Optimal Contracts for Lenient Supervisors*

Thomas Giebe, *Humboldt University Berlin***,

Oliver Gürtler, *University of Bonn****

Abstract
We consider a situation where an agent's effort is monitored by a supervisor who cares for the agent's well–being. This is modeled by incorporating the agent's utility into the utility function of the supervisor. The first–best solution can be implemented even if the supervisor's preferences are unknown. The corresponding optimal contract is similar to what we observe in practice: The supervisor's wage is constant and independent of his report. It induces one type of supervisor to report the agent's performance truthfully, while all others report favorably independent of performance. This implies that overstated performance (leniency bias) may be the outcome of optimal contracts under informational asymmetries.

Keywords: Subjective performance evaluation, leniency, supervisor, private information
JEL classification: D82, D86, J33, M52

* We would like to thank Patrick Bolton, Florian Englmaier, Veronika Grimm, Nicolas Klein, Johannes Koenen, Matthias Kräkel, Sebastian Kranz, Tymofiy Mylovanov, Judith Przemeck, Ferdinand von Siemens, Dirk Sliwka, Roland Strausz, Achim Wambach and Elmar Wolfstetter for helpful comments. Financial support by the Deutsche Forschungsgemeinschaft (DFG), grants SFB/TR 15 ("Governance and the Efficiency of Economic Systems"), is gratefully acknowledged.
** Thomas Giebe, Department of Economics, Economic Theory I, Humboldt University Berlin, Spandauer Straße 1, D-10178 Berlin, Germany, e-mail: thomas.giebe@wiwi.hu-berlin.de, phone: +49-30-20935773, fax: +49-30-20935619.
*** Oliver Gürtler, Department of Economics, BWL II, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, e-mail: oliver.guertler@uni-bonn.de, phone: +49-228-739214, fax: +49-228-739210.

1 Introduction

In many situations, employers are reluctant to tie an employee's compensation to objective (and verifiable) measures of performance because these measures rarely reflect the employee's true performance. Instead they may be affected by random factors, efforts of other employees, or measure only some of the demanded activities. An incentive contract based on such a measure is problematic, because the employee would demand a high risk-premium, engage in free-riding, or focus on the activities that are captured by the performance measure.[1]

Subjective performance measures are often argued to be more precise and to reflect an employee's performance better than objective measures. A subjective performance evaluation may therefore constitute a solution to the aforementioned problems.[2] However, these measures are by definition subjective and, hence, non-verifiable by third parties. As a consequence, it is often observed that supervisors distort performance ratings by not sufficiently differentiating good from bad performance.[3] Of particular relevance in this respect is "leniency bias" which refers to a practice of systematically

[1] These problems have been analyzed in a variety of papers. See e.g. Gibbons (1998, 2005) or Prendergast (2002a) for the risk–incentives trade-off, Holmström (1982) for the free-rider problem under limited liability and Kerr (1975), Holmström & Milgrom (1991), Baker (1992) or Feltham & Xie (1994) for the multi-tasking problem.

[2] See, for instance, Baker et al. (1994).

[3] See e.g. the excellent survey by Prendergast (1999).

overstating the employees' performance.[4]

This paper analyzes a principal – supervisor – agent relationship. In order to incorporate leniency into the model we assume the agent's utility to enter the supervisor's utility function.[5] This approach captures a setting where the supervisor is lenient because of a personal relationship with the agent. It also represents a scenario where the propensity to be lenient is a fixed characteristic of the supervisor. The relevance of the latter has been indicated by several field studies and psychological research.[6] In the first scenario the supervisor does not learn his preferences until after he has worked with the agent for some time. In the second scenario he knows his preferences ex ante. In principle, a contract can eliminate the leniency bias by punishing the supervisor for good effort ratings. This, however, requires knowledge of the supervisor's exact preference (type). The present paper, in contrast, starts from the assumption that the supervisor's type is unknown. We derive optimal contracts for this setting.[7]

Despite these informational asymmetries, the first-best solution can be implemented if the agent is unlimitedly liable. The corresponding optimal contract pays the supervisor a fixed wage, independent of how he evaluates the agent.

[4] This effect is particularly strong for poor performing employees.

[5] See Rotemberg & Saloner (1993), Prendergast & Topel (1996), Prendergast (2002b), Rotemberg (2004, 2007) or Sliwka (2007) for similar specifications.

[6] See, for example, Guilford (1954) or Kane et al. (1995).

[7] Most of the literature on alternative preferences assumes the parties' preferences to be common knowledge and so does not consider the problem we address. See, however, von Siemens (2008a, b) on the screening of inequity-averse agents.

This induces a supervisor who does not care for the agent's well-being to always report the agent's performance truthfully, while all other supervisor types report favorably independent of performance. Thus, the principal extracts all rents without eliminating the leniency bias, while the agent's incentives are reinstalled by a sufficiently high bonus payment for a good report. This suggests that leniency bias may be part of an optimal contractual arrangement under informational asymmetries.[8]

We also show that, under limited liability, the first-best solution is no longer attainable. Still, a contract paying the supervisor a fixed wage and not eliminating the leniency bias may be optimal in the class of pure–strategy direct mechanisms. A menu of contracts inducing the supervisor to report his type truthfully and to correctly evaluate the agent is feasible but may lead to excessive rents and may thus be suboptimal.

The paper is organized as follows: In the next section, we relate the paper to the literature. In Section 3, we present the model and two benchmark cases. Section 4 solves the model when the supervisor type is unknown to principal and agent. Section 5 concludes. All proofs are in the Appendix.

2 Related Literature

The paper is closely related to the literature on subjective performance evaluation. Particularly relevant is the work on compression of ratings, e.g. Baker

[8] For different explanations see MacLeod (2003) or Grund & Przemeck (2008).

et al. (1988), Murphy (1992), Harris (1994), Prendergast & Topel (1996) or Prendergast (2002b). In contrast to the present text, none of those papers deals with eliciting the supervisor's preferences.

The paper is clearly related to the literature on optimal contracts in three-tiered hierarchies consisting of a principal, a supervisor and an agent.[9] There are two crucial differences between those papers and ours. First, they typically assume hard information, i.e. a supervisor may conceal but not misrepresent information. In our model, information is soft.[10] Second, in those models the supervisor has known standard (egoistic) preferences. Eliciting the supervisor's preferences is, however, a core problem of the present paper.

Finally, the paper is also related to literature combining adverse selection and moral hazard, such as Laffont & Tirole (1986), McAfee & McMillan (1987) or Lewis & Sappington (1997). These papers assume the player choosing an unobservable action also to have superior information. In contrast, our supervisor has private information (his type) and a hidden action (to report effort truthfully or not) while the action of the agent is hidden from the principal but not from the supervisor.

[9] See e.g. Tirole (1986), Laffont & Martimort (1997), Strausz (1997) or Vafai (2004).

[10] See Faure-Grimaud et al. (2003) for a model with soft information.

3 The model

3.1 Description of the model and notation

Consider a stage game with a principal (P), a supervisor (S) and an agent (A). A exerts effort $e \geq 0$ at a cost $c(e)$ in order to produce output $f(e)$ that accrues to P.[11] The cost $c(e)$ is twice differentiable and satisfies $c(0) = 0$, $c'(0) = 0$, $c'(e) > 0$ for $e > 0$, $c'(e) = \infty$ for $e \to \infty$, and $c''(e) > 0$. Output $f(e)$ satisfies $f(0) = 0$, $f'(e) > 0$ and $f''(e) \leq 0$. Effort and output are unobservable by P and not verifiable by third parties. Moreover, there is no other objective measure of e. Hence, P cannot write an explicit performance contract to motivate A. She hires S whose only task is to monitor A and to observe and report the agent's effort choice.

P demands a certain effort $\hat{e}$ from A. As P cannot observe this effort, she asks S for a message or report $m \in \{y, n\}$. By reporting $m = y$, the supervisor states that A has chosen the required effort, i.e. $e \geq \hat{e}$. Instead, $m = n$ has the opposite meaning $e < \hat{e}$. Although S's observation is unobservable and unverifiable private information, the report is verifiable by third parties. For simplicity, S observes A's effort perfectly at no cost.[12]

[11] Note that we sometimes write e, f or c with subscripts that, depending on the context, either denote the supervisor's type or an optimal effort corresponding to a certain class of contracts.

[12] As an example, consider a supervisor and an agent who share an office, possibly working on the same project. Here, S might observe A's effort without additional effort or cost. We assume effort to be perfectly observable to present our results in the most

All parties are risk-neutral and have zero reservation value. A central assumption is that S cares for A's well-being and thus has an incentive to distort the effort report. We model S's preferences as[13]

$$U_S = w_S + \lambda U_A, \tag{1}$$

where w_S denotes the income of S and U_A the agent's utility. The parameter $\lambda \in [0, \bar{\lambda}]$, $\bar{\lambda} < 1$, measures how strongly S cares for A's well-being and is therefore multiplied with A's utility, which in turn is given by

$$U_A = w_A - c(e), \tag{2}$$

with w_A as the agent's income. Thus, the supervisor cares for A's effort cost and for payments the agent receives.

Except for a benchmark case in Subsection 3.3, the supervisor's preference, or type, is private information. In particular, the supervisor of type i ($i = 1, ..., M$), S_i, knows the parameter λ_i, while for P and A the type λ is a random variable with $\Pr\{\lambda = \lambda_i\} =: q_i$. Moreover, we assume that there are supervisors S_1 who do not care for the agent at all (i.e. $\lambda_1 = 0$): $q_1 > 0$.[14]

definite way. With an imperfect effort signal, a limitedly liable agent would earn a rent even if the signal was verifiable. This would make it harder to identify the rents induced by the combination of hidden information, hidden action, unverifiable information and the supervisor's leniency.

[13] As mentioned before, see Rotemberg & Saloner (1993), Prendergast & Topel (1996), Prendergast (2002b), Rotemberg (2004, 2007) or Sliwka (2007) for similar specifications.

[14] Note that $q_1 = 1$ is the standard case of egoistic preferences. Thus, we consider $q_1 > 0$ to be a reasonable assumption.

We confine analysis to pure strategies and direct incentive–compatible contracts where S reports his type truthfully and A makes the demanded effort.[15] A contract designed for supervisor type i is of the form $\{a_i, b_i, \hat{e}_i, w_{yi}, w_{ni}\}$. It specifies supervisor S_i's wages $\{w_{yi}, w_{ni}\}$ conditional on his reports $m = y$ resp. $m = n$. Moreover, the agent receives a base wage a_i, and, if the supervisor reports $m = y$, an additional bonus payment b_i. Contracts may be the same for different supervisor types (pooling).

The timing of the game is as follows: At stage 1, P offers a set of contracts in the presence of both S and A. At stage 2, S and A sign this contract set simultaneously (or the game ends). At stage 3, P may ask S to reveal his type, either publicly or in private (if P wants to conceal the type from A).[16] By announcing his type, S is understood to select one of the contracts from the set that has been signed. We assume that the set of signed contracts as well as the contract selected by S's type announcement can be verified by third parties. At stage 4, A chooses and S observes A's effort. At stage 5, S sends a message about the agent's effort to P. At stage 6, payments are made.

[15] W.l.o.g. we ignore contracts where P demands a certain effort but anticipates that A will choose a different effort. Then, equivalently, P can as well demand the effort that A is going to choose.

[16] Note that it does not matter whether S knows his type ex ante or learns it after signing the contracts but before A chooses effort. This is because contracts are incentive compatible for all types of supervisors. In this sense, it is irrelevant whether λ depends on the particular agent or describes a fixed characteristic of S.

Finally, we are well aware that collusion is an important issue in three-tiered hierarchies. A combined analysis of collusion and informational asymmetries, however, is beyond the scope of the paper. We thus assume the cost of collusion to be prohibitively high (e.g. due to legal sanctions).

3.2 The first-best solution

As a benchmark case, we derive the first-best solution assuming that effort is contractible and no contractual frictions arise. Here, the principal does not need a supervisor.[17] She simply imposes the effort level that solves

$$\max_{\hat{e},w_A} \pi = f(\hat{e}) - w_A \qquad s.t.\ w_A - c(\hat{e}) \geq 0$$

Hence, the first-best effort, e_{FB}, is given by

$$f'(e_{FB}) = c'(e_{FB}) \tag{3}$$

In the following we will repeatedly refer to the first–best problem resp. profit, $\max_{\hat{e}} f(\hat{e}) - c(\hat{e})$, and the corresponding first–best effort e_{FB} for comparison. Denote $f_{FB} := f(e_{FB})$ and $c_{FB} := c(e_{FB})$.

3.3 Complete Information

As a further benchmark, consider the case where the supervisor's type λ is common knowledge.[18] The optimal contracts solve the following program

[17] Note that P cannot do better by leaving a rent R to A and hiring a supervisor at a negative wage $w_S = -\lambda R$. This is due to $\lambda < 1$.

[18] Because preferences are known we drop the subscript i here.

where the supervisor reports the agent's effort truthfully:

$$\max_{a,b,w_y,w_n,\hat{e}} \pi = f(\hat{e}) - a - b - w_y \tag{4}$$

$$\begin{aligned}
s.t. \quad & (IC_A) \quad && a + b - c(\hat{e}) \geq a \\
& (IC_{S1}) \quad && w_n + \lambda (a - c(\tilde{e})) \geq w_y + \lambda (a + b - c(\tilde{e})), \ \forall\, \tilde{e} < \hat{e} \\
& (IC_{S2}) \quad && w_y + \lambda (a + b - c(\bar{e})) \geq w_n + \lambda(a - c(\bar{e})), \ \forall\, \bar{e} \geq \hat{e} \\
& (IR_A) \quad && a + b - c(\hat{e}) \geq 0 \\
& (IR_S) \quad && w_y + \lambda (a + b - c(\hat{e})) \geq 0
\end{aligned}$$

The incentive constraint (IC_A) ensures that A chooses the requested effort and does not deviate to $e = 0$. (IC_{S1}) and (IC_{S2}) make the supervisor report A's effort honestly. Finally, (IR_A) and (IR_S) ensure participation.

Proposition 1 *The optimal contract under complete information is* $a^* = 0$, $b^* = c_{FB}$, $\hat{e}^* = e_{FB}$, $w_y^* = 0$, $w_n^* = \lambda c_{FB}$. *It implements the first-best effort and profit.*

The supervisor's wages make him indifferent between reports $m \in \{y, n\}$, he earns a higher wage for reporting poor performance of the agent. This induces truthful reporting and the agent's effort becomes basically verifiable. Thus, the first–best solution is implemented. The proposed wages are non-negative which means that the contract is feasible even if we impose wealth constraints on S and A.

Interestingly, the first-best solution can also be implemented if the supervisor's type is unknown, as the following section shows.

4 Incomplete Information

Suppose the distribution of supervisor types is common knowledge while the realization λ_i is S_i's private information. The optimal contracts depend on whether the players are wealth–constrained. We analyze unlimited liability in subsection 4.1, while subsection 4.2 deals with the limited liability case.

4.1 Unlimited liability

Proposition 2 *If the players are unlimitedly liable, the optimal contract is unique and satisfies* $a^* = \frac{q_1 - 1}{q_1} c_{FB}$, $b^* = \frac{c_{FB}}{q_1}$, $\hat{e}^* = e_{FB}$, $w_y^* = w_n^* = 0$. *It implements the first-best effort and profit.*[19]

The supervisor's wage, $w_y^* = w_n^*$, is independent of his report. This induces a supervisor that does not care for A's well-being to report truthfully[20], while all other types report $m = y$ independent of A's effort choice. As a consequence, the leniency bias is not eliminated. This, however, does not

[19] When talking about uniqueness, we only consider sets of contracts where each constituent contract has positive acceptance probability. Given an optimal set of contracts, it is irrelevant if one adds contracts that are never accepted in equilibrium.

[20] One may criticize that, due to the indifference of S, he may as well misrepresent his information. If, however, we assume the supervisor to dislike lying only a little bit, the indifference is broken. Recent experimental evidence suggests that this is indeed a reasonable assumption (see e.g. Gneezy 2005). Note that this would also imply that one contract could induce several types of supervisors to report the agent's effort truthfully. Under a contract satisfying $w_n = w_y + \hat{\lambda} b$ supervisor types in an interval around $\hat{\lambda}$ would report truthfully.

cause any problems, since the bonus of the agent is adjusted in order to provide appropriate effort incentives, while the base wage is negative in order to avoid any rents for A.

The contract is unique, since all other types of contracts either leave a rent to certain types of supervisors or induce S not to accept the contract. The only contract that is accepted while not leaving a rent for S has a fixed wage equal to S's reservation utility. This contract does not work under limited liability since A's base wage is negative.

4.2 Limited liability

If wages must be non-negative, the contract of Proposition 2 is no longer feasible. We analyze the optimal contracts under wealth constraints. The problem, however, becomes far more complicated. To simplify the exposition but still highlight the effects at work we analyze the case of two supervisor types, 1 and 2, with $0 = \lambda_1 < \lambda_2 < 1$ that each occur with positive probability and $q_1 = 1 - q_2$. This is common knowledge.

With any contract, S_i is either indifferent between reports $m \in \{y, n\}$ (in which case truthful reporting is optimal for S_i), or he strictly prefers one of the reports in which case he reports independent of A's effort. Thus, with any optimal set of contracts where A exerts positive effort, at least one of the supervisor types must be induced to report effort truthfully. Therefore, any equilibrium contract must satisfy $w_y + \lambda_i b = w_n$ for at least one type of supervisor $i \in \{1, 2\}$. A contract where only one type, i, reports truthfully, is denoted by "S-pool-i". Contracts that induce both types to report truthfully are denoted by "S-sep".

The agent may get a contract that depends on the supervisor's type ("A-sep") or a contract that is independent of the supervisor's type ("A-pool"). In the latter case, P may reveal ("rev") or not reveal ("norev") the type to A.[21] Thus we have nine potential cases:

	A-rev-sep	A-rev-pool	A-norev-pool
S-pool-1	(A)	(B)	(C)
S-pool-2	(D)	(E)	(F)
S-sep	(G)	(H)	(I)

However, we can restrict attention to three of these cases.

Lemma 1 *Under limited liability the search for the optimal contract can be restricted to cases (C), (F), (G). The remaining cases are profit–dominated.*

Under contract set (C) both supervisor types get the same contract, S_1 reports truthfully, while S_2 always reports $m = y$ and the type is not revealed to A. Contract set (F) differs only in the fact that here S_2 tells the truth while S_1 reports $m = n$. Note that under both contracts, the type report of the supervisor is not used and thus not needed.

In contrast, under contract set (G), all wages and A's requested effort are type–dependent and the type is revealed to the agent.

[21] A's contract cannot be type–dependent if the type is to be concealed from him. And, obviously, it does not make sense to reveal only one of the types.

Which of these contract sets is optimal depends on the distribution of types (q_1), the parameter λ_2 and the output and cost function (f and c). Recall P's profits (see the proof of Lemma 1, subscripts C, F, G denote the cases):

$$(C)\ \pi_C = f_C - \frac{c_C}{q_1} \tag{5}$$

$$(F)\ \pi_F = f_F - c_F\left(1 + \frac{q_1\lambda_2}{1-q_1}\right) \tag{6}$$

$$(G)\ \pi_G = q_1\left(f_{FB} - c_{FB}\right) + (1-q_1)f_G - c_G\left(2 - q_1(2-\lambda_2)\right) \tag{7}$$

It is easy to see that for $q_1 \to 0$, $\pi_F \to \pi_{FB}$, while for $q_1 \to 1$, $\pi_G \to \pi_{FB}$ (with $e_G = 0$ and hence $f_G = c_G = 0$) and $\pi_C \to \pi_{FB}$.

Proposition 3 *Let the players be limitedly liable. Depending on the parameters of the model, any one of the contract sets (C), (F), (G) can be optimal.*

In the following we discuss why we cannot further reduce the set of possibly optimal contracts.

A contract of form (C) leaves an income rent to the agent, since S_2 reports $m = y$, i.e he does not "punish" underprovision of effort. This problem is especially severe if S_2 is very likely to occur. Accordingly, π_C is increasing in q_1.

Under a contract of form (F), S_1 receives an income rent (w_n) from lying and reporting $m = n$. This is of particular relevance if q_1 is high. Obviously, π_F is decreasing in q_1.

Finally, under contract set (G) S is presented with two contracts from which he chooses one. Recall that the two supervisor types are equal except for their care for A's well-being. In particular, their utility from a wage payment is the same. This implies that the only way to induce self-selection is to treat A differently depending on whether the supervisor claims to be of type 1 or 2. In fact, the principal has to give A a rent, if the supervisor chooses the contract designed for type 2. By limited liability, S_2 then has positive utility. This, in turn, requires $w_{n2} > 0$ in order to induce him to report truthfully. As the other supervisor does not care for the agent's well-being, such a contract is less desirable for him. Still, he has an incentive to claim to be of type 2, to report $m = n$, and get wage $w_{n2} > 0$. In order to destroy this incentive, type 1 receives an informational rent. Hence, there are two types of rent. If these rents become too high, a contract of form (G) is suboptimal.[22]

We have shown that under wealth constraints and informational asymmetries it can be optimal to have the simple contract (C) that pays S a fixed wage and does not eliminate the leniency bias. There, only the supervisor of type 1 tells the truth while all others report favorably independent of performance. Again, this simple contract combined with leniency bias can be observed in practice.

[22] If there are more than two types of supervisors ($M > 2$), the optimal separating contracts, (G), have similar characteristics. There, each type has a binding incentive to imitate higher types and report $m = n$ and at the same time the binding incentive to imitate lower types and report $m = y$. Nevertheless, full separation (self–selection) of types is feasible but very costly, which makes low or zero efforts for most types optimal. A corresponding proof can be downloaded from the authors' websites.

5 Conclusion

In this paper, we have considered the problem of designing optimal contracts in a principal-supervisor-agent relationship where the supervisor cares for the agent's well–being. We have shown that a simple contract paying the supervisor a fixed wage equal to his reservation value may be optimal and may even implement the first-best solution. Thus, not eliminating a leniency bias may be an optimal contractual arrangement under informational asymmetry. We discussed private information and limited liability for the case of two supervisor types. This restriction is due to the fact that the number of possible contracts grows exponentially in the number of types. Some types of contracts can be straightforwardly solved for arbitrary numbers of types or continuous type space. Contract type (F) generalizes to contracts in which some or all supervisor types are pooled at some higher type, i.e. this type reports truthfully. Optimality of these contracts unrealistically requires that this type has more probability mass than others, including, in particular, the type that does not care for the agent's well-being.[23]

Contract type (G) is always feasible, but with many types of supervisors it becomes prohibitively expensive to satisfy the many incentive constraints. Then very low or zero efforts are optimal for most types. Contract type

[23]It seems intuitive to assume that the latter type is the most probable type to occur. One could use a utility function similar to (1) to explain findings in experiments on dictator games. In these experiments, the mode is typically at zero. This means that the amount of money most often donated is zero. See e.g. Figure 4 in Bolton et al. (1998).

(C), on the other hand, is always feasible and becomes better with higher probability of the supervisor type that does not care for the agent. Thus we consider contract (C) to be an important candidate for the optimal contract. It also coincides with contracts observed in practice.

6 Appendix

Proof of Proposition 1. It is easily verified that the proposed contract satisfies (4). The principal's profit simplifies to the first–best profit, $\pi = f_{FB} - c_{FB}$. ■

Proof of Proposition 2. The proof is done in two steps. First, we show that the proposed contract yields the first-best solution and is therefore optimal. Second, we show that it is the only contract with this property.

1. If $w_y^* = w_n^* = 0$, a supervisor of type $\lambda_1 = 0$ reports the agent's effort truthfully, while all other types report $m = y$. Thus, A chooses effort $\hat{e}$, if

$$a + b - c(\hat{e}) \geq a + (1 - q_1)\, b \quad \Leftrightarrow \quad b \geq \frac{c(\hat{e})}{q_1}$$

Thus, under the proposed contract, A exerts the first–best effort $\hat{e} = e_{FB}$. Moreover, S and A accept the contract since their payoffs are $U_A = a^* + b^* - c_{FB} = 0$ and $U_S = \lambda U_A = 0$. Finally, P receives the first–best profit: $\pi = f_{FB} - b^* - a^* = f_{FB} - c_{FB}$.

2. The principal's expected profit can be written as

$$\pi = f(e) - E[w_S] - E[w_A]$$

where $E[w_S]$ and $E[w_A]$ are the expected wage payments to S and A. From S's participation constraints it follows that $E[w_S] \geq -\bar{\lambda} U_A$. Hence, the principal's profit is not higher than $f(e) + \bar{\lambda} U_A - E[w_A]$ or $f(e) + \bar{\lambda} U_A - (U_A + c(e))$. As $\bar{\lambda} < 1$, the first-best profit can only be achieved if $U_A = 0$, i.e. A must not receive a rent.

Assume now that the contract is such that no supervisor type reports the agent's output truthfully. Then the agent chooses $e = 0$, which is clearly suboptimal.

Moreover, a contract can only be optimal if S accepts it and no type of supervisor receives a positive rent. The latter requirement implies $\max\{w_y, w_n\} \leq 0$. Otherwise, a supervisor of type $\lambda_1 = 0$ receives a strictly positive rent. As $U_A = 0$, the only contract satisfying $\max\{w_y, w_n\} \leq 0$, inducing at least one type of S to report y truthfully and ensuring participation of S has $w_y = w_n = 0$.

■

Proof of Lemma 1. We analyze the cases separately and then show that only the cases (C), (F), (G) remain. Denote $f_i := f(e_i)$ and $c_i := c(e_i)$ for any subscript i.

(A) Here, both types of supervisor get the same wages with $w_y = w_n$ while A gets a type-dependent contract $(a_1, b_1, \hat{e}_1)$ and $(a_2, b_2, \hat{e}_2)$. The following result also holds for $M > 2$ types and also for continuous type space.

By $w_y = w_n$, supervisor 2 (or higher) prefers report $m = y$ whenever A has positive utility. Since A learns the type, he will never exert positive effort if type 2 (or higher) occurs. Thus, $e_2 = 0$ and $a_2 = b_2 = 0$.

Type 1 only cares about wages, thus there is no incentive to imitate type 2. Type 2, however, would like to imitate type 1 if A's utility from contract 1 is above that from contract 2. Given the above derived contract 2, we need $a_1 + b_1 - c_1 = 0$ to make supervisor 2 indifferent. The cost–minimizing way to achieve this with positive effort is $a_1 = 0$, $b_1 = c_1$. The cost–minimizing wages are $w_y = w_n = 0$. With these contracts, P attains zero profit if type 2 occurs and the first–best profit if type 1 occurs. The optimal contracts and the corresponding profit are:

$$w_{yi} = w_{ni} = a_i = 0, \; i = 1, \ldots, M \tag{8}$$

$$b_i = \begin{cases} c_{FB} & i = 1 \\ 0 & i = 2, \ldots, M \end{cases}, \; e_i = \begin{cases} e_{FB} & i = 1 \\ 0 & i = 2, \ldots, M \end{cases} \tag{9}$$

$$\pi_A = q_1 \left(f_{FB} - c_{FB} \right) \tag{10}$$

(The same result applies for continuous type space as long as type 1 has positive probability mass).

(B) Here, $w_y = w_n$ and the agent is informed about the type but gets a type–independent contract, i.e. P demands the same effort no matter which type is announced.

First, type 2 prefers $m = y$ and will lie about effort unless the requested effort is zero. Second, the agent knows the type and will exert zero effort if type 2 occurs. Thus, there is no uniform positive effort that the agent would exert in presence of both types. This means that in case (B) the only implementable effort is zero.[24]

(C) Here, $w_y = w_n$, the type is not revealed and A gets a type–independent contract $(a, b, \hat{e}_C)$. S_2 prefers report $m = y$.

The agent knows that type 1 reports truthfully while type 2 always reports $m = y$. Thus, the agent will exert effort if

$$a + b - c_C \geq a + q_2 b \;\Rightarrow\; b \geq \frac{c_C}{1 - q_2}$$

There is no imitation issue here because the type information is not used, so P would not have to ask for S's type at all. The optimal

[24] One may argue that P could offer a pooling contract with a strictly positive requested effort, but accept that the agent deviates if the supervisor is of type 2. But then P can as well demand different type–dependent efforts and obtain the same outcome (see (A)). As mentioned before, w.l.o.g. we restrict attention to contracts where A does not have an incentive to deviate from the requested effort.

contract and P's profit are

$$a = 0,\ b = \frac{c_C}{1 - q_2},\ \hat{e}_C > 0,\ w_y = w_n = 0 \tag{11}$$

$$\pi_C = f_C - \frac{c_C}{1 - q_2} = f_C - \frac{c_C}{q_1} \tag{12}$$

(D) Here, both types of supervisor get the same wages with $w_y + \lambda_2 b_2 = w_n$ while A gets a type-dependent contract $(a_1, b_1, \hat{e}_1)$ and $(a_2, b_2, \hat{e}_2)$.

Since $\lambda_1 = 0$, supervisor 1 prefers w_n and thus A will not exert positive effort if type 1 is announced. Thus, cost minimization implies $a_1 = b_1 = \hat{e}_1 = 0$. Supervisor 2 must be prevented from imitating type 1 with report $m = n$. This would give him a utility of w_n since A's utility is zero. Thus contract 2 must give him at least the same utility. However, the only way to achieve this is to increase a_2. Note that changing w_y or b_2 only increases w_n and therefore does not solve the problem. Moreover, supervisor 1 is indifferent towards changes in A's wage. The optimal contracts for case (D) are:

$$a_1 = b_1 = \hat{e}_1 = 0,\ a_2 = b_2 = c_2,\ \hat{e}_2 > 0,\ w_y = 0,\ w_n = \lambda_2 c_2$$

where c_2 is the cost associated with the profit–maximizing effort. Denoting $c_D := c_2$ and $f_D := f_2$ we get

$$\pi_D = -q_1 \lambda_2 c_2 + q_2 \left(f_2 - 2c_2\right) \tag{13}$$

$$= \left(1 - q_1\right)\left(f_D - c_D\right) - c_D \left(1 - q_1(1 - \lambda_2)\right) \tag{14}$$

(E) Here, $w_y + \lambda_2 b = w_n$, and the agent is informed about the type and gets a type-independent contract. An argument similar to the one in (B) applies and the only implementable effort is again zero.

(F) Here, the supervisors get the type–independent contract $w_y + \lambda_2 b = w_n$, and the type is not revealed to A. Thus, P doesn't need to ask the supervisor for his type and there is no imitation issue. Type 1 always reports $m = n$ while type 2 tells the truth. A gets a type–independent contract $(a, b, \hat{e}_F)$ and exerts positive effort if his bonus satisfies

$$a + q_2 b - c_F \geq a \;\Rightarrow\; b \geq \frac{c_F}{1 - q_1}$$

We get the optimal contract by setting the cost–minimizing wages:

$$a = 0,\; b = \frac{c_F}{1 - q_1},\; \hat{e}_F > 0,\; w_y = 0,\; w_n = \lambda_2 \frac{c_F}{1 - q_1}$$

The principal's profit is

$$\pi_F = f_F - q_1 \lambda_2 \frac{c_F}{1 - q_1} - q_2 \frac{c_F}{1 - q_1} = f_F - c_F \left(1 + \frac{q_1 \lambda_2}{1 - q_1}\right)$$

(G) Here, the supervisors get contracts with $w_{y1} = w_{n1}$ and $w_{y2} + \lambda_2 b_2 = w_{n2}$, the agent knows the type and gets type-dependent contracts $(a_1, b_1, \hat{e}_1)$ and $(a_2, b_2, \hat{e}_2)$.

Supervisor S_i prefers selecting the right contract if

$$\begin{aligned} S_2: \quad & w_{y2}+\lambda_2(a_2+b_2-c_2) \geq \\ & \qquad \max\left\{w_{y1}+\lambda_2(a_1+b_1-c_1), w_{n1}+\lambda_2(a_1-c_1)\right\}, \qquad (15) \\ S_1: \quad & w_{y1}+\lambda_1(a_1+b_1-c_1) \geq \\ & \qquad \max\left\{w_{y2}+\lambda_1(a_2+b_2-c_2), w_{n2}+\lambda_1(a_2-c_2)\right\}. \qquad (16) \end{aligned}$$

Applying the wages from above, these conditions simplify to

$$w_{y2}+\lambda_2(a_2+b_2-c_2) \geq w_{y1}+\lambda_2(a_1+b_1-c_1), \tag{17}$$

$$w_{y1} \geq w_{y2}+\lambda_2 b_2 \tag{18}$$

which implies

$$a_1+b_1-c_1 \leq a_2-c_2$$

As the left-hand-side of the inequality is non-negative, (IR_A), we obtain $a_2 \geq c_2$. By cost minimization, $a_2 = c_2$ and $a_1 = 0$ (limited liability). Moreover, from A's incentive compatibility constraints, we get $b_1 = c_1$ and $b_2 = c_2$. Finally, $w_{y2} = 0$, $w_{n2} = \lambda_2 c_2$, and using (18), $w_{y1} = \lambda_2 c_2$ and $w_{n1} = \lambda_1 c_1 + \lambda_2 c_2$. The principal's profit is:

$$\begin{aligned} \pi_G &= q_1\left(f_1-c_1-\lambda_2 c_2\right)+q_2\left(f_2-2c_2\right) & (19) \\ &= q_1\left(f_1-c_1\right)+\left(1-q_1\right)\left(f_2-c_2\right)-c_2\left(1-q_1(1-\lambda_2)\right) & (20) \end{aligned}$$

Obviously, $\hat{e}_1 = e_{FB}$ and $\hat{e}_2 < e_{FB}$. The profit is below the first–best profit since S_1 gets a rent and A gets a rent if S_2 occurs. Denote

$c_G := c_2$ and $f_G := f_2$. Thus,

$$\pi_G = q_1 \left(f_{FB} - c_{FB}\right) + (1 - q_1) \left(f_G - c_G\right) - c_G \left(1 - q_1(1 - \lambda_2)\right) \quad (21)$$

$$= q_1 \left(f_{FB} - c_{FB}\right) + (1 - q_1) f_G - c_G \left(2 - q_1(2 - \lambda_2)\right) \quad (22)$$

(H),(I) Here, the type–dependent contracts for S must satisfy $w_{y1} = w_{n1}$ and $w_{y2} + \lambda_2 b = w_{n2}$ while the agent gets a type–independent contract $(a, b, \hat{e})$. In order for supervisors to reveal their type truthfully, it is necessary that

$$w_{y1} \geq w_{n2}, \quad (23)$$

$$w_{y2} + \lambda_2(a + b - c(e)) \geq w_{y1} + \lambda_2(a + b - c(e)) \;\Rightarrow\; w_{y2} \geq w_{y1} \quad (24)$$

However, from these we get $w_{y2} \geq w_{n2}$, which contradicts $w_{y2} + \lambda_2 b = w_{n2}$, since positive effort requires $b \geq c(\hat{e}) > 0$. Accordingly, the only implementable effort is zero.

The contract types (B), (E), (H), and (I) are not optimal since they cannot induce the agent to choose positive effort. Moreover, $\pi_G \geq \pi_A$, since by setting $\hat{e}_G = 0$ we get $\pi_G = \pi_A$. Thus, (A) is dominated. In addition, (D) is dominated since $\pi_G \geq \pi_D$ if the same effort is chosen in both cases, $\hat{e}_G = \hat{e}_D$. Thus, only (C), (F), (G) remain. ■

Proof of Proposition 3. Consider the following example. Output and cost are $f(e) = e$, $c(e) = \frac{e^3}{3}$ and $q_1 = \frac{3}{4}$. We get $e_{FB} = 1$, $\pi_C = \hat{e}_C - \frac{4\hat{e}_C^3}{9}$, $\pi_F = \hat{e}_F - \frac{1}{3}\hat{e}_F^3(1+3\lambda_2)$ and $\pi_G = \frac{1}{12}\left(6 + 3\hat{e}_G - \hat{e}_G^3(2 + 3\lambda_2)\right)$. The principal's maximum profits are

$$\pi_C = \frac{1}{\sqrt{3}},\ \pi_F = \frac{2}{3\sqrt{1+3\lambda_2}},\ \pi_G = \frac{1}{6}\left(3 + \frac{1}{\sqrt{2+3\lambda_2}}\right) \tag{25}$$

It can be verified that (F) is optimal for $0 < \lambda_2 \leq 0.061$, (G) is optimal for $0.061 < \lambda_2 < 0.881$ and (C) is optimal for $\lambda_2 > 0.881$. ■

References

Baker, G. (1992): Incentive Contracts and Performance Measurement, *Journal of Political Economy*, 100, pp. 598-614.

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

Baker, G., Jensen, M. C. and K.J. Murphy (1988): Compensation and Incentives; Practice vs. Theory, *Journal of Finance*, 3, pp. 593-616.

Bolton, G.E., Katok, E. and R. Zwick (1998): Dictator game giving: Rules of fairness versus acts of kindness, *International Journal of Game Theory*, 27, pp. 269-299.

Faure-Grimaud, A., Laffont, J.J. and D. Martimort (2003): Collusion, Delegation and Supervision with Soft Information, *Review of Economic Studies*, 70, pp. 253-279.

Feltham, G.A. and J. Xie (1994): Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations, *The Accounting Review*, 69, 429-453.

Gibbons, R. (1998): Incentives in Organizations, *Journal of Economic Perspectives*, 12, pp. 115-132.

Gibbons, R. (2005): Incentives Between Firms (and Within), *Management Science*, 51, pp. 2-17.

Gneezy, U. (2005): Deception: The Role of Consequences, *American Economic Review, 95,* pp. 384-394.

Grund, C. and J. Przemeck (2008): Subjective Performance Evaluation and Inequality Aversion, Mimeo, University of Bonn and University of Würzburg.

Guilford, J.P. (1954): Psychometric Methods, New York: McGraw-Hill.

Harris, M. M. (1994): Rater Motivation in the Performance Appraisal Context: A Theoretical Framework, *Journal of Management*, 20, pp. 737-756.

Holmström, B. (1982): Moral Hazard in Teams, *Bell Journal of Economics*, 13, pp. 324-340.

Holmström, B. and P. Milgrom (1991): Multitask principal-agent analysis: Incentive contracts, asset ownership, and job design, *Journal of Law,* *Economics, and Organization*, 7, pp. 24-52.

Kane, J.S., Bernardin, H.J., Villanova, P. and J. Peyrefitte (1995): Stability of Rater Leniency: Three Studies, *Academy of Management Journal,* 38, 1036-1051.

Kerr, S. (1975): On the Folly of Rewarding A While Hoping for B, *Academy of Management Journal,* 18, 769-783.

Laffont, J.J. and D. Martimort (1997): Collusion under Asymmetric Information, *Econometrica*, 65, pp. 875-911.

Laffont, J.J. and J. Tirole (1986): Using Cost Observation to Regulate Firms, *Journal of Political Economy*, 94, pp. 614-641.

Lewis, T.R. and D.E.M. Sappington (1997): Information Management in Incentive Problems, *Journal of Political Economy*, 105, pp. 796-821.

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

McAfee, R.P. and J. McMillan (1987): Competition for agency contracts, *RAND Journal of Economics*, 18, 296-307.

Murphy, K.J. (1992): Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance. William Bruns (ed.): Performance Measurement, Evaluation, and Incentives. Harvard Business School 1992, pp. 37-62.

Prendergast, C. (1999): The Provision of Incentives in Firms, *Journal of Economic Literature*, 37, pp. 7-63.

Prendergast, C. (2002a): The Tenuous Trade-off between Risk and Incentives, *Journal of Political Economy*, 110, pp. 1071-1100.

Prendergast, C. (2002b): Uncertainty and Incentives, *Journal of Labor Economics*, 20, pp. S115-S137.

Prendergast, C. and R. Topel (1996): Favoritism in Organizations, *Journal of Political Economy*, 104, pp. 958-978.

Rotemberg, J. J. (2004): Fair Pricing, NBER Working Paper Series, No. w10915.

Rotemberg, J. J. (2007): Attitude-Dependent Altruism, Turnout and Voting, mimeo, Harvard University.

Rotemberg, J. J. and G. Saloner (1993): Leadership Style and Incentives, *Management Science,* 39, pp. 1299-1318.

Sliwka, D. (2007): Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes, *American Economic Review,* 97, pp. 999-1012.

Strausz, R. (1997): Delegation of Monitoring in a Principal-Agent Relationship, *Review of Economic Studies*, 64, 337-357.

Tirole, J. (1986): Hierarchies and Bureaucracies: On the Role of Collusion in Organization, *Journal of Law, Economics, and Organization*, 2, pp.

181-214.

Vafai, K. (2004): Delegation and Opportunism, *Journal of Institutional and Theoretical Economics*, 160, pp. 498-521.

von Siemens, F. (2008a): Heterogeneous Social Preferences, Screening, and Employment Contracts, mimeo, University of Munich.

von Siemens, F. (2008b): Social Preferences, Sorting, and Competition, mimeo, University of Munich.