{LL}}{\partial g} < 0$ and $\frac{\partial q_{s_A}^*}{\partial b} > 0$. From Appendix A.4 we know that $\frac{\partial b}{\partial g} < 0$. Thus we conclude that $\frac{\partial \Pi}{\partial g}$ is unambiguously positive. As a result, welfare is increasing in the quality of the principal's signal when he is not bounded by truthtelling constraints.

Part (ii)(a) We have just shown that $\frac{\partial \Pi}{\partial g} > 0$. Now fix any $\tau^{max} \in (0,1)$ with a ψ in accordance with Result 5 such that $\frac{d\tau^{max}}{dg} < 0$. There exists a project value ϕ' such that $\frac{\partial \Pi(\tau)}{\partial \tau}|_{\tau=\tau^{max}} > 0$. Then $\frac{\partial \Pi(\tau)}{\partial \tau}|_{\tau=\tau^{max}}\frac{d\tau^{max}}{dg} < 0$. Since $\frac{d\tau^{max}}{dg}$ and $\frac{\partial \Pi(\tau)}{\partial g}$ are independent of ϕ and $\frac{\partial \Pi(\tau)}{\partial \tau}$ is linearly increasing in ϕ for a fixed $\tau = \tau^{max}$, there exists a project value ϕ'' such that $\frac{d\Pi(\tau^*)}{dg} < 0$ for all $\phi > \tilde{\phi} \equiv \max(\phi', \phi'')$.

Part (ii)(b) Fix any $\tau^{min} \in (0,1)$ with a ψ in accordance with Result 5 such that $\frac{d\tau^{min}}{dg} > 0$ and a positive real number z. There exist a project value ϕ' such that $\tau^* = \tau^{min}$ and there exists an effort cost function such that $\frac{v'(\tau^{min})}{v''(\tau^{min})} > z$. Notice that $\frac{\partial \Pi(\tau)}{\partial g}$ is independent of $v(\tau)$ and its derivatives. $\frac{\partial \Pi(\tau)}{\partial \tau}\frac{d\tau^{min}}{dg}$ on the other hand is increasingly negative in $\frac{v'(\tau)}{v''(\tau)}$. For this reason there exists a z large enough such that $\frac{d\Pi(\tau)}{dg}$ is negative.