

TI 2013-196/VII
Tinbergen Institute Discussion Paper



Leniency Bias in Long-Term Workplace Relationships

Jan Tichem

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Leniency Bias in Long-Term Workplace Relationships*

Jan Tichem[†]

December 11, 2013

Abstract

This paper studies how firms can efficiently incentivize supervisors to truthfully report employee performance. To this end, I develop a dynamic principal-supervisor-agent model. The supervisor is either selfish or altruistic towards the agent, which is observable to the agent but not to the principal. The analysis yields two key results. First, supervisor altruism sometimes provides a net incentive to report performance truthfully, rather than to bias evaluations upward. The intuition is that an altruistic supervisor values his job because of his good relationship with the agent, and puts his job at risk by overrating the agent's performance. Second, I show that by screening for one supervisor type, firms can incentivize the supervisor to truthfully report performance at the lowest possible costs. For this reason, screening may be optimal, even though it reduces the probability that vacancies are filled.

Keywords: Altruism, incentives, leniency bias, screening, subjective performance evaluation, supervisor.

JEL Classification: D86, J33, M52, M55

*I gratefully acknowledge helpful comments and suggestions by Robert Dur, Michiel Souverijn, Robin Zoutenbier, and participants to the ESE Brown Bag Seminar and TI PhD Lunch Seminar.

[†]Department of Economics, Erasmus University Rotterdam, and Tinbergen Institute.
E-mail: tichem@ese.eur.nl.

1 Introduction

Many firms incentivize their employees on the basis of subjective performance evaluation (MacLeod and Malcomson 1998, Prendergast 1999, MacLeod and Parent 2000, Gibbs et al. 2004, Gibbs 2012). Moreover, the actual appraisal of employee performance is often not conducted by firm owners, but delegated to supervisors. There is ample evidence that supervisors exert bias in appraising performance.¹ Such bias may be harmful to firms because it weakens the link between employees' actual performance and their performance evaluation. In turn, this reduces employees' incentives to exert effort. Firms therefore have an interest in stimulating supervisors to report performance truthfully. This paper develops and analyzes a model to study how firms can achieve this through designing optimal contracts for supervisors and employees.

In the model, the agent chooses a privately observable, continuous effort level that generates a high or low performance. To induce the agent to exert effort, the principal wishes to pay the agent a bonus if his performance is high. The agent's performance is not verifiable, but can be subjectively evaluated. Since the principal does not observe the agent's performance herself, she hires a supervisor to perform this task. The supervisor is either selfish or altruistic towards the agent, and may therefore bias performance evaluations. Supervisor's altruism is observed by the agent but not by the principal. For simplicity, I focus on the case where the unconditional probability that the supervisor is altruistic equals one half.

The agency problem is studied in an infinitely repeated game. As is usual in the repeated games literature, the agent is assumed to play a trigger strategy that prescribes to shirk in all future periods if the supervisor has once biased his performance evaluation (see e.g. Bull 1987). Next, as in Baker et al. (1994), I assume there is a verifiable signal about the agent's performance available to the principal. The verifiable signal is necessary for incentivizing the supervisor to report performance truthfully, as will be explained below. Allowing for objective incentives for the agent on the basis of this verifiable signal does not change the results of the analysis qualitatively, and so for simplicity I abstract from this in the analysis. Further, I assume there is a large pool of supervisors and agents. Therefore, if a current labor relationship

¹Well-known biases are the leniency bias and the centrality bias. See for example Medoff and Abraham (1980), Jawahar and Williams (1997), Prendergast (1999), Moers (2005), Berger et al. (2011), and Bol (2011).

breaks down, the principal will be rematched to a new agent and/or supervisor in the next period. Finally, throughout the paper I restrict attention to the case where the principal offers a uniform contract to the supervisor.

The analysis yields two key results. First, given the optimal contracts, supervisor's altruism may provide a net incentive to report performance truthfully, rather than to bias evaluations. Second, for relatively low and relatively high values of the discount factor, the principal optimally designs the supervisor's contract such that it only attracts altruistic types. In the remainder of the Introduction, the results are described and explained in more detail.

In case the supervisor is altruistic, he enjoys to some extent the well-being of the agent. Therefore, the altruistic supervisor has an incentive to report the agent's performance to be high, irrespective of the agent's actual performance. The reason is that reporting high performance will yield the agent a valuable bonus. However, the agent only has an incentive to exert effort if he expects the supervisor to report performance truthfully. In the repeated game, the supervisor will do so only if the following holds: reporting performance truthfully yields future rents, the discounted value of which exceeds the current benefit from overreporting. To satisfy this condition in the most efficient way, the principal optimally uses the following three compensation elements.

First, note that the lower the altruistic supervisor's current benefit from overreporting, the lower the rents he requires to report performance truthfully. The principal therefore optimally makes it costly to the supervisor to report high performance. This is achieved by tying the supervisor's compensation to his evaluation report, as in Giebe and Gürtler (2012). Clearly, a monetary penalty following a high performance rating makes it attractive for the selfish supervisor type to bias performance ratings downward. This is costly to the principal since now the selfish type must be paid rents to report performance truthfully. The optimal penalty for reporting high performance therefore trades off the rents paid to the selfish type against the rents paid to the altruistic type, and depends on whether or not the principal screens for a supervisor type, as will become clear below.

Second, to ensure that reporting performance truthfully yields future rents, the supervisor receives a monetary reward if the verifiable signal about the agent's performance is high. To understand why, note that the agent responds to a biased performance evaluation by shirking in the future. If the agent shirks, the verifiable signal will be low. Therefore, biasing the agent's performance evaluation reduces the supervisor's compensation in the future.

Third, to maximize the supervisor's incentives to report performance truthfully, his utility after biasing a performance evaluation is set as low as possible. This is achieved by paying the supervisor a low base salary. The base salary can be set as low as necessary. This holds because a low base salary can be compensated for by paying the supervisor a higher reward if the verifiable signal about the agent's performance is high. Because the supervisor can always take his outside option, the implication is that his utility after biasing an evaluation optimally equals at most his reservation utility.

In equilibrium, the principal sets the three compensation elements discussed above such that the supervisor's uniform contract is attractive to both types, or only to the altruistic type. In the former case, I say that the supervisor's contract is a *pooling contract*, whereas in the latter case I say that the supervisor's contract is a *screening contract*. The crucial feature of the screening contract is that it reveals the supervisor's type. This information is valuable to the principal because it enables her to eliminate the supervisor's current benefit from biasing the performance evaluation. This is achieved by setting the right penalty for giving the agent a high performance evaluation. The benefit to the principal is that the supervisor does not require a rent to report performance truthfully. Further, since in this case inducing correct performance evaluations from the supervisor does not generate agency costs, it is also optimal to provide the agent the first-best level of the bonus. Screening for one supervisor type thus yields the principal first-best profits, conditional on being matched. Assuming that the principal optimally screens, it holds that she optimally screens for the altruistic type. The reason is that, in contrast to the selfish type, the altruistic type enjoys working with the agent. The altruistic type therefore demands a lower total compensation level to accept the job. This feature also explains why the screening contract does not attract the selfish type.

Next, if the principal offers the pooling contract to the supervisor, at least one supervisor type enjoys a benefit from biasing the performance evaluation in the current period. This is inevitable because one type is selfish and the other is altruistic, but the penalty for giving a high performance evaluation can only take one value. It follows that at least one type must enjoy future rents to be willing to report performance truthfully. Interestingly, the altruistic supervisor enjoys future rents from reporting truthfully precisely because he is altruistic. The reason is that he enjoys altruistic utility from working with the agent. This utility cannot be extracted from the altruistic type, since doing so would make the job unattractive to the selfish type. It

follows that, as long as he reports performance truthfully, the altruistic type's utility is strictly above his reservation utility. However, as explained above, the optimal contract ensures that any type of supervisor earns at most his reservation utility after biasing a performance evaluation. Biasing an evaluation thus implies that the altruistic type will lose the future rents derived from working with the agent. Therefore, supervisor altruism not only gives an incentive to overreport performance, but it also provides an incentive to report performance truthfully. Moreover, if the supervisor is sufficiently patient, the latter effect outweighs the former. In that case, the net incentive from supervisor altruism is to report performance truthfully, rather than to bias performance evaluations.

The optimal contracts in case the principal attracts both supervisor types are as follows. If the supervisor is very patient, then even if the agent's bonus is high, the altruistic type prefers to report performance truthfully, as described above. The principal therefore only needs to incentivize the selfish type to report performance truthfully. By the same intuition as in the screening case, the optimal way to do so is by letting the supervisor's compensation be independent of the supervisor's evaluation report. This ensures that the selfish type is indifferent to his evaluation report, and hence does not have to be paid costly rents for rating truthfully. As in the screening case, agency costs from subjective performance evaluation are now fully avoided, implying it is optimal to provide the agent with the first-best level of the bonus.

If the supervisor is less patient, and the agent's bonus is high, the altruistic supervisor prefers to bias performance evaluations. The principal can respond to this situation by reducing the agent's bonus, such that supervisor altruism again provides a net incentive to report performance truthfully. Alternatively, she may increase the supervisor's compensation to ensure reporting truthfully yields sufficient rents. The principal can also pursue both strategies simultaneously. However, in any case it holds that the lower the supervisor's patience, the lower the principal's profits. Therefore, if the supervisor is sufficiently impatient, it becomes optimal to screen. Screening is costly because the supervisor may reject the job offer, in which case the principal earns his reservation utility. Yet, these costs are outweighed by the benefits from increasing the agent's incentives, and eventually reducing the supervisor's compensation, conditional on the supervisor accepting his contract. I call this the 'incentive motive' for screening.

Finally, screening is also optimal if players are sufficiently patient. This may seem counterintuitive, since it has just been explained that attracting

both supervisor types yields high profits in this case. However, if the principal does not screen, she cannot enjoy the compensating wage differential that hiring only the altruistic supervisor gives rise to. If the principal is sufficiently patient, she is willing to incur the costs of screening to enjoy higher profits once having found an altruistic supervisor. I call this the ‘wage differential motive’ for screening.

The paper is organized as follows. Section 2 discusses the related literature. Section 3 describes the model, which is solved in section 4. Section 5 finishes with concluding remarks.

2 Related literature

Supervisor altruism as a source of bias in performance evaluation has received considerable attention. A seminal paper in the field is Prendergast and Topel (1996). The authors show that in response to biased performance ratings, firms may optimally reduce incentives for agents, limit authority of supervisors, and use bureaucratic rules in pay and promotion decisions. Lee and Persson (2011) extend Prendergast and Topel (1996) by allowing the agent and the supervisor to be mutually altruistic. They show that supervisor’s altruism induces leniency, whereas agent’s altruism induces loyalty. Moreover, the benefits from the agent’s loyalty may outweigh the negative effects from the supervisor’s leniency, such that the principal is better off with good social relationships in the workplace. Another extension of Prendergast and Topel (1996) is Grund and Przemec (2012), who consider a situation where two inequality averse agents must be supervised. In this case, the altruistic supervisor not only inflates ratings, but also compresses ratings in order to reduce inequality between the agents.²

As already noted by Prendergast and Topel (1996), incentives can be used to align supervisor’s behavior with the firm’s interest, rather than with his social preferences. This claim receives empirical support from Bandiera et

²Other supervisor preferences potentially causing distorted ratings have been analyzed in the economic literature. Müller and Weinschenk (2011) derive optimal contracts when the supervisor suffers from the ‘horns and halo’ effect, that is, the tendency to give performance evaluations close to the ones given in the past. Golman and Bhatia (2012) assume the supervisor observes noisy signals of the agent’s performance, and feels worse about unfavorable mistakes compared to favorable ones. Kamphorst and Swank (2012) analyze supervisor’s rating behavior when the supervisor simultaneously wants to strengthen the agent’s self-confidence, and come across as a capable evaluator of the agent’s performance.

al. (2009). In a field experiment, the authors exogenously vary supervisors' compensation from a fixed wage to a bonus based on the productivity of supervised workers. Bandiera et al. (2009) report that supervisors respond to this change by assisting workers in a more firm-efficient way, rather than assisting the workers they are socially connected to. In line with this, Berger et al. (2011) find that supervisors who share in company profits or receive performance pay make better promotion decisions. Studying this issue theoretically, Giebe and Gürtler (2012) derive optimal incentive contracts for employees and supervisors when it is the supervisor's task to evaluate the agent's performance. In their model, the supervisor may be either selfish or altruistic towards the agent, which is private information as in Prendergast and Topel (1996). The authors show that inducing correct performance evaluations from the altruistic type may only be possible at the cost of reducing the agent's incentives, and paying the supervisor costly rents. For this reason, the principal may optimally choose not to eliminate leniency from the altruistic supervisor at all. In this case the agent will still exert effort because with some probability she will face the selfish supervisor, who reports performance truthfully.

The present paper contributes to the existing literature in the following ways. First, the model developed and analyzed here is dynamic which enables me to study workplace relationships where interaction is repeated over time. This feature of the model gives rise to the first key result, namely that supervisor altruism may provide a net incentive to report performance truthfully, rather than to bias performance evaluations. Second, as is realistic in long-term workplace relationships, I assume that the agent knows the supervisor's type, and therefore knows whether or not the supervisor will be lenient in rating his performance. Last, this paper highlights an optimal response to the possibility of supervisor bias in performance evaluations that has not been considered before, namely to attract only one supervisor type. A screening contract reveals the supervisor's type, which consequently enables the principal to more efficiently incentivize the supervisor to report performance truthfully. To the best of my knowledge, this is the first study to analyze optimal screening for supervisor types.³

³Since the seminal analysis by Spence (1973), screening for productive agents has become a well-known concept. Recently, screening for agents with some social preference also has attracted quite some attention. See for example Sliwka (2007), Delfgaauw and Dur (2007, 2008), Kosfeld and Von Siemens (2010, 2011), Non (2012), and Von Siemens (2011, 2012).

By studying an infinitely repeated game, this paper is also related to the literature on relational contracts (see e.g. Bull 1987, MacLeod and Malcolmson 1989 and 1998, Baker et al. 1994, and Levin 2003). Closely related are Cordero Salas and Roe (2012) and Dur and Tichem (forthcoming). Both studies show that a more altruistic principal finds it easier to credibly promise a bonus to the agent. Similarly, the analysis here shows that supervisor altruism may provide an incentive to report performance truthfully. In contrast to the present paper, the relational contracts literature generally assumes that a residual claimant principal evaluates the agent’s performance. An exception is Thiele (2013), who studies a principal-supervisor-agent-model that allows for the possibility of collusion between the supervisor and the agent. This analysis differs from Thiele (2013) in that I abstract from collusion. Instead, it is assumed that bias in performance evaluations stems only from the supervisor’s altruism towards the agent. Also, I study the role of uncertainty about the supervisor’s type to the principal, and assume there is verifiable information about the agent’s performance available to the principal.

Finally, Sol (2010) studies the use of peer evaluations to incentivize employees, and allows peers to be altruistic or spiteful towards each other. Altruism and spite lead peers to internalize part of their co-workers’ well-being, which implies that peers have an incentive to bias performance evaluations. Sol (2010) shows that the principal can ensure truthful peer evaluations by reducing the bonus tied to a positive evaluation, which is sometimes found to be optimal in the present analysis as well. Sol (2010) also shows that if peer evaluation becomes severely constrained because feelings of altruism or spite are strong, the principal optimally uses team incentives in addition to peer evaluation.

3 The model

Consider a world with one principal, many agents and many supervisors. All players live an infinite number of periods, denoted by $t = 0, 1, 2, \dots$. Future periods are discounted at a common discount factor δ . Each period the principal may employ one agent and one supervisor. When employed by the principal, the agent chooses a privately observable effort level, e . Effort yields a non-verifiable performance, denoted by $V \in \{0, 1\}$. The probability that the agent’s performance is high is $\Pr[V = 1] = e$. V is observed by the

agent and the supervisor, but not by the principal.⁴ The agent's effort also produces a verifiable signal $z \in \{0, 1\}$, which is modeled in the same way as in Baker et al. (1994). The probability that the verifiable signal takes a high value is $\Pr[z = 1] = \mu e$. The parameter μ has an expectation $E[\mu] = 1$, a variance $\sigma_\mu^2 > 0$, and μ is privately observed by the agent before exerting effort. The costs of exerting effort are given by $\frac{1}{2}\theta e^2$. Note that the first-best effort level is therefore equal to $e^{FB} = \frac{1}{\theta}$. To ensure that the probability that $V = 1$ is properly defined, it is assumed that $\theta \geq 1$. It is also assumed that θ and the support of μ are such that $\Pr[z = 1] \in [0, 1]$.

In the analysis, I assume that the principal incentivizes the agent only on the basis of subjective evaluation of the agent's performance V .⁵ Since the principal does not observe the agent's performance herself, it is the supervisor's task to evaluate the agent's performance. The supervisor's evaluation report is denoted by $d \in \{0, 1\}$, where $d = 0$ means " $V = 0$ ", and $d = 1$ means " $V = 1$ ". Importantly, the principal can commit to reward the agent on the basis of the supervisor's report.⁶ The principal can therefore offer the agent a fully enforceable contract, which consists of a base salary a , and a bonus b to be paid if the supervisor reports performance is high ($d = 1$). The agent's utility in period t can now be written as:

$$U_A = w_A - \frac{1}{2}\theta e^2, \quad (1)$$

where $w_A \equiv a + db$ is the agent's total compensation. The agent's reservation

⁴It is a feature of many large firms that firm-owners do not observe the performance of individual employees. See also e.g. Prendergast and Topel (1996) and Giebe and Gürtler (2012), who make the same assumption.

⁵Sometimes the principal can attain strictly higher profits by also incentivizing the agent on the basis of the verifiable signal z (namely in cases II and III of Proposition 2). However, if σ_μ^2 is sufficiently large but finite, the key results of the paper are qualitatively the same as in the case where the principal only uses subjective incentives for the agent. A formal proof of this claim is provided in Appendix B. The intuition behind it is as follows (also see Baker et al. 1994). If the agent is incentivized on the basis of the verifiable signal z , he will vary his effort level with the realization of the parameter μ . However, variation in the agent's effort level is costly because the agent's effort costs are convex. For these costs the agent must be compensated through his wage. Therefore, if σ_μ^2 is sufficiently high, using objective incentives is so costly such that it is always optimal to also include subjective incentives in the agent's contract, in which case all results of the paper hold.

⁶This is not a strong assumption, since the supervisor's report d can easily be made verifiable information (for example, d can be a written report). See also e.g. Prendergast and Topel (1996), Giebe and Gürtler (2012), and Thiele (2013), who make the same assumption.

utility equals $\overline{U}_A > 0$. Following Baker et al. (1994), the agent is assumed to play a trigger strategy that prescribes to always shirk if the supervisor once evaluated performance incorrectly ($d \neq V$).

The supervisor's utility in period t is given by:

$$U_S = w_S + \gamma U_A, \quad (2)$$

where w_S denotes the supervisor's compensation, and γ denotes the supervisor's altruism towards the agent. The supervisor's compensation is defined as $w_S \equiv \alpha + \eta z - \beta db$. The variable α is the supervisor's base salary. β ties the supervisor's compensation to his evaluation of the agent's performance. Finally, η is a bonus which is paid out if the verifiable signal about the agent's performance, z , is high ($z = 1$). Note that the supervisor's contract is also fully enforceable.

The supervisor may have standard selfish preferences, in which case $\gamma = \underline{\gamma} = 0$. Alternatively, the supervisor may be altruistic towards the agent, in which case $\gamma = \overline{\gamma} > 0$. It holds that $\overline{\gamma} < 1$, implying the altruistic supervisor always cares more for his own compensation than for the agent's utility. I assume that the supervisor's altruism is specific to the agent, and not a general trait towards all people. It follows that the supervisor's reservation utility, $\overline{U}_S > 0$, is independent of his type.⁷ Further, the agent and the supervisor learn γ before they accept or reject the principal's job offer.⁸ Also, it is assumed that the principal does not observe γ .⁹ Finally, for simplicity the probability that the supervisor is altruistic towards the agent equals $\Pr[\gamma = \overline{\gamma}] = \frac{1}{2}$, which is common knowledge.

The principal's profits are given by the agent's performance V , minus the agent's and the supervisor's compensation:

$$\Pi = V - w_S - w_A. \quad (3)$$

⁷Of course, \overline{U}_S may incorporate that the supervisor expects to have some degree of altruism towards another agent he may work with when taking his outside option.

⁸In light of the dynamic nature of the model, this is a realistic assumption. Alternatively, one could imagine that the agent and supervisor have been co-workers at the firm. During that period they formed some social relationship. The formal model then starts at the point when the principal offers one of the co-workers a promotion to the position of supervisor.

⁹As will become clear in the analysis, unless the supervisor's contract screens for one type, the principal cannot learn the supervisor's type over time either.

The principal's reservation utility $\bar{\Pi}$ is normalized to zero. To make the problem interesting, it is assumed that:

$$\frac{1}{2\theta} - \bar{U}_A - \bar{U}_S > 0. \quad (4)$$

Assumption (4) ensures that, if the agent exerts the first-best effort level and the agent and the supervisor don't earn a rent, hiring an agent and a supervisor is attractive.

The order of the game is as follows.

1) In some period t , the principal is matched to one agent and one supervisor.

2) The supervisor's type γ is randomly drawn, and observed by the agent and the supervisor.

3) The principal offers a take-it-or-leave-it contract to the agent and the supervisor. I impose that the supervisor's contract is uniform, that is, the principal cannot offer a menu of contracts to the supervisor.¹⁰ All players observe the content of both contracts.

4) The agent and the supervisor decide whether or not to accept their contract. Without loss of generality, I assume that all players receive their reservation utility in period t unless both the agent and the supervisor accept their contract.

5) If both the agent and the supervisor accept their contract, the agent exerts effort, the supervisor gives a performance evaluation, and payoffs realize.

6) Period $t + 1$ starts. If both the agent and the supervisor accepted their contract, the game is played again from the third stage. Otherwise, the game is played again from the first stage. In this case, the principal is matched to a new agent and a new supervisor.

4 Analysis

In this section I solve for the subgame perfect equilibrium of the game by using backward induction. Subsection 4.1 introduces the central agency problem, namely that the supervisor may want to bias performance evaluations.

¹⁰A motivation for this assumption is the 'equal pay for equal work' principle. The assumption also rules out that, within a firm, some supervisors are imposed a penalty for giving high performance evaluations while others are not.

After that, I derive the agent's optimal effort choice and participation decision. In subsection 4.3 the optimal contracts for the supervisor and the agent are derived. At this stage, the principal's choice whether or not to screen for a supervisor type is still kept exogenous. The principal's screening decision is analyzed in subsection 4.4.

4.1 The supervisor's evaluation decision

After the agent has exerted effort and his performance V has realized, the supervisor decides to give a high or low performance rating to the agent. The evaluation of the agent's performance affects the supervisor's utility both in the current period t and in the future. The supervisor's marginal utility in period t from giving a high performance rating is $(\gamma - \beta)b$, where γ reflects the supervisor's altruistic utility from giving the agent a high performance evaluation and β reflects the supervisor's monetary costs from doing so. Depending on $\gamma \gtrless \beta$ and the realization of the agent's performance $V \in \{0, 1\}$, four cases can be distinguished. In two of these cases, the supervisor may have an incentive to bias performance evaluations. If $\gamma > \beta$ and $V = 0$, the supervisor may have an incentive to give an unwarranted high performance evaluation. If $\gamma < \beta$ and $V = 1$, the supervisor may have an incentive to give an unwarranted low performance evaluation. In both cases, the supervisor's marginal benefit in period t from biasing the performance evaluation can be written as $|\gamma - \beta|b \geq 0$. In the remaining two cases, the marginal benefit in period t from biasing the performance evaluation equals $-|\gamma - \beta|b \leq 0$.

In the repeated game, the supervisor will only report performance truthfully in period t if the following holds: reporting performance truthfully yields rents in the future, the discounted value of which exceeds the current marginal benefit from biasing the performance evaluation. Formally, this requires:

$$\frac{\delta}{1-\delta} \{U_S^{tr} - \max [U_S^b; \overline{U}_S]\} \geq |\gamma - \beta|b, \text{ and} \quad (5)$$

$$\frac{\delta}{1-\delta} \{U_S^{tr} - \max [U_S^b; \overline{U}_S]\} \geq -|\gamma - \beta|b. \quad (6)$$

where U_S^{tr} is the supervisor's per-period utility from the job, given that he reports performance truthfully in period t , and U_S^b is the supervisor's per-period utility from the job, given that he biases the performance evaluation in period t . Note that, because he can always leave the firm, the supervisor

earns at least his reservation utility after biasing the performance evaluation. Condition (5) therefore ensures that the supervisor enjoys a weakly positive rent from his job. In turn, this implies that the supervisor is always willing to accept the job. In addition to this, if his marginal benefit from biasing the performance evaluation equals $-|\gamma - \beta|b \leq 0$, the supervisor will always report performance truthfully. In other words, condition (6) is slack. In the remainder of this subsection, U_S^{tr} and U_S^b are derived explicitly.

Given that the supervisor reports performance truthfully, the agent will exert some optimal effort level e^* (to be derived in the next subsection). The probability that $V = 1$ equals e , and given that $E[\mu] = 1$, the expected probability that $z = 1$ also equals e . It follows that the supervisor's utility from the job when rating performance correctly is equal to:

$$U_S^{tr} = \alpha + e^*(\eta - \beta b) + \gamma U_A. \quad (7)$$

If the supervisor biased a performance evaluation, the agent will henceforth shirk. This implies that the signal z will always be low in the future. Even though the agent shirks, if $\gamma > \beta$, the supervisor finds it nevertheless attractive to give a high evaluation report. In this case, the agent always earns the bonus b , implying he will stay with the firm.¹¹ However, if $\gamma < \beta$, the supervisor always claims performance was low, implying the agent only earns the base salary a . Depending on whether $a < \overline{U}_A$ or $a \geq \overline{U}_A$, the agent will leave or stay with the firm. In case the agent leaves the firm, the supervisor also loses his job by assumption. Hence, the supervisor's per-period utility from the job after biasing the performance evaluation is equal to:

$$U_S^b = \begin{cases} \alpha + \gamma a + (\gamma - \beta)b & \text{if } \gamma \geq \beta \\ \alpha + \gamma a & \text{if } \gamma < \beta \text{ and } a \geq \overline{U}_A \\ \overline{U}_S & \text{if } \gamma < \beta \text{ and } a < \overline{U}_A \end{cases}. \quad (8)$$

¹¹To see this, suppose that the supervisor will report performance truthfully, implying the agent will optimally exert some effort level e^* (to be derived in the next section). To make the job attractive to the agent, the principal must then pay the agent a base salary at least equal to:

$$a^* \geq \overline{U}_A + \frac{1}{2}\theta(e^*)^2 - e^*b.$$

Next, suppose that, after both players have accepted the job, the supervisor chooses to bias the performance evaluation upward. The agent then earns the bonus irrespective of his performance, and so it becomes optimal for him to shirk. Given the minimum base salary derived above, the agent's utility from the job now becomes $a^* + b \geq \overline{U}_A + \frac{1}{2}\theta(e^*)^2 + (1 - e^*)b$. Since $e \leq 1$ by assumption, this is always greater than the agent's reservation utility \overline{U}_A .

Finally, the supervisor already enjoys either U_S^{tr} or U_S^b in the first period he is employed. The reason is that the agent knows whether or not the supervisor will evaluate performance truthfully. This follows from the assumptions that the agent observes the supervisor's type and contract.

4.2 The agent's effort and participation decision

Given that the supervisor reports performance truthfully, the agent's optimal effort choice is:

$$e = \frac{b}{\theta}. \quad (9)$$

Note that, since the first best effort-level equals $e^{FB} = \frac{1}{\theta}$, the first-best level of the bonus equals $b^{FB} = 1$. If the supervisor does not report performance truthfully, the agent will exert zero effort. The reason is that the supervisor will give either a high or low performance evaluation, independent of the agent's performance. In that case, exerting effort yields the agent no marginal benefit.

Any contract offered to the agent should satisfy the agent's participation constraint. Given the agent's effort choice, and assuming the supervisor reports performance truthfully, this constraint reads:

$$a + \frac{1}{2} \frac{b^2}{\theta} \geq \overline{U}_A. \quad (10)$$

4.3 Optimal contracts

In this subsection the optimal contracts for the agent and supervisor are derived. The principal's screening decision is still kept exogenous. It is analyzed in the following subsection. I start this subsection by establishing the following intermediary result. Suppose that the principal wants to attract some supervisor type and have him report performance truthfully. Then, depending on the supervisor's contract, it may be that $U_S^b > \overline{U}_S$ or $U_S^b \leq \overline{U}_S$ (see (8)). However, the principal is better off the lower U_S^b is. The intuition is that, the lower the supervisor's utility after biasing a performance evaluation, the less total compensation he requires to be willing to report performance truthfully. Since the supervisor can always take his outside option, a lower bound on U_S^b is given by \overline{U}_S . It follows that the principal strictly prefers $U_S^b \leq \overline{U}_S$ to $U_S^b > \overline{U}_S$.

The next thing to note is that the principal can set the supervisor's base salary α as low as necessary to ensure that $U_S^b \leq \overline{U}_S$ holds (see (8)). Moreover, doing so is costless. This holds because the base salary α and the bonus η are substitutes in the supervisor's compensation, given that he reports performance truthfully (see (7)). A low base salary can thus be compensated for by a higher value of η . Lemma 1 follows:

Lemma 1 *The principal sets the supervisor's base salary α sufficiently low, such that $U_S^b \leq \overline{U}_S$ holds for each supervisor type the principal attracts.*

Before we turn to the optimal contracts, note that Lemma 1 implies that the supervisor will report performance truthfully if:

$$\frac{\delta}{1-\delta} [\alpha + e^* (\eta - \beta b) + \gamma U_A - \overline{U}_S] \geq |\gamma - \beta| b. \quad (11)$$

4.3.1 Attracting one supervisor type (screening)

If screening for one supervisor type is optimal, the principal offers the following contracts to the agent and supervisor.

Proposition 1 *Suppose screening for one supervisor type is optimal. The supervisor's optimal contract only attracts the altruistic type. The supervisor's base salary α^* satisfies Lemma 1, and it holds that:*

$$\beta^* = \overline{\gamma} > 0, \text{ and } \eta^* \geq \frac{1}{2} \overline{\gamma} > 0.$$

The supervisor's total compensation equals $w_S^ = \overline{U}_S - \overline{\gamma} \overline{U}_A$.*

The agent's optimal contract is given by:

$$a^* = \overline{U}_A - \frac{1}{2\theta} < \overline{U}_A \text{ and } b^* = b^{FB} = 1.$$

The agent earns his reservation utility from the job, $U_A = \overline{U}_A$.

Proof. Appendix A. ■

The optimal contracts can be explained as follows. First, screening reveals the supervisor's type which is valuable information to the principal.

The reason is that it enables her to make the supervisor indifferent to his evaluation report. This is achieved by equating β to the supervisor's type γ . The benefit is that the supervisor will report performance truthfully without requiring costly rents to do so (one can easily see this in condition (5)). Moreover, since the supervisor is indifferent to his evaluation report for any value of b , it is optimal to pay the agent the first-best level of the bonus, $b^* = b^{FB} = 1$. Note that because the optimal screening contract attracts only the altruistic type, the contract contains a monetary cost for reporting high performance ($\beta^* = \bar{\gamma} > 0$).

Second, the reason that the supervisor's contract screens for the altruistic type is as follows. As explained above, given that $\beta = \gamma$, the supervisor requires the lowest possible compensation level to report performance truthfully, which equals $w_S = \bar{U}_S - \gamma U_A$. Importantly, since it holds that $\bar{\gamma} > \underline{\gamma} = 0$, the altruistic type requires a lower compensation level than the selfish type. The intuition is that in contrast to the selfish type, the altruistic type enjoys working with the agent. Next, since both types are equally likely to occur, the supervisor is equally likely to accept his contract independent of the type screened for. It follows that the principal's expected profits from screening are highest if she screens for the altruistic supervisor.¹²

Third, the supervisor's compensation depends positively on the signal z via the bonus $\eta^* > 0$. The intuition is that the supervisor is optimally paid a low base salary α^* , as implied by Lemma 1. However, it must still hold that reporting performance truthfully is sufficiently attractive for the altruistic type. This is ensured by the bonus η^* .

Fourth, to see why the supervisor's contract only attracts the altruistic type, note that the wage differential explained above implies that reporting performance truthfully yields the selfish supervisor less than his reservation utility. Hence, the selfish type may only be hired given that he biases the agent's performance evaluation. If the selfish supervisor biases the perfor-

¹²If, contrary to the assumption made, the selfish type is sufficiently more abundant than the altruistic type, the principal may prefer to screen for the selfish type. It is possible to show that such a screening contract exists if it holds that $\frac{\delta}{1-\delta}\bar{U}_A < 1$. If this contract is accepted by the supervisor, the principal's profits are the same as in Proposition 1, minus the compensating wage differential the supervisor's altruism gives rise to. The possibility that screening for the selfish type dominates screening for the altruistic type only affects the first result of Proposition 3: for low values of the discount factor, it may be optimal to screen for the selfish type rather than for the altruistic type. However, the intuition behind the optimality of screening per se remains the same, namely that screening enables the principal to incentivize the supervisor at lower costs.

mance evaluation, he will bias the evaluation downward. The reason is that giving a high performance evaluation imposes a monetary cost upon the supervisor ($\beta^* = \bar{\gamma} > 0$). However, expecting this, the agent will not accept his contract in the first place. This is true because the agent will only earn the base salary from the job, which is below his reservation utility. If the agent rejects his contract, I assume that the principal does not hire the supervisor either. It follows that the screening contract only attracts the altruistic supervisor type.

Last, the principal pays the agent such a low base salary that the agent earns only his reservation utility from the job. This is optimal even though raising the agent's utility decreases the altruistic supervisor's compensation. The reason is that the supervisor always cares more for his own utility than the agent's utility ($\bar{\gamma} < 1$). It is therefore not possible for the principal to gain from paying the agent a higher compensation and cut the supervisor's compensation more than proportionally.

By way of summary, if the principal screens for the altruistic supervisor, the supervisor nor the agent earn a rent, and the agent exerts the first-best effort level. Therefore, screening yields the principal first-best profits, conditional on the supervisor and the agent accepting their contract.

4.3.2 Attracting both supervisor types (pooling)

If the principal optimally attracts both supervisor types, she offers the following contracts to the agent and supervisor.

Proposition 2 *Suppose the principal optimally attracts both supervisor types. Denote by δ^L , δ^M , and δ^H threshold values of the discount factor, where it holds that $0 < \delta^L < \delta^M < \delta^H < 1$. There are three different cases:*

Case I applies for $\delta \in [\delta^H, 1)$, in which it holds that:

$$\begin{aligned}\beta_I^{**} &= 0, \\ w_{S,I}^{**} &= \overline{U_S}, \text{ and} \\ b_I^{**} &= 1.\end{aligned}$$

Case II applies for $\delta \in (0, \delta^L]$ and $\delta \in [\delta^M, \delta^H)$. The only difference with case I is that:

$$b_{II}^{**} = \frac{\delta}{1 - \delta} \overline{U_A} < 1.$$

Case III applies for $\delta \in (\delta^L, \delta^M)$, in which it holds that:

$$\begin{aligned}\beta_{III}^{**} &= \frac{\bar{\gamma}}{2} \left(1 - \frac{\delta}{1 - \delta} \frac{\bar{U}_A}{b_{III}^{**}} \right), \\ w_{S,III}^{**} &= \bar{U}_S - \frac{\bar{\gamma}}{2} \bar{U}_A + \frac{\bar{\gamma}}{2} \frac{1 - \delta}{\delta} b_{III}^{**}, \text{ and} \\ b_{III}^{**} &= 1 - \theta \frac{\bar{\gamma}}{2} \frac{1 - \delta}{\delta} < 1.\end{aligned}$$

It holds that $b_I^{**} > b_{II}^{**} > b_{III}^{**} > 0$, $w_{S,I}^{**} = w_{S,II}^{**} < w_{S,III}^{**}$, and $\beta_{III}^{**} > 0$. Furthermore, for all cases it holds that 1) the agent's base salary a^{**} is set such that, given his effort choice, the bonus b^{**} , and the supervisor's contract, the agent earns his reservation utility from the job ($U_A = \bar{U}_A$), 2) the agent's base salary is below the agent's reservation utility ($a^{**} < \bar{U}_A$), 3) the supervisor's base salary α^{**} satisfies Lemma 1, and 4) the supervisor receives a bonus conditional on the signal about the agent's performance being high ($\eta^{**} > 0$).

Proof. Appendix A. ■

All cases from Proposition 2 share the property that the bonus $\eta^{**} > 0$. The intuition is identical to the one given in the previous subsection. The supervisor is paid a low base salary to make biasing performance ratings unattractive (Lemma 1). The bonus $\eta^{**} > 0$ compensates the supervisor for the low base salary, given that he reports performance truthfully. Also, the agent always earns his reservation utility from the job. This is optimal, simply because it minimizes the agent's total compensation level. In the remainder of this subsection, the distinguishing properties of each case from Proposition 2 are discussed in detail.

Case I. The most remarkable property of case I is that it sets $\beta_I^{**} = 0$. The supervisor's compensation is thus independent of his evaluation of the agent's performance. By the intuition from Proposition 1, this is the most efficient way to induce truthful ratings from the selfish type. However, by the same logic, $\beta_I^{**} = 0$ is very inefficient when inducing truthful ratings from the altruistic type. To see why $\beta_I^{**} = 0$ is nevertheless optimal, recall first that the supervisor will report performance truthfully if:

$$\frac{\delta}{1 - \delta} (w_S + \gamma U_A - \bar{U}_S) \geq |\gamma - \beta| b. \quad (12)$$

Second, given $\beta = 0$, this condition is only satisfied for the selfish type ($\gamma = 0$) if it holds that $w_S \geq \overline{U}_S$. Note that the supervisor cannot be paid a lower compensation level, since otherwise the selfish type will not accept the job. Third, given $\beta = 0$ and $w_S = \overline{U}_S$, condition (12) for the altruistic type reduces to:

$$\frac{\delta}{1-\delta}U_A \geq b. \quad (13)$$

Hence, if the above condition holds, the altruistic type will report performance truthfully even if there is no monetary penalty attached to reporting high performance, and the supervisor is paid the lowest possible compensation level.

The intuition behind this result is that the altruistic supervisor enjoys working with the agent, as measured by γU_A . As long as he reports performance truthfully, this utility cannot be extracted from the altruistic type because doing so would violate the selfish type's participation constraint. However, after biasing a performance evaluation, the optimal contract ensures that the supervisor's utility equals at most his reservation utility (Lemma 1). The altruistic supervisor thus enjoys a rent as long as he rates performance correctly. The implication is that altruism not only provides an incentive to bias performance evaluations, it also provides a motivation to report performance truthfully. Moreover, if the supervisor is sufficiently patient, as measured by δ , the latter effect outweighs the former. If this is the case, altruism provides a net incentive to report performance truthfully, rather than to bias performance evaluations. This is the first key result of the paper:

Corollary 1 *If condition (13) holds, supervisor's altruism provides a net incentive to report performance truthfully, rather than to bias performance evaluations.*

Next, in case I, the supervisor is assumed to be so patient ($\delta \geq \delta^H$) that the altruistic type prefers to report performance truthfully, even if $b = 1$ and $U_A = \overline{U}_A$. It follows that both supervisor types report performance truthfully, the agent is provided first-best incentives, and the agent and the supervisor are paid the lowest possible compensation level. There is no set of contracts that could yield higher profits. Finally, note that for reasonable values of the discount factor, say $\delta \geq 0.8$, condition (13) is satisfied even if the agent's first-best bonus is a factor four higher than the agent's outside option utility. Therefore, case I is likely to apply in real workplace settings.

Case II. This case is almost identical to case I, except that it applies for lower values of the supervisor's patience ($\delta < \delta^H$). The implication is that, if $b = 1$ and $U_A = \overline{U}_A$, supervisor altruism no longer provides a net incentive to report performance truthfully, but rather to bias evaluations. In case II, the principal deals with this problem in a simple way: she reduces the agent's bonus till the point where the supervisor's altruism again provides a net incentive to report performance truthfully. Hence, $b_{II}^{**} = \frac{\delta}{1-\delta} \overline{U}_A < 1$.

One may note that an alternative way to solve this problem is to raise the agent's utility above \overline{U}_A . This is never optimal, however. The reason is that the supervisor always cares more for his own compensation than for the agent's utility ($\overline{\gamma} < 1$). It is therefore more efficient to increase the supervisor's rents directly by raising his own compensation, rather than indirectly by raising the agent's utility. This brings us to the final case.

Case III. As does case II, case III applies for relatively low values of the discount factor. However, in contrast to case II, the principal does not design the agent's and supervisor's contract such that the supervisor's altruism provides a net incentive to report performance truthfully. Instead, she simply raises the supervisor's compensation to ensure that reporting truthfully yields sufficient rents for the altruistic type. This response gives rise to two trade-offs.

First, the principal optimally lowers the agent's bonus below the first-best level. The reason is that the supervisor would enjoy a lower marginal benefit in period t from biasing the performance evaluation. The benefit to the principal is that the supervisor now requires lower rents from reporting truthfully, which in turn lowers his compensation.

Second, the principal imposes a penalty on the supervisor for giving a high performance evaluation, that is, $\beta_{III}^{**} > 0$. To see why this is optimal, note that raising the supervisor's compensation not only increases the altruistic type's rents from reporting truthfully, but also the selfish type's. The principal can therefore impose a penalty upon the supervisor for giving a high performance evaluation, without changing the selfish type's decision to report truthfully. The benefit of doing so is that the altruistic type enjoys a lower marginal benefit in period t from biasing performance evaluations. This, in turn, lowers the supervisor's compensation. Whether case III yields higher profits than case II depends on the value of the discount factor. In the proof of the Proposition, the threshold values of the discount factor are explicitly derived.

Summarizing this subsection, offering the pooling contract to the supervisor yields first-best profits if the discount factor is high (case I). For lower values of the discount factor the principal's profits are below first-best because the agent's incentives are reduced (case II), or because the agent's incentives are reduced and the supervisor earns a rent (case III).

4.4 Optimal screening

In this subsection, the principal's decision whether or not to offer the screening contract to the supervisor is analyzed. Denote by $E[\Pi]$ the principal's per-period profits given that the agent and the supervisor accepted their contract, which will be defined later on. Next, if the principal offers the pooling contract to the supervisor, there is a probability $\varphi = 0$ that the supervisor will reject his contract, whereas if the principal offers the screening contract there is a probability of $\varphi = \frac{1}{2}$ that the supervisor rejects his contract. If the supervisor rejects his contract, all players by assumption earn their outside option in period t . In that case the principal is matched to a new agent and a new supervisor in period $t + 1$. The principal's expected lifetime utility from being matched to one agent and one supervisor can now be written as:

$$E[\Pi^L] = \varphi \{0 + \delta E[\Pi^L]\} + (1 - \varphi) \frac{1}{1 - \delta} E[\Pi], \quad (14)$$

which is equivalent to:

$$E[\Pi^L] = \frac{1 - \varphi}{(1 - \delta)(1 - \delta\varphi)} E[\Pi]. \quad (15)$$

Let's now define the value of $E[\Pi]$. Using the results from Proposition 1, if the principal screens for the altruistic supervisor, $E[\Pi]$ equals:

$$E[\Pi^*] = \frac{1}{2\theta} - (1 - \bar{\gamma}) \bar{U}_A - \bar{U}_S > 0, \quad (16)$$

where the sign follows from assumption (4) and $\bar{\gamma} > 0$. If the principal attracts both supervisor types, the value of $E[\Pi]$ depends on which case from Proposition 2 applies. For each case, the principal's profits are given

by:

$$E[\Pi_I^{**}] = \frac{1}{2\theta} - \overline{U}_A - \overline{U}_S > 0, \quad (17)$$

$$E[\Pi_{II}^{**}] = \frac{b_{II}^{**}}{\theta} - \frac{1}{2} \frac{(b_{II}^{**})^2}{\theta} - \overline{U}_A - \overline{U}_S, \text{ and} \quad (18)$$

$$E[\Pi_{III}^{**}] = \frac{b_{III}^{**}}{\theta} - \frac{1}{2} \frac{(b_{III}^{**})^2}{\theta} - \frac{1-\delta}{\delta} \frac{\bar{\gamma}}{2} b_{III}^{**} - \left(1 - \frac{\bar{\gamma}}{2}\right) \overline{U}_A - \overline{U}_S, \quad (19)$$

respectively. It is now possible to derive when screening is optimal by comparing the expected lifetime utility from screening to the expected lifetime utility from pooling.

By Proposition 2, if the principal offers the pooling contract to the supervisor and $\delta \geq \delta^H$, the contracts from case I are optimal. It follows that, conditional on $\delta \geq \delta^H$, screening is optimal if:

$$\frac{1 - \frac{1}{2}}{(1-\delta)(1 - \frac{1}{2}\delta)} E[\Pi^*] \geq \frac{1}{1-\delta} E[\Pi_I^{**}]. \quad (20)$$

Rewriting this condition yields that screening is optimal if the discount factor is sufficiently high:

$$\delta \geq 2 - \frac{E[\Pi^*]}{E[\Pi_I^{**}]} \quad (21)$$

Note that the threshold value of the discount factor is smaller than 1 since $E[\Pi^*] > E[\Pi_I^{**}]$. The intuition behind this result is as follows. If the principal screens for the altruistic type, conditional on the supervisor accepting his contract, she enjoys the compensating wage differential that altruism gives rise to. I call this motive the ‘wage differential motive’ for screening. As revealed by condition (21), the wage differential motive only makes screening attractive if the principal is sufficiently patient. The reason is that screening implies that the supervisor may reject his job a number of times. In expectation, the wage differential is therefore only enjoyed somewhere in the future.

Next, by Proposition 2, if the principal offers the pooling contract to the supervisor and $\delta < \delta^H$, the contracts from case II or case III are optimal. Unfortunately, the equivalent of condition (20) for cases II and III is difficult to solve explicitly for the discount factor. However, by the following argument it can still be shown that screening is optimal for low values of

the discount factor. First, recall from Proposition 2 that, if the principal attracts both supervisor types and $\delta \in (0, \delta^L]$, case II applies. Hence, the principal's profits from attracting both supervisor types are given by $E[\Pi_{II}^{**}]$. Second, $E[\Pi_{II}^{**}]$ is lower for lower values of the discount factor.¹³ The reason is that the agent's bonus b_{II}^{**} is lower for lower values of the discount factor. As explained in the previous subsection, this is done to ensure that the altruistic supervisor prefers to report performance truthfully. Third, in the limit where the discount factor is zero, the agent's bonus is optimally equal to zero. It follows that $E[\Pi_{II}^{**}]$ is strictly negative. However, the principal's profits from screening are always strictly positive. Therefore, there must be a range of low discount factors for which screening is optimal. The intuition behind this result is that screening allows the principal to give the agent stronger incentives. I call this the 'incentive motive' for screening. The final Proposition follows:

Proposition 3 *There exist two ranges of discount factors, $\delta \in (0, \underline{\delta}]$ and $\delta \in [\bar{\delta}, 1)$ where $\underline{\delta} > 0$ and $\bar{\delta} < 1$, for which the principal optimally screens for altruistic supervisors.*

Finally, note that a similar limit argument as given above for case II cannot be established for case III. The simple reason is that case III only applies for intermediate values of the discount factor (see Proposition 2). However, as in case II, it holds that the principal's profits from attracting both supervisor types are lower, the lower the value of the discount factor.¹⁴ The reason is that, in order to reduce the supervisor's compensation, the agent is optimally paid a lower bonus if the discount factor is lower. Hence, also in case III the incentive motive may imply that screening is optimal for relatively low values of the discount factor. In addition to this, note that in

¹³Taking derivatives yields:

$$\frac{\partial E[\Pi_{II}^{**}]}{\partial \delta} = \frac{1 - b_{II}^{**}}{\theta(1 - \delta)^2} > 0,$$

where the sign follows from the fact that $b_{II}^{**} < 1$.

¹⁴Taking derivatives yields:

$$\frac{\partial E[\Pi_{III}^{**}]}{\partial \delta} = \frac{\bar{\gamma}}{4\delta^3} [2\delta - (1 - \delta)\bar{\gamma}\theta] > 0,$$

where the sign follows from the fact that $b_{III}^{**} > 0$.

both cases II and III, the wage differential motive may imply that screening is optimal for relatively high values of the discount factor. However, due to computational limitations, it cannot be established when these motives are sufficiently strong such that screening is optimal.

5 Concluding remarks

This paper studies how firms can optimally deal with the possibility of leniency bias through designing contracts for employees and potentially altruistic supervisors. Two key results emerge from this analysis. First, the optimal contracts sometimes ensure that supervisor's altruism provides a net incentive to report performance truthfully, rather than to bias performance evaluations. Second, if the discount factor is relatively low or high, the principal optimally screens for the altruistic supervisor.

The present analysis assumes that the principal can only offer a uniform contract to the supervisor, which seems to be a realistic assumption in many workplace settings. Nevertheless, if the principal has the possibility to offer a menu of contracts to the supervisor it could be optimal for her to exploit this possibility, and hence the results found here may be affected. I leave this potentially interesting case for further research.

Another avenue for further research is to allow for the possibility of collusion. A detailed analysis of collusion proof contracts in an infinitely repeated game is provided by Thiele (2013). Interestingly, the analysis by Thiele (2013) and the one presented here display close similarities. Specifically, an altruistic supervisor can be interpreted as a supervisor who is prone to colluding with the agent against the principal. Both Thiele (2013) and the present analysis reveal that, to eliminate bias in performance evaluation, the supervisor may require a rent, and the agent's incentives may be optimally reduced. It would be interesting to study whether and how social relationships and the possibility of collusion interact in the workplace.

Appendix A. Proofs

This Appendix contains the proofs of the results described in Propositions 1 and 2. To facilitate the reading of the proofs, I start by listing the constraints the principal faces when designing the agent's and supervisor's contracts.

First, given that the supervisor reports performance truthfully, the agent's incentive compatibility constraint is:

$$e = \frac{b}{\theta}, \quad (22)$$

and, second, the agent's participation constraint is:

$$U_A = a + \frac{1}{2} \frac{b^2}{\theta} \geq \overline{U}_A. \quad (23)$$

Last, the altruistic and selfish supervisor types' truthtelling constraints are given by:

$$\frac{\delta}{1-\delta} \left[\alpha + \frac{b}{\theta} (\eta - \beta b) + \bar{\gamma} U_A - \overline{U}_S \right] \geq |\bar{\gamma} - \beta| b, \text{ and} \quad (24)$$

$$\frac{\delta}{1-\delta} \left[\alpha + \frac{b}{\theta} (\eta - \beta b) - \overline{U}_S \right] \geq |0 - \beta| b, \quad (25)$$

respectively. Note that (24) and (25) ensure that reporting performance truthfully yields the supervisor at least his reservation utility. Finally, recall that α must satisfy Lemma 1. That is, α must be set sufficiently low such that $U_S^b \leq \overline{U}_S$ holds for each supervisor type the principal attracts, where U_S^b is given by (8).

Proof of Proposition 1

In this proof the optimal contracts are derived, given that the principal optimally screens for one supervisor type. The proof relies on the assumption that the principal optimally screens for the altruistic supervisor type. In the main text it is proved that, and explained why this assumption holds.

The principal's problem is to design the contracts that maximize her expected profits subject to the agent's incentive compatibility constraint (22), the agent's participation constraint (23), and the altruistic supervisor type's truthtelling constraint (24). Moreover, the supervisor's contract must not

attract the selfish supervisor type. I start by deriving the optimal contracts assuming the latter condition is satisfied. After that, it is checked whether this condition holds, given the optimal contract for the supervisor. The principal's problem can now be written as:

$$\max_{a,b,\beta,\eta} E[\Pi] = \frac{b}{\theta} - \left(a + \frac{b^2}{\theta} \right) - \left[\alpha + \frac{b}{\theta} (\eta - \beta b) \right] \text{ s.t. (23) and (24).}$$

The first-order conditions to the problem are:

$$\frac{\partial E[\Pi]}{\partial a} = -1 + \lambda_1 + \lambda_2 \frac{\delta}{1-\delta} \bar{\gamma} = 0, \quad (26)$$

$$\frac{\partial E[\Pi]}{\partial b} = \frac{1}{\theta} - 2\frac{b}{\theta} - \left(\frac{1}{\theta} \eta - 2\frac{b}{\theta} \beta \right) + \lambda_1 \frac{b}{\theta} + \quad (27)$$

$$\lambda_2 \left[\frac{\delta}{1-\delta} \left(\frac{1}{\theta} \eta - 2\frac{b}{\theta} \beta + \bar{\gamma} \frac{b}{\theta} \right) - |\bar{\gamma} - \beta| \right] = 0,$$

$$\frac{\partial E[\Pi]}{\partial \beta} = \frac{b^2}{\theta} - \lambda_2 \left[\frac{\delta}{1-\delta} \frac{b^2}{\theta} + \frac{\beta - \bar{\gamma}}{|\bar{\gamma} - \beta|} b \right] = 0, \text{ and} \quad (28)$$

$$\frac{\partial E[\Pi]}{\partial \eta} = -\frac{b}{\theta} + \lambda_2 \frac{\delta}{1-\delta} \frac{b}{\theta} = 0, \quad (29)$$

where λ_1 and λ_2 are the Lagrange multipliers on the agent's participation constraint (23) and the altruistic supervisor's truthtelling constraint (24), respectively. Solving the system of first-order conditions yields $\lambda_1 = 1 - \bar{\gamma} > 0$, $\lambda_2 = \frac{1-\delta}{\delta} > 0$, $\beta^* = \bar{\gamma}$, and $b^* = 1$. From $\lambda_1 > 0$ it follows that the agent's participation constraint (23) binds. This implies the agent's base salary equals $a^* = \bar{U}_A - \frac{1}{2\theta} < 0$, where the sign follows from assumption (4). Also, the agent's utility from the job equals his reservation utility, $U_A = \bar{U}_A$. From $\lambda_2 > 0$ it follows that the altruistic supervisor's truthtelling constraint (24) binds. Solving (24), given the solutions found so far, yields $\eta^* = \bar{\gamma} + \theta (\bar{U}_S - \alpha - \bar{\gamma} \bar{U}_A)$, where the supervisor's base salary α must satisfy Lemma 1. Next, α satisfies Lemma 1 if $U_S^b \leq \bar{U}_S$ holds for the altruistic supervisor. Taking into account that $\beta^* = \bar{\gamma}$, equation (8) yields this is the case if $\alpha \leq \bar{U}_S - \bar{\gamma} a^*$. It follows that $\eta^* \geq \frac{1}{2} \bar{\gamma} > 0$. Further, for any α that satisfies Lemma 1, the supervisor's total compensation level can now be computed to be equal to $w_S^* = \bar{U}_S - \bar{\gamma} \bar{U}_A$.

It must still be checked whether the supervisor's contract indeed only attracts the altruistic type. First note that, given the solutions found so far,

the selfish type's truthtelling constraint (25) reduces to:

$$-\frac{\delta}{1-\delta}\overline{U}_A \geq 1, \quad (30)$$

which can never hold. The selfish type will therefore give biased performance evaluations if he accepts the job. Next, since it holds that the supervisor's monetary cost for reporting high performance is positive, $\beta^* = \overline{\gamma} > 0$, the selfish supervisor will always give a low performance evaluation. Therefore, the agent only earns his base salary from his job, which is below his reservation utility, $a^* < \overline{U}_A$. It follows that the agent will not accept the job if the supervisor is selfish. Finally, by assumption, in this case the selfish supervisor is not hired either.

Proof of Proposition 2

In this proof the optimal contracts are derived, given that the principal optimally attracts both supervisor types. I start by establishing the following two intermediate results.

First, it can never be optimal to attract both supervisor types, but have only one of them report performance truthfully. The reason is that this strategy is always dominated by the strategy to screen for one supervisor type. To see this, note that if the principal screens, she can enjoy two possible profit levels both of which occur with a probability of $\frac{1}{2}$. If the supervisor accepts his job, profits equal the first-best level (Proposition 1). If the supervisor rejects his job, by assumption the principal enjoys her reservation utility $\overline{\Pi} = 0$. In case both supervisor types are attracted but only one of them reports performance truthfully, there are also two possible profit levels that each occur with a probability of $\frac{1}{2}$. If the supervisor reports performance truthfully, the principal enjoys at most the first-best profit level. However, if the supervisor biases performance evaluations, profits are strictly negative. The reason is that the agent will exert zero effort in this case, while hiring the agent and supervisor is costly. Hence, it can never be optimal to attract both supervisor types, but incentivize only one of the types to report performance truthfully. It follows that both supervisor types' truthtelling constraints (24 and 25) must be satisfied in equilibrium.

Second, it must always hold that $\beta \in [0, \overline{\gamma}]$. The reason is that if $\beta < 0$ or $\beta > \overline{\gamma}$, the principal could increase or decrease β and thereby lower both types' marginal benefit in period t from rating performance incorrectly. This

would be strictly profit-increasing, since then both types require lower rents to be willing to report performance truthfully. Because it turns out that only the lower constraint on β may bind in equilibrium, the following constraint is added to the principal's maximization problem:

$$\beta \geq 0. \quad (31)$$

The principal's problem can now be written as:

$$\max_{a,b,\beta,\eta} E[\Pi] = \frac{b}{\theta} - \left(a + \frac{b^2}{\theta} \right) - \left[\alpha + \frac{b}{\theta} (\eta - \beta b) \right] \text{ s.t. (23), (24), (25), and (31).}$$

The first-order conditions to the problem are:

$$\frac{\partial E[\Pi]}{\partial a} = -1 + \lambda_1 + \lambda_2 \frac{\delta}{1-\delta} \bar{\gamma} = 0, \quad (32)$$

$$\frac{\partial E[\Pi]}{\partial b} = \frac{1}{\theta} - 2\frac{b}{\theta} - \left(\frac{1}{\theta} \eta - 2\frac{b}{\theta} \beta \right) + \lambda_1 \frac{b}{\theta} + \quad (33)$$

$$\lambda_2 \left[\frac{\delta}{1-\delta} \left(\frac{1}{\theta} \eta - 2\frac{b}{\theta} \beta + \bar{\gamma} \frac{b}{\theta} \right) - |\bar{\gamma} - \beta| \right] +$$

$$\lambda_3 \left[\frac{\delta}{1-\delta} \left(\frac{1}{\theta} \eta - 2\frac{b}{\theta} \beta \right) - |0 - \beta| \right] = 0,$$

$$\frac{\partial E[\Pi]}{\partial \beta} = \frac{b^2}{\theta} - \lambda_2 \left[\frac{\delta}{1-\delta} \frac{b^2}{\theta} + \frac{\beta - \bar{\gamma}}{|\bar{\gamma} - \beta|} b \right] - \quad (34)$$

$$\lambda_3 \left[\frac{\delta}{1-\delta} \frac{b^2}{\theta} + \frac{\beta - 0}{|0 - \beta|} b \right] + \lambda_4 = 0, \text{ and}$$

$$\frac{\partial E[\Pi]}{\partial \eta} = -\frac{b}{\theta} + \lambda_2 \frac{\delta}{1-\delta} \frac{b}{\theta} + \lambda_3 \frac{\delta}{1-\delta} \frac{b}{\theta} = 0, \quad (35)$$

where λ_1 , λ_2 , λ_3 , and λ_4 are the Lagrange multipliers on (23), (24), (25), and (31), respectively. There are three qualitatively different sets of solutions to the system of first-order conditions. Which one applies will be shown to depend on the value of the discount factor δ . I start by solving the case that has been labelled case III in the main text.

Case III. Let's first assume that the constraint on β does not bind, that is, $\beta > 0$ and $\lambda_4 = 0$. Solving (34) and (35) then yields $\lambda_2 = \lambda_3 = \frac{1}{2} \frac{1-\delta}{\delta} > 0$, implying that both supervisor types' truthtelling constraints (24 and 25) bind

in equilibrium. Solving these constraints for β and η yields:

$$\beta_{III}^{**} = \frac{\bar{\gamma}}{2} \left(1 - \frac{\delta}{1 - \delta} \frac{U_A}{b_{III}} \right), \text{ and} \quad (36)$$

$$\eta_{III}^{**} = \frac{1 - \delta}{\delta} (\bar{\gamma} - \beta_{III}^{**}) \theta - \frac{\theta}{b_{III}} \left[\alpha_{III} - \frac{b_{III}^2}{\theta} \beta_{III}^{**} + \bar{\gamma} U_A - \bar{U}_S \right], \quad (37)$$

where the supervisor's base salary α_{III} must satisfy Lemma 1. Solving (32) yields $\lambda_1 = 1 - \frac{1}{2}\bar{\gamma} > 0$. From this it follows that the agent's participation constraint (23) binds, implying that $a_{III}^{**} = \bar{U}_A - \frac{1}{2} \frac{(b_{III}^{**})^2}{\theta}$ and $U_A = \bar{U}_A$. Given the solutions found so far, solving (33) yields:

$$b_{III}^{**} = 1 - \theta \frac{\bar{\gamma} 1 - \delta}{2 \delta} < 1. \quad (38)$$

One can now compute the supervisor's total compensation level, which is equal to:

$$w_{S,III}^{**} = \bar{U}_S - \frac{\bar{\gamma}}{2} \bar{U}_A + \frac{\bar{\gamma} 1 - \delta}{2 \delta} b_{III}^{**}. \quad (39)$$

Next, recall that Lemma 1 states that it must hold for all supervisor types attracted that the utility from biasing the performance evaluation is weakly below the supervisor's reservation utility, that is, $U_S^b \leq \bar{U}_S$. Taking into account that $\bar{\gamma} > \beta_{III}^{**} > 0$ and $a_{III}^{**} < \bar{U}_A$, equation (8) reveals this is the case for the altruistic type if it holds that the supervisor's base salary satisfies the following condition: $\alpha_{III} \leq \bar{U}_S - \bar{\gamma} a_{III}^{**} - (\bar{\gamma} - \beta_{III}^{**}) b_{III}^{**}$. In case the supervisor is selfish, it always holds that $U_S^b \leq \bar{U}_S$. The reason is that the agent will not accept the job given that the selfish type biases the performance evaluation, as the selfish type biases the performance evaluation downwards. By assumption, in this case the selfish supervisor is not hired either. Further, the condition imposed on the base salary α_{III} by Lemma 1 implies that $\eta_{III}^{**} > 0$. To see this, note that given the condition, η_{III}^{**} can be written as:

$$\eta_{III}^{**} \geq \frac{\theta}{\delta} (\bar{\gamma} - \beta_{III}^{**}) - \left(\frac{1}{2} \bar{\gamma} - \beta_{III}^{**} \right) b_{III}^{**} > 0, \quad (40)$$

where the sign follows from $\theta \geq 1$, $0 < \delta < 1$, and $0 < b_{III}^{**} < 1$.

Finally, recall that case III applies under the assumption that $\beta > 0$. Rewriting $\beta_{III}^{**} > 0$ yields that this assumption only holds if:

$$b_{III}^{**} = 1 - \theta \frac{\bar{\gamma} 1 - \delta}{2 \delta} > \frac{\delta}{1 - \delta} \bar{U}_A. \quad (41)$$

In turn, condition (41) is satisfied if and only if $\delta \in (\delta^L, \delta^M)$, where the threshold values of the discount factor are defined as:

$$\delta^L \equiv \frac{1 + \theta\bar{\gamma} - \sqrt{1 - 2\bar{U}_A\theta\bar{\gamma}}}{2 + \theta\bar{\gamma} + 2\bar{U}_A}, \text{ and} \quad (42)$$

$$\delta^M \equiv \frac{1 + \theta\bar{\gamma} + \sqrt{1 - 2\bar{U}_A\theta\bar{\gamma}}}{2 + \theta\bar{\gamma} + 2\bar{U}_A}. \quad (43)$$

Note that $\delta^L > 0$ and $\delta^L < \delta^M$ because $1 + \theta\bar{\gamma} > \sqrt{1 - 2\bar{U}_A\theta\bar{\gamma}} > 0$. The first inequality follows from $\bar{U}_A > 0$, $\theta > 1$, and $\bar{\gamma} > 0$. The second inequality follows from assumption (4) and $0 < \bar{\gamma} < 1$.

Cases I and II. Let's now assume that the condition on β binds, that is, $\beta = 0$ and $\lambda_4 > 0$. If $\beta = 0$, there are two subcases. The first subcase is where the altruistic supervisor's truthtelling constraint (24) binds, that is, $\lambda_2 > 0$. This leads to what has been labeled case II in the main text. In solving this subcase, it is assumed that also the selfish supervisor's truthtelling constraint (25) binds, that is, $\lambda_3 > 0$. Along the way, it is shown that this assumption indeed holds.

Solving the supervisor types' truthtelling constraints (24 and 25), given $\beta = 0$, yields:

$$b_{II}^{**} = \frac{\delta}{1 - \delta} U_A, \text{ and} \quad (44)$$

$$\eta_{II}^{**} = \frac{\theta(1 - 2\delta + \delta^2)(\bar{U}_S - \alpha_{II}) + \bar{\gamma}\delta^2 U_A^2}{\delta(1 - \delta)U_A}. \quad (45)$$

One can now compute the supervisor's compensation to be:

$$w_{S,II}^{**} = \bar{U}_S. \quad (46)$$

Given the solutions found so far, the first-order conditions (32), (33), and (35) can be solved for the Lagrange multipliers. This yields:

$$\lambda_1 = \frac{1}{\delta(1 - \delta)\theta} \left[\frac{\delta}{1 - \delta} U_A - 1 + \theta \frac{1 - \delta}{\delta} \right] > 0, \quad (47)$$

$$\lambda_2 = \frac{1}{(1 - \delta)\bar{\gamma}\theta} [1 - \delta(1 + U_A)], \text{ and} \quad (48)$$

$$\lambda_3 = \frac{1}{\bar{\gamma}\theta} \left[\frac{\delta}{1 - \delta} U_A - 1 + \theta \frac{1 - \delta}{\delta} \bar{\gamma} \right] > 0. \quad (49)$$

The signs of λ_1 and λ_3 follow from the assumptions that $\beta = 0$, which has been shown above to imply that $\frac{\delta}{1-\delta}\overline{U}_A - 1 + \theta\frac{\overline{\gamma}}{2}\frac{1-\delta}{\delta} > 0$ (see condition (41)), and $\overline{\gamma} < 1$. Moreover, since $\lambda_1 > 0$, the agent's participation constraint (23) binds, from which it follows that $a_{II}^{**} = \overline{U}_A - \frac{1}{2}\frac{(b_{II}^{**})^2}{\theta}$ and $U_A = \overline{U}_A$. The proof that $\eta_{II}^{**} > 0$ is analogous to the one given in case III. Proceeding in the same way as there, one can check that:

$$\eta_{II}^{**} \geq \frac{\overline{\gamma}}{2\delta(1-\delta)} [2\theta(1-\delta) + \delta^2\overline{U}_A] > 0. \quad (50)$$

To conclude case II, recall that it has been assumed that $\lambda_2 > 0$ and $\lambda_3 > 0$. The latter assumption indeed holds, as shown in (49). However, inspecting (48) reveals that the former assumption only holds if $\delta < \delta^H$, where the threshold value of the discount factor is defined as:

$$\delta^H \equiv \frac{1}{1 + \overline{U}_A}. \quad (51)$$

One can check that $\delta^M < \delta^H$. This inequality follows from $\theta > 1$, $\overline{U}_A > 0$, and $\overline{\gamma} > 0$. Therefore, case II applies for values of the discount factor such that $\delta \in (0, \delta^L]$ and $\delta \in [\delta^M, \delta^H)$.

The second subcase is where the altruistic supervisor's truthtelling constraint (24) is slack, that is, $\lambda_2 = 0$. As shown above, this holds for values of the discount factor $\delta > \delta^H$. The set of contracts that results has been labeled case I in the main text. Solving the first-order conditions (32), (33), and (35) yields $\lambda_1 = 1 > 0$, $\lambda_3 = \frac{1-\delta}{\delta} > 0$, and $b_I^{**} = 1$. From $\lambda_1 > 0$ it follows that the agent's participation constraint (23) binds, implying that $a_I^{**} = \overline{U}_A - \frac{1}{2\theta}$ and $U_A = \overline{U}_A$. From $\lambda_3 > 0$ it follows that the selfish supervisor type's truthtelling constraint (25) binds. Rewriting (25), given that $\beta_I^{**} = 0$, yields:

$$\eta_I^{**} = \theta(\overline{U}_S - \alpha_I). \quad (52)$$

One can now compute the supervisors's compensation to be:

$$w_{S,I}^{**} = \overline{U}_S. \quad (53)$$

The proof that $\eta_I^{**} > 0$ is analogous to the one given in case III. Proceeding in the same way as there, one can check that:

$$\eta_I^{**} \geq \frac{1}{2}\overline{\gamma} + \theta\overline{\gamma}(1 + \overline{U}_A) > 0. \quad (54)$$

By conclusion, case III applies for $\delta \in (\delta^L, \delta^M)$, case II applies for $\delta \in (0, \delta^L]$ and $\delta \in [\delta^M, \delta^H)$, and case I applies for the remaining values of the discount factor, $\delta \in [\delta^H, 1)$. Also, it has been shown to hold that $0 < \delta^L < \delta^M < \delta^H < 1$.

Appendix B. Allowing for Objective Incentives for the Agent

In this Appendix I show that, as long as σ_μ^2 is sufficiently large, allowing for objective incentives for the agent does not change the key results derived in this paper. The way I model the verifiable information about the agent's performance is identical to Baker et al. (1994).

Suppose that the agent's contract also contains a bonus v to be paid if the verifiable signal is high ($z = 1$). Given that the supervisor reports performance truthfully, and some realization of μ which the agent has privately observed, the agent's optimal effort choice is:

$$e = \frac{b + \mu v}{\theta}, \quad (55)$$

which varies with μ . The agent's participation constraint now reads:

$$E \left[a + \frac{b + \mu v}{\theta} (b + \mu v) - \frac{1}{2} \theta \left(\frac{b + \mu v}{\theta} \right)^2 \right] \geq \overline{U}_A, \quad (56)$$

where the expectation runs over μ . Using $E[\mu] = 1$ and $E[\mu^2] = 1 + \sigma_\mu^2$, the participation constraint can be written as:

$$a + \frac{1}{2} \frac{b^2 + 2bv + (1 + \sigma_\mu^2) v^2}{\theta} \geq \overline{U}_A. \quad (57)$$

Rewriting (57) in another way yields that the agent must be paid at least a total expected compensation level of:

$$w_A \geq \overline{U}_A + \frac{1}{2} \frac{b^2 + 2bv + (1 + \sigma_\mu^2) v^2}{\theta}. \quad (58)$$

Equation (58) reveals that, if $\sigma_\mu^2 > 0$, the agent's compensation rises faster in v compared to b . The reason is as follows. If $v > 0$, the agent will sometimes

exert a high effort level and sometimes a low effort level depending on the realization of μ . Because effort costs are convex, average effort costs increase in the variation in e , and therefore in σ_μ^2 . For these costs the agent must be compensated. Using objective incentives for the agent thus generates its own kind of agency costs. I assume that using only objective incentives for the agent does not generate sufficient value to compensate the agent for his reservation utility. One can show that this assumption holds if:

$$\sigma_\mu^2 > \frac{1 - 2\theta\overline{U}_A}{2\theta\overline{U}_A} > 0, \quad (59)$$

where the sign follows from assumption (4). In the following I derive the optimal contracts in case the principal can also use objective incentives for the agent.

First, note that the sets of contracts derived in case the principal screens (Proposition 1) and in case the principal attracts both supervisor types and $\delta \geq \delta^H \equiv \frac{1}{1+\overline{U}_A}$ (Proposition 2, case I), are still optimal. The reason is that given these sets of contracts, the supervisor nor the agent earns a rent, and the agent is induced to exert the first-best effort level. It follows that there is no way to improve upon these sets of contracts by using objective incentives for the agent. This observation already enables us to confirm some key results. First, given the contracts from Proposition 2, case I, the supervisor's altruism provides a net incentive to report performance truthfully, rather than to bias performance evaluations (Corollary 1). Second, the proof that screening is optimal for high values of the discount factor, presented in subsection 4.4, depends only on the aforementioned sets of contracts. Since these sets of contracts do not change, it follows that this result is insensitive to allowing for objective incentives for the agent. In the remainder of this Appendix, I derive the optimal contracts in case the principal attracts both supervisor types and it holds that $\delta < \delta^H$. The derivations are largely similar to the ones in the proof of Proposition 2.

Given the agent's effort e (55), $E[\mu] = 1$, and $E[\mu^2] = 1 + \sigma_\mu^2$, the altruistic and selfish supervisor will report performance truthfully if:

$$\frac{\delta}{1-\delta} \left[\alpha + \frac{(b+v)(\eta - b\beta) + v\sigma_\mu^2\eta}{\theta} + \overline{\gamma}U_A - \overline{U}_S \right] \geq |\overline{\gamma} - \beta|b, \text{ and} \quad (60)$$

$$\frac{\delta}{1-\delta} \left[\alpha + \frac{(b+v)(\eta - b\beta) + v\sigma_\mu^2\eta}{\theta} - \overline{U}_S \right] \geq |0 - \beta|b, \quad (61)$$

respectively. Further, recall from the proof of Proposition 2 that it must hold that:

$$\beta \geq 0. \quad (62)$$

The principal's problem can now be written as:

$$\begin{aligned} \max_{a,b,v,\beta,\eta} \Gamma &= \frac{b+v}{\theta} - \left[a + \frac{b^2 + 2bv + (1 + \sigma_\mu^2) v^2}{\theta} \right] - \\ &\quad \left[\alpha + \frac{(b+v)(\eta - b\beta) + v\sigma_\mu^2 \eta}{\theta} \right], \\ \text{s.t. } &(57), (60), (61), \text{ and } (62). \end{aligned}$$

The first-order conditions to the problem are:

$$\frac{\partial \Gamma}{\partial a} = -1 + \lambda_1 + \lambda_2 \frac{\delta}{1-\delta} \bar{\gamma} = 0, \quad (63)$$

$$\frac{\partial \Gamma}{\partial b} = \frac{1}{\theta} - \frac{2b+2v}{\theta} - \left[\frac{1}{\theta} (\eta - \beta b) - \beta \frac{b+v}{\theta} \right] + \lambda_1 \frac{b+v}{\theta} + \quad (64)$$

$$\lambda_2 \left\{ \frac{\delta}{1-\delta} \left[\frac{\eta - \beta b}{\theta} - \beta \frac{b+v}{\theta} + \bar{\gamma} \frac{b+v}{\theta} \right] - |\bar{\gamma} - \beta| \right\} +$$

$$\lambda_3 \left\{ \frac{\delta}{1-\delta} \left[\frac{\eta - \beta b}{\theta} - \beta \frac{b+v}{\theta} \right] - |0 - \beta| \right\} = 0,$$

$$\frac{\partial \Gamma}{\partial v} = \frac{1}{\theta} - \frac{2b+2(1+\sigma_\mu^2)v}{\theta} - \frac{\eta(1+\sigma_\mu^2) - b\beta}{\theta} + \lambda_1 \frac{b+v(1+\sigma_\mu^2)}{\theta} + \quad (65)$$

$$\lambda_2 \frac{\delta}{1-\delta} \left[\frac{\eta(1+\sigma_\mu^2) - b\beta}{\theta} + \bar{\gamma} \frac{b+v(1+\sigma_\mu^2)}{\theta} \right] + \lambda_3 \frac{\delta}{1-\delta} \frac{\eta(1+\sigma_\mu^2) - b\beta}{\theta} = 0,$$

$$\frac{\partial \Gamma}{\partial \beta} = \frac{b+v}{\theta} b - \lambda_2 \left(\frac{\delta}{1-\delta} \frac{b+v}{\theta} b + \frac{\beta - \bar{\gamma}}{|\bar{\gamma} - \beta|} b \right) - \quad (66)$$

$$\lambda_3 \left(\frac{\delta}{1-\delta} \frac{b+v}{\theta} b + \frac{\beta - 0}{|0 - \beta|} b \right) + \lambda_4 = 0, \text{ and}$$

$$\frac{\partial \Gamma}{\partial \eta} = -\frac{b+(1+\sigma_\mu^2)v}{\theta} + \lambda_2 \frac{\delta}{1-\delta} \frac{b+(1+\sigma_\mu^2)v}{\theta} + \lambda_3 \frac{\delta}{1-\delta} \frac{b+(1+\sigma_\mu^2)v}{\theta} = 0, \quad (67)$$

where λ_1 , λ_2 , λ_3 , and λ_4 are the Lagrange multipliers on (57), (60), (61), and (62), respectively. As in Proposition 2, if $\delta < \delta^H$, two cases can be

distinguished. Since the cases are closely comparable to cases II and III from Proposition 2, they are again denoted by subscripts *II* and *III*. I start by solving case III.

Case III. Let's first assume that condition (62) does not bind, that is, $\beta > 0$ and $\lambda_4 = 0$. Solving (63), (66), and (67) then yields $\lambda_1 = 1 - \frac{1}{2}\bar{\gamma} > 0$ and $\lambda_2 = \lambda_3 = \frac{1-\delta}{2\delta} > 0$. Hence, the agent's participation constraint (57) binds. This implies that $U_A = \bar{U}_A$ and $a_{III}^{**} = \bar{U}_A - \frac{1}{2} \frac{b_{III}^2 + 2b_{III}v_{III} + (1+\sigma_\mu^2)v_{III}^2}{\theta}$. Further, both supervisor types' truthtelling constraints (60 and 61) also bind. Solving these constraints yields:

$$\beta_{III}^{**} = \frac{\bar{\gamma}}{2} \left(1 - \frac{\delta}{1-\delta} \frac{\bar{U}_A}{b_{III}} \right), \text{ and} \quad (68)$$

$$\eta_{III}^{**} = \frac{\theta \left(\frac{1-\delta}{\delta} \beta_{III}^{**} b_{III} + \bar{U}_S - \alpha_{III}^{**} \right) + (b_{III} + v_{III}) b_{III} \beta_{III}^{**}}{b_{III} + (1 + \sigma_\mu^2) v_{III}}, \quad (69)$$

where α_{III}^{**} must satisfy Lemma 1. Next, solving (64) and (65) for b_{III} and v_{III} yields:

$$b_{III}^{**} = 1 - \theta \frac{\bar{\gamma}}{2} \frac{1 - \delta}{\delta} \frac{1 + \sigma_\mu^2}{\sigma_\mu^2}, \text{ and} \quad (70)$$

$$v_{III}^{**} = \frac{\bar{\gamma}}{2} \frac{1 - \delta}{\delta} \frac{\theta}{\sigma_\mu^2} > 0. \quad (71)$$

Thus, in case III the principal optimally uses objective incentives for the agent in order to reduce the rents the supervisor requires to report performance truthfully.

To conclude case III, recall that it applies as long as it holds that $\beta_{III}^{**} > 0$. This inequality can be written as:

$$b_{III}^{**} = 1 - \theta \frac{\bar{\gamma}}{2} \frac{1 - \delta}{\delta} \frac{1 + \sigma_\mu^2}{\sigma_\mu^2} > \frac{\delta}{1 - \delta} \bar{U}_A. \quad (72)$$

Solving the above condition yields that it must hold that $\delta \in (\delta^L, \delta^M)$, where the threshold values of the discount factor are given by:

$$\delta^L \equiv \frac{\sigma_\mu^2 + \bar{\gamma}\theta(1 + \sigma_\mu^2) - \sqrt{\sigma_\mu^2 [\sigma_\mu^2 - 2\bar{U}_A\bar{\gamma}\theta(1 + \sigma_\mu^2)]}}{2\sigma_\mu^2 + \bar{\gamma}\theta(1 + \sigma_\mu^2) + 2\bar{U}_A\sigma_\mu^2}, \text{ and} \quad (73)$$

$$\delta^M \equiv \frac{\sigma_\mu^2 + \bar{\gamma}\theta(1 + \sigma_\mu^2) + \sqrt{\sigma_\mu^2 [\sigma_\mu^2 - 2\bar{U}_A\bar{\gamma}\theta(1 + \sigma_\mu^2)]}}{2\sigma_\mu^2 + \bar{\gamma}\theta(1 + \sigma_\mu^2) + 2\bar{U}_A\sigma_\mu^2}. \quad (74)$$

Note that δ^L and δ^M only exist if the root terms in the expressions above are well-defined. This is only the case if $\sigma_\mu^2 - 2\overline{U}_A\overline{\gamma}\theta(1 + \sigma_\mu^2) > 0$, which can be written as:

$$\sigma_\mu^2 > \frac{2\overline{U}_A\overline{\gamma}\theta}{1 - 2\overline{U}_A\overline{\gamma}\theta}. \quad (75)$$

The above inequality is not necessarily implied by the assumption on the minimum level of σ_μ^2 (assumption (59)). It follows that case III may exist or not exist, depending on the level of σ_μ^2 .

Case II. The next step is to derive the optimal contracts for $\delta \notin (\delta^L, \delta^M)$, given that $\delta < \delta^H$ (case II). This case always exists. To see this, first note that if δ^L and δ^M are not well-defined, case II applies for all $\delta < \delta^H$. Second, given that it is well-defined, it can be easily checked that $\delta^L > 0$. For the remainder of the proof, it does not matter whether or not $\delta^M < \delta^H$.

First note that a property of case II is that $\beta_{II}^{**} = 0$, since for any $\beta > 0$ case III is optimal. Next, case II can be easily solved by the following argument. Let's suppose that the agent's participation constraint (57) and both supervisor types' truthtelling constraints (60 and 61) bind. If the agent's participation constraint (57) binds, it holds that $U_A = \overline{U}_A$ and $a_{II}^{**} = \overline{U}_A - \frac{1}{2} \frac{b_{II}^2 + 2b_{II}v_{II} + (1 + \sigma_\mu^2)v_{II}^2}{\theta}$. Moreover, given $\beta_{II}^{**} = 0$, solving both supervisor types' truthtelling constraints (60 and 61) yields:

$$b_{II}^{**} = \frac{\delta}{1 - \delta} \overline{U}_A, \text{ and} \quad (76)$$

$$\eta_{II}^{**} = \frac{\theta(\overline{U}_S - \alpha_{II}^{**})}{b_{II}^{**} + (1 + \sigma_\mu^2)v_{II}}, \quad (77)$$

where α_{II}^{**} must satisfy Lemma 1. Note that $b_{II}^{**} < 1$ by the assumption that $\delta < \delta^H$. Next, for any base salary α_{II}^{**} that satisfies Lemma 1 and the solutions found so far, the supervisor's compensation can be computed to be $w_{S,II}^{**} = \overline{U}_S$. The supervisor cannot be paid a lower compensation level, since otherwise the selfish type will reject the job. It follows directly that it is indeed optimal to have the agent's participation constraint (57) and both supervisor types' truthtelling constraints (60 and 61) bind. The reason is that the agent earns the lowest possible compensation level (since (57) binds), the supervisor earns the lowest possible compensation level, and the agent receives the highest possible subjective bonus (note that b_{II}^{**} cannot be greater than $\frac{\delta}{1 - \delta} \overline{U}_A$, as this would imply that case III is optimal, see condition (72)).

Importantly, in case II the optimal contracts imply that supervisor's altruism provides a net incentive to report performance truthfully (Corollary 1). To see this, note that given $\beta_{II}^{**} = 0$ and $w_{S,II}^{**} = \overline{U}_S$, the altruistic type's truthtelling constraint (60) reduces to $\frac{\delta}{1-\delta}\overline{U}_A \geq b_{II}^{**}$, which always holds.

Finally, one can now derive the optimal value of the objective bonus v_{II} . Using $E[\mu] = 1$, $E[\mu^2] = 1 + \sigma_\mu^2$, and the solutions found so far, the optimal value of v is given by:

$$\max_{v_{II}} E[\Pi_{II}^{**}] = \frac{\frac{\delta}{1-\delta}\overline{U}_A + v_{II}}{\theta} - \overline{U}_S - \left[\overline{U}_A + \frac{1}{2} \frac{\left(\frac{\delta}{1-\delta}\overline{U}_A\right)^2 + 2\frac{\delta}{1-\delta}\overline{U}_A v_{II} + (1 + \sigma_\mu^2) v_{II}^2}{\theta} \right], \quad (78)$$

which yields:

$$v_{II}^{**} = \frac{1 - \frac{\delta}{1-\delta}\overline{U}_A}{1 + \sigma_\mu^2} > 0, \quad (79)$$

where the sign follows from $b_{II}^{**} = \frac{\delta}{1-\delta}\overline{U}_A < 1$. Hence, in case II, the principal avoids paying the supervisor rents, but at the costs of setting the subjective bonus below the first-best level. To get the agent's effort level closer to the first-best, the principal adds objective incentives to the agent's contract.

It is now possible to confirm the final key result, namely that screening is optimal for relatively low values of the discount factor. The proof is analogous to the one given in subsection 4.4. Recall that the principal's per-period utility from screening, conditional on the supervisor and agent accepting their contract, equals $\frac{1}{2\theta} - (1 - \overline{\gamma})\overline{U}_A - \overline{U}_S$. This profit is strictly positive by assumption (4) and $\overline{\gamma} > 0$. Next, suppose that $\delta \rightarrow 0$. For $\delta = 0$ it holds that the subjective bonus equals $b_{II}^{**} = 0$. By assumption (59), the principal's utility is strictly negative if she only uses an objective incentive scheme. It follows that there must be a range of discount factors $\delta \in (0, \underline{\delta})$ for which screening is optimal, where $\underline{\delta} > 0$.

References

- [1] Baker, G., R. Gibbons, and K.J. Murphy (1994), "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics*, 109(4): 1125-1156.

- [2] Bandiera, O., I. Barankay, and I. Rasul (2009), "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, 77(4): 1047-1094.
- [3] Berger, J., C. Herbertz, and D. Sliwka (2011), "Managerial Incentives and Favoritism in Promotion Decisions: Theory and Field Evidence." IZA Discussion Paper, No. 5543.
- [4] Bol, J.C. (2011), "The Determinants and Performance Effects of Managers' Performance Evaluation Biases." *Accounting Review*, 85(5): 1549-1575.
- [5] Bull, C. (1987), "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics*, 102(1): 147-159.
- [6] Cordero Salas, P. and B. Roe (2012), "The Role of Cooperation and Reciprocity in Structuring Carbon Sequestration Contracts in Developing Countries." *American Journal of Agricultural Economics*, 94(2): 411-418.
- [7] Delfgaauw, J. and R. Dur (2007), "Signaling and Screening of Worker's Motivation." *Journal of Economic Behavior and Organization*, 62(4): 605-624.
- [8] Delfgaauw, J. and R. Dur (2008), "Incentives and Workers' Motivation in the Public Sector." *Economic Journal*, 118(525): 171-191.
- [9] Dur, R. and J. Tichem, "Altruism and Relational Incentives in the Workplace." *Journal of Economics and Management Strategy*, forthcoming.
- [10] Gibbs, M. (2012), "Design and Implementation of Pay for Performance." IZA Discussion Paper No. 6322.
- [11] Gibbs, M., K.A. Merchant, W.A. Van der Stede, and M.E. Vargus (2004), "Determinants and Effects of Subjectivity in Incentives." *Accounting Review*, 79(2): 409-436.
- [12] Giebe, T. and O. Gürtler (2012), "Optimal Contracts for Lenient Supervisors." *Journal of Economic Behavior and Organization*, 81(2): 403-420.

- [13] Golman, R. and S. Bhatia (2012), "Performance Evaluation Inflation and Compression." *Accounting, Organizations and Society*, 37(8): 534-543.
- [14] Grund, C. and J. Przemeczek (2012), "Subjective Performance Evaluation and Inequality Aversion." *Applied Economics*, 44(17): 2149-2155.
- [15] Jawahar, I.M. and C.R. Williams (1997), "Where All The Children Are Above Average: The Performance Appraisal Purpose Effect." *Personnel Psychology*, 50(4): 905-926.
- [16] Kamphorst, J.J.A. and O.H. Swank (2012), "The Role of Performance Appraisals in Motivating Employees." Tinbergen Institute Discussion Paper, TI2012-034/1.
- [17] Kosfeld, M. and F.A. Von Siemens (2009), "Worker Self-selection and the Profits from Cooperation." *Journal of the European Economic Association*, 7(2-3): 573-582.
- [18] Kosfeld, M. and F.A. Von Siemens (2011), "Competition, Cooperation, and Corporate Culture." *RAND Journal of Economics*, 42(1): 23-43.
- [19] Lee, S. and P. Persson (2011), "Authority versus Loyalty: Social Incentives and Governance." NYU Working Paper, No. FIN-10-001.
- [20] Levin, J. (2003), "Relational Incentive Contracts." *American Economic Review*, 97(3): 835-857.
- [21] MacLeod, W.B. and J.M. Malcomson (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, 57(2): 447-480.
- [22] MacLeod, W.B. and J.M. Malcomson (1998), "Motivation and Markets." *American Economic Review*, 88(3): 388-411.
- [23] MacLeod, W.B. and D. Parent, (2000) "Job Characteristics and the Form of Compensation." In: Polachek, S.W. and J. Robst (Eds.), *Research in Labor Economics*, Vol. 18, London: JAI Press.
- [24] Medoff, J.L. and K. G. Abraham (1980), "Experience, Performance, and Earnings." *Quarterly Journal of Economics*, 95(4): 703-736.

- [25] Moers, F. (2005), "Discretion and Bias in Performance Evaluation: The Impact of Diversity and Subjectivity." *Accounting, Organizations and Society*, 30(1): 67-80.
- [26] Müller, D. and P. Weinschenk (2011), "Rater Bias in Performance Appraisal: on Horns, Halos, and Incentive Provision." Mimeo, University of Bonn.
- [27] Non, A. (2012), "Gift-Exchange, Incentives, and Heterogeneous Workers." *Games and Economic Behavior*, 75(1): 319-336.
- [28] Prendergast, C. (1999), "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7-63.
- [29] Prendergast, C. and R.H. Topel (1996), "Favoritism in Organizations." *Journal of Political Economy*, 104(5): 958-978.
- [30] Sliwka, D. (2007), "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *American Economic Review*, 97(3): 999-1012.
- [31] Sol, J. (2010), "Peer Evaluation: Incentives and Co-Worker Relations." Tinbergen Institute Discussion Paper, TI2010-055/1.
- [32] Spence, M. (1973), "Job Market Signalling." *Quarterly Journal of Economics*, 87(3): 355-374.
- [33] Thiele, V. (2013), "Subjective Performance Evaluation, Collusion, and Organizational Design." *Journal of Law, Economics, and Organization*, 29(1): 35-59.
- [34] Von Siemens, F.A. (2011), "Heterogeneous Social Preferences, Screening and Employment Contracts." *Oxford Economic Papers*, 63(3): 499-522.
- [35] Von Siemens, F.A. (2012), "Social Preferences, Sorting and Competition." *Scandinavian Journal of Economics*, 114(3): 780-807.