

Gürtler, Marc; Gürtler, Oliver

**Working Paper**

## The interaction of explicit and implicit contracts: A signaling approach

Working Paper Series, No. IF38V1

**Provided in Cooperation with:**

Technische Universität Braunschweig, Institute of Finance

*Suggested Citation:* Gürtler, Marc; Gürtler, Oliver (2012) : The interaction of explicit and implicit contracts: A signaling approach, Working Paper Series, No. IF38V1, Technische Universität Braunschweig, Institut für Finanzwirtschaft, Braunschweig, <https://doi.org/10.2139/ssrn.2031582>

This Version is available at:

<https://hdl.handle.net/10419/57177>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Interaction of Explicit and Implicit Contracts: A Signaling Approach

Marc Gürtler\*, *Braunschweig Institute of Technology*

Oliver Gürtler\*\*, *University of Cologne*

## Abstract

We analyze the interaction of explicit and implicit contracts in a model with selfish and fair principals. Fair principals are willing to honor implicit agreements, whereas selfish principals are not. Principals are privately informed about their types. We investigate a separating equilibrium in which principals reveal their type through the contract offer to the agent. If this equilibrium is played, explicit and implicit contracts are substitutes. Since the agent learns the principal's type, a selfish principal has to rely on explicit incentives. A fair principal, by contrast, can effectively induce implicit incentives and hence does not need to use explicit incentives. Interestingly, if a selfish principal can rely on more effective explicit incentives, a fair principal becomes more likely to be able to separate from the selfish type and, hence, to make better use of implicit incentives. In this sense, there is a strategic complementarity between explicit and implicit incentives.

**Keywords:** explicit contracts, implicit contracts, separating equilibrium, substitutes, strategic complementarity

**JEL classification:** D82, D86, M52

\*Marc Gürtler, Department of Economics, Braunschweig Institute of Technology, Abt-Jerusalem-Str. 7, D-38106 Braunschweig, Germany. Phone: +49-531-3912895; E-mail: marc.guertler@tu-bs.de

\*\*Oliver Gürtler, Department of Economics, University of Cologne, Albertus Magnus Platz, D-50923 Cologne, Germany. Phone: +49-221-4701450; E-mail: oliver.guertler@uni-koeln.de

# 1 Introduction

Explicit (or formal) contracts are based on verifiable information and thus can be enforced by a court. Implicit (or informal) contracts instead rely on observable but nonverifiable information. Therefore, they cannot be enforced by a court and must be self-enforcing. In the economics literature, implicit contracts have been modeled in two ways. In one class of models the game is played infinitely often and implicit contracts are sustained by supergame strategies that involve some type of punishment if the contract is reneged on.<sup>1</sup> A second class of models assumes that players have some form of other-regarding preference and therefore find it in their interest not to betray their co-player.<sup>2</sup>

Both explicit and implicit contracts are prevalent in practice. For instance, firms often tie employee compensation to objective measures of performance (such as total sales or the number of items produced) but also evaluate (and reward) employee performance subjectively. Competing firms may enter a formal contract to govern a research cooperation and, at the same time, use implicit contracts to sustain high prices for their products. Because of the prevalence of the two types of contract, it is important to understand whether the existence of formal contracts simplifies the sustainment of implicit contracts or, in other words, whether formal and informal contracts are complements or substitutes. Several papers have addressed this question. Baker et al. (1994) and Schmidt and Schnitzer (1995), for example, demonstrate that, depending on the economic environment, both types of relationship are feasible and that the existence of explicit contracts may make implicit contracts easier or more difficult to sustain. Papers analyzing the interaction of explicit and implicit contracts typically belong to the first class of models, in which implicit contracts are sustained

---

<sup>1</sup>See, among many others, MacLeod and Malcomson (1989), Baker et al. (1994, 1999, 2002), Che and Yoo (2001) and Levin (2002, 2003).

<sup>2</sup>See, for example, Fehr and Schmidt (1999), James Jr (2002) and Sliwka (2007). Furthermore, note that implicit incentives may result from a player's career concerns as well. In corresponding models, the player has an incentive to choose a certain action to favorably affect the market perception regarding his characteristics (Holmström 1982, Dewatripont et al. 1999a,b, Irlenbusch and Sliwka 2006). A contract is not needed to incentivize the player so that the problem of contract enforcement does not arise.

within the framework of an infinitely repeated game. Instead, the current paper addresses the problem by considering a model with players who honor implicit contracts because of other-regarding preferences. In this way, we aim to highlight different effects and obtain new insights into the functioning of implicit contracts and their relation to explicit ones. This approach is also relevant because many experimental results suggest that other-regarding preferences are an important driver of individual behavior.<sup>3</sup>

In our model there is a principal and an agent. The agent chooses effort to produce output that accrues to the principal. Output is nonverifiable and can only be used in an implicit contract. The principal can, however, monitor the agent, in which case the latter must choose a certain minimum level of effort that is enforceable by a court. There are two types of principal. One type is selfish and would not keep any promise to pay the agent an output-dependent reward. The second type also cares for the agent's wellbeing. This type of principal is called fair and is willing to reward the agent for high output. Initially, the agent does not know the principal's type. Therefore, implicit incentives based on output are not very effective, because the agent encounters the risk of facing a selfish principal and not being rewarded for good performance. Under certain circumstances, however, there exists a separating equilibrium, in which both types of principal offer different contracts to the agent so that the agent learns the principal's type from the contract offer. We study this type of equilibrium in detail and find that explicit and implicit contracts are substitutes. Since the agent learns the principal's type, a selfish principal cannot credibly promise to honor an implicit contract. Instead, he has to rely on monitoring (i.e., on using explicit incentives). A fair principal, by contrast, can effectively induce implicit incentives and hence does not need to monitor the agent. Interestingly, if a selfish principal can rely on more effective explicit incentives, it becomes more likely that a fair principal can be separated from the selfish type. This means that the possibility of offering better explicit incentives for one type of principal allows the other type of principal to offer more effective implicit incentives. In this sense, there is a strategic complementarity between explicit and implicit incentives (Bulow et al.

---

<sup>3</sup>See Fehr and Schmidt (1999) and references therein.

1985).

The remainder of the paper is organized as follows. The next section discusses related literature on this topic. In Section 3, the model is presented and in Section 4 it is solved. Section 5 concludes. All formal proofs can be found in the Appendix.

## 2 Related literature

Most closely related to the current paper is the study by Sliwka (2007), who considers a principal–agent model with three types of agent: selfish agents who want to maximize their payoff, fair agents who take the principal’s payoff into account besides their own payoff, and conformist agents who are either selfish or fair, depending on what they believe most other agents to be. It is assumed that the principal privately knows the proportion of each agent type. Sliwka (2007) shows that under certain conditions there exists a separating equilibrium in which the principal does not offer a performance contract if and only if the majority of the agents are fair. By not offering a performance contract, the principal thus “turns” the conformists into fair agents. By offering a performance contract, the principal instead signals that the majority of agents are selfish and that he distrusts the agents. If, in the current paper, the principal decides to monitor the agent, he signals that he himself is selfish and cannot be trusted to fulfill an implicit contract. In this sense, the result is exactly opposite to that in Sliwka (2007).<sup>4</sup> Since implicit contracts cannot be sustained, there is a cost for monitoring the agent. Therefore, the paper is also related to the study by Falk and Kosfeld (2006), who use experimental data to show that there is a cost of control. However, they also argue that control signals distrust of the co-player and not of the player himself.

The paper is also related to literature that analyzes the interaction of explicit contracts (or other modes of governance such as a particular ownership structure) and implicit contracts. Examples include Baker et al. (1994, 2002), Schmidt and Schnitzer (1995), Halonen

---

<sup>4</sup>Note furthermore that in the Sliwka (2007) model the existence of conformist players is crucial to the results. In the current paper, we confine our attention to fair and selfish players and still are able to demonstrate the existence of a separating equilibrium with different principal types offering different contracts.

(2002) and Schöttner (2008). These papers analyze how the presence of contractible modes of governance affects the ability of the parties to rely on implicit contracts. The answer is not straightforward. In certain situations, implicit contracts become easier to sustain. In other situations it becomes more difficult to sustain such contracts. Note that all the papers mentioned above consider models in which a stage game is played infinitely often and implicit contracts are sustained by punishment strategies in the corresponding supergame. By considering a one-period model with other-regarding principals, the current paper deviates from this approach and therefore highlights different effects regarding the interaction of explicit and implicit contracts.

In response to the development of formal theories of other-regarding preferences, research has started to incorporate these types of preferences into contract-theoretic models. Grund and Sliwka (2005), for instance, consider a tournament model with inequality-averse contestants. A dynamic team production model with inequality-averse players is dealt with by Mohnen et al. (2008). Neilson and Stowe (2010) consider a model in which two inequality-averse agents are hired by a principal and offered piece-rate contracts. Whereas these models assume (horizontal) inequality aversion between several agents, Englmaier and Wambach (2010) consider a model with (vertical) inequality aversion between a principal and an agent. Our paper is of course related to this literature. It differs by assuming that the principal's type is his private information. Most of the papers on other-regarding preferences assume that preferences are commonly known.<sup>5</sup> Therefore, these papers do not analyze the possibility of signaling, which is at the core of the current paper.

Finally, the paper is related to the literature on informed-principal problems (Maskin and Tirole 1990, 1992, Spier 1992, Beaudry 1994, Jost 1996). This literature investigates how a principal can reveal his private information through the choice of contract. In our model, the principal uses a contract offer to signal whether he is selfish or cares for the agent's wellbeing. This is different from previous papers, which focus on other aspects of the economic environment (such as the productivity of the agent's effort).

---

<sup>5</sup>There are a few notable exceptions, such as von Siemens (2011, 2012) and Giebe and Gürtler (2012).

### 3 Model description and notation

A principal  $P$  hires an agent  $A$  to produce some output  $y$ . Output accrues to  $P$  and is observable by  $P$  and  $A$ , but is nonverifiable by a third party (e.g., a court). Output is either high ( $y = y_H$ ) or low ( $y = y_L < y_H$ ). The probability of high output realization depends on the agent's effort  $e \in [0, 1]$ . W.l.o.g. we set this probability equal to  $e$ . Effort is costly to the agent and these costs are given by  $c(e)$  with  $c(0) = 0$ ,  $c' \geq 0$ , and  $c'' > 0$ . Furthermore, we assume that (weakened) Inada conditions hold, i.e.,  $\lim_{e \rightarrow 0} c'(e) = 0$  and  $\lim_{e \rightarrow 1} c'(e) > \frac{y_H - y_L}{2}$ , implying that  $A$  never finds it optimal to choose  $e = 1$  in equilibrium. Let  $w \in \mathbb{R}$  denote the (total) wage payment from  $P$  to  $A$ . The agent's utility is additively separable in income and effort costs and he is risk-neutral, i.e.,

$$u_A(w, e) = w - c(e).$$

His reservation utility is given by  $\bar{u} = 0$ .

The principal is risk-neutral and his utility function is given by

$$u_p(y, w, e) = y - w - \alpha \cdot \max\{w - (y - w), 0\} - \beta \cdot \max\{(y - w) - w, 0\}.$$

The parameters  $\alpha \geq 0$  and  $\beta \geq 0$  characterize the type of principal. There are two possible types characterized by either  $\alpha + \beta = 0$  or  $\alpha + \beta > 0$ . If  $\alpha + \beta = 0$ , we have the standard model in which relative comparisons do not matter. This parameter set characterizes a selfish principal. For  $\alpha + \beta > 0$ , the principal is inequality-averse. To be more specific,  $\alpha > 0$  measures the envy he feels when being worse off than the agent and  $\beta > 0$  measures the compassion he feels when he is relatively better off (Fehr and Schmidt 1999). We call the inequality-averse principal fair. The principal's type is his private information. At the outset, the agent assumes that  $P$  is selfish with probability  $q_s \in (0, 1)$  and inequality-averse with probability  $1 - q_s$ . He uses Bayes' rule to update his belief. Throughout the analysis, we assume that the utility function for an inequality-averse principal satisfies  $\alpha \geq \beta > 0.5$ . This is in line with the findings in Table III of Fehr and Schmidt (1999).

The principal can affect the agent's effort by deciding to monitor him. We assume that monitoring is costless (to highlight indirect costs of control) and, following Falk and Kosfeld

(2006), that it restricts the agent's choice set. To be more specific, let  $\hat{e} \in \{0, \bar{e}\}$  describe the minimum level of effort the agent must choose. If the agent is not monitored, we have  $\hat{e} = 0$ . Otherwise, the agent must provide effort of at least  $\bar{e} \in (0, 1)$ . We assume

$$c'(\bar{e}) \leq \min \left\{ \frac{y_L}{2}, \frac{y_H - y_L}{2} \right\} \quad (1)$$

to ensure that implicit incentives can work better than monitoring, as shown later.

The agent's effort can also be affected by means of incentive pay. The principal can promise to pay the agent a bonus  $\hat{b}_i \in \mathbb{R}_+$  in addition to a fixed wage  $\hat{w}$  if the agent produces output  $y_i$  ( $i = H, L$ ). We impose  $\hat{b}_H \geq \hat{b}_L$ , as this inequality is required to incentivize the agent. Note, however, that there is a potential commitment problem. If  $A$  produces high output, the principal may claim that output was low in order to lower his wage costs. The commitment problem occurs since output is nonverifiable by third parties so a court cannot assess the agent's true output. In other words, the principal is contractually obliged to pay the agent  $\hat{b}_L$ , but not to pay  $\hat{b}_H$ . W.l.o.g., we can set  $\hat{b}_L = 0$  in the contract, because  $\hat{b}_L$  is a type of fixed wage that can be incorporated into  $\hat{w}$ . Note that the bonuses that are actually paid in equilibrium can deviate from the contractually specified bonuses  $\hat{b}_L = 0$  and  $\hat{b}_H \geq 0$ . If the principal finds it optimal to pay the agent  $\hat{b}_H$  or more, he can simply do so (the agent will never refuse to take the money). Similarly, if he wants to pay the agent a bonus between zero and  $\hat{b}_H$ , he can do so by claiming that output was low (in which case the contractually specified bonus is zero), but pay the agent a positive amount of money. To summarize, there is no way to structure the contract such that the bonus payment is restricted; every positive amount is feasible. We therefore do not explicitly account for bonus payments in the contract and assume that a contract is given by a tuple  $(\hat{w}, \hat{e})$ . Still, we keep in mind that the principal may voluntarily decide to reward the agent for his performance.

The timing of the model is as follows. In the first stage, nature determines the principal's type. In the second stage, the principal offers a contract to the agent. If the agent rejects the contract, the game ends. Otherwise it proceeds to stage 3, when the agent chooses his effort. In stage 4, output is realized. In stage 5, the principal decides whether to pay the agent a bonus in addition to the fixed wage he is required to pay by the contract.



## 4 Model solution

The model is solved by backward induction, so we start with stage 5. Let  $b_i$  denote the bonus that is actually paid by the principal if  $y_i$  is observed. Obviously, a selfish principal will never pay anything to the agent because a bonus payment is made voluntarily. Things are different for the inequality-averse principal with  $\alpha + \beta > 0$ , whose optimization problem corresponds to

$$\max_w [y - w - \alpha \cdot \max\{w - (y - w), 0\} - \beta \cdot \max\{(y - w) - w, 0\}].$$

Consequently, the objective function is differentiable at all points  $w$  with  $w \neq \frac{y}{2}$  and the marginal utility corresponds to

$$-1 - 2 \cdot \alpha \cdot \chi_{(y/2, \infty)}(w) + 2 \cdot \beta \cdot \chi_{(-\infty, y/2)}(w),$$

where  $\chi_A$  denotes the characteristic function of set  $A$ , i.e.  $\chi_A(w) = 0$  if  $w \notin A$  and  $\chi_A(w) = 1$  if  $w \in A$ . Thus, the marginal utility is positive (negative) iff  $w < \frac{y}{2}$  ( $w > \frac{y}{2}$ ) and the maximum of the objective function is at  $w_{ia}^* = \frac{y}{2}$ . Bearing in mind that the bonus payment must be non-negative, we obtain

$$b_i(\hat{w}) = \max\left\{\frac{y_i}{2} - \hat{w}, 0\right\} \quad (2)$$

as the bonus payment by an inequality-averse principal if output level  $y_i$  is observed.

Now consider the third stage. We denote by  $p_s \in [0, 1]$  the agent's posterior belief regarding the principal type (after being offered a contract). This means that the agent believes to meet a selfish principal with probability  $p_s$ . Then the agent chooses effort  $e$  to solve

$$\max_{e \in [\hat{e}, 1]} [\hat{w} - c(e) + (1 - p_s) \cdot (e \cdot b_H(\hat{w}) + (1 - e) \cdot b_L(\hat{w}))].$$

In the case of an interior solution, the optimal effort  $e^*$  thus satisfies

$$(1 - p_s) \cdot (b_H(\hat{w}) - b_L(\hat{w})) = c'(e^*).$$

Consequently,<sup>6</sup>

$$e^*(p_s) = \max\{c'^{-1}((1 - p_s) \cdot (b_H(\hat{w}) - b_L(\hat{w}))), \hat{e}\}.$$

---

<sup>6</sup>Note that  $c' : [0, 1] \rightarrow \mathbb{R}$  is injective. Thus, the inverse  $c'^{-1} : c'([0, 1]) \rightarrow [0, 1]$  exists.

From the above condition, a dilemma for the inequality-averse principal becomes evident. The agent's equilibrium effort is increasing in  $(1 - p_s)$  and thus the inequality-averse principal would gain by revealing his type. However, the selfish principal clearly has an incentive to hide his type. In the second stage of the model, to which we turn now, we therefore focus on the following question: can there be a separating equilibrium in which the two types of principal offer different contracts and the contract offer reveals the principal's type?<sup>7</sup>

The next proposition characterizes the optimal contract for a selfish principal and his expected profit in such a separating equilibrium.

**Proposition 1** *In a separating equilibrium, the selfish principal offers a contract  $(\hat{w}_s, \hat{e}_s) = (c(\bar{e}), \bar{e})$  and receives expected profit equal to  $y_L + \bar{e} \cdot (y_H - y_L) - c(\bar{e})$ . In this case,  $\hat{w}_s \leq \frac{y_L}{2}$  and  $b_i(\hat{w}_s) = \frac{y_i}{2} - \hat{w}_s$  ( $i \in \{L, H\}$ ).*

Proposition 1 simplifies the analysis by specifying the contract that the selfish principal offers in the separating equilibrium. Because his type is revealed in such an equilibrium, he can never credibly promise to pay the agent a bonus and therefore must rely on explicit incentives.

We denote the optimal contract of the inequality-averse principal in the separating equilibrium by  $(\hat{w}_{ia}, \hat{e}_{ia}) \neq (\hat{w}_s, \hat{e}_s)$ . To ensure the existence of a separating equilibrium, each type of principal has to prefer his own contract to the contract of the other type.<sup>8</sup> The optimal contract  $(\hat{w}_{ia}, \hat{e}_{ia})$  therefore satisfies the following program:

$$\begin{aligned} & \Phi_{ia}(\hat{w}_{ia}, \hat{e}_{ia}) \\ &= e^*(0) \cdot [y_H - b_H(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_H(\hat{w}_{ia}) + \hat{w}_{ia}) - y_H)] \\ & \quad + (1 - e^*(0)) \cdot [y_L - b_L(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_L(\hat{w}_{ia}) + \hat{w}_{ia}) - y_L)] - \hat{w}_{ia} \\ & \rightarrow \max_{\hat{w}_{ia}, \hat{e}_{ia}}! \end{aligned}$$

---

<sup>7</sup>In Section 5, we briefly address other types of equilibrium.

<sup>8</sup>Of course, each type of principal must prefer his own contract to any other contract as well. This issue is addressed in the Appendix.

subject to

$$\begin{aligned}
& y_L + \bar{e} \cdot (y_H - y_L) - \hat{w}_s \geq y_L + e^*(0) \cdot (y_H - y_L) - \hat{w}_{ia} \\
& \iff (e^*(0) - \bar{e}) \cdot (y_H - y_L) \leq \hat{w}_{ia} - \hat{w}_s \quad (IC_s) \\
& e^*(0) \cdot [y_H - b_H(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_H(\hat{w}_{ia}) + \hat{w}_{ia}) - y_H)] \\
& + (1 - e^*(0)) \cdot [y_L - b_L(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_L(\hat{w}_{ia}) + \hat{w}_{ia}) - y_L)] - \hat{w}_{ia} \\
& \geq \bar{e} \cdot [y_H - b_H(\hat{w}_s) - \alpha \cdot (2 \cdot (b_H(\hat{w}_s) + \hat{w}_s) - y_H)] \\
& + (1 - \bar{e}) \cdot [y_L - b_L(\hat{w}_s) - \alpha \cdot (2 \cdot (b_L(\hat{w}_s) + \hat{w}_s) - y_L)] - \hat{w}_s \quad (IC_{ia}) \\
& b_i(\hat{w}_{ia}) = \max \left\{ \frac{y_i}{2} - \hat{w}_{ia}, 0 \right\} \quad (BP_i) \\
& e^*(0) = \max \{ c'^{-1}((b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia})), \hat{e}_{ia}), \hat{e}_{ia} \} \quad (IC_A) \\
& \hat{w}_{ia} + e^*(0) \cdot b_H(\hat{w}_{ia}) + (1 - e^*(0)) b_L(\hat{w}_{ia}) - c(e^*(0)) \geq \bar{u} = 0 \quad (IR_A)
\end{aligned}$$

The constraints ensure that each type of principal prefers his own contract to that of the other type ( $(IC_s)$  and  $(IC_{ia})$ ).  $(BP_i)$  specifies the optimal bonus payment for the inequality-averse principal, whereas  $(IC_A)$  characterizes the agent's optimal effort. Finally,  $(IR_A)$  ensures that the agent accepts the contract offer by the inequality-averse principal. Note that we apply the relationship  $\max \{ y_j - 2 \cdot (b_j(\hat{w}) + \hat{w}), 0 \} = 0$  ( $j \in \{L, H\}$ ) to simplify the single expressions implying no compassion on the part of the inequality-averse principal.

The optimization problem can be simplified, as the following lemma shows.

**Lemma 1** *Condition  $(IR_A)$  is always satisfied.*

Lemma 1 indicates that the agent's individual rationality constraint is always met when he faces an inequality-averse principal. This is intuitive: because the principal takes the agent's wellbeing into account, the agent should never be worse off than when facing the selfish type of principal.

In the following, we analyze the constraints  $(IC_{ia})$  and  $(IC_s)$  since their simultaneous fulfillment is necessary for the existence of a separating equilibrium. This is addressed in the next two propositions.

**Proposition 2** (i)  $\frac{\partial e^*(0)}{\partial \hat{w}_{ia}} < 0$  if  $\frac{y_L}{2} < \hat{w}_{ia} < \frac{y_H}{2}$  and  $\frac{\partial e^*(0)}{\partial \hat{w}_{ia}} = 0$  if  $\hat{w}_{ia} < \frac{y_L}{2}$  or  $\hat{w}_{ia} > \frac{y_H}{2}$ .

(ii) The objective function for the inequality-averse principal is strictly decreasing in  $\hat{w}_{ia}$  if  $\hat{w}_{ia} > \frac{y_L}{2}$  and constant in  $\hat{w}_{ia}$  if  $\hat{w}_{ia} < \frac{y_L}{2}$ .

(iii) There exists a unique  $\hat{w}'_{ia} \in (\frac{y_L}{2}, \frac{y_H}{2} - c'(\bar{e}))$  such that  $(IC_{ia})$  is fulfilled for  $\hat{w}_{ia}$  if and only if  $\hat{w}_{ia} \leq \hat{w}'_{ia}$ .

(iv)  $\hat{w}_{ia} = \hat{w}'_{ia}$  is implicitly defined by

$$\left(1 - c'^{-1}\left(\frac{y_H}{2} - \hat{w}_{ia}\right)\right) \cdot \left(1 + (1 + 2 \cdot \alpha) \cdot \left(\frac{2 \cdot \hat{w}_{ia} - y_L}{y_H - y_L}\right)\right) = 1 - \bar{e}.$$

(v)  $\frac{\partial \hat{w}'_{ia}}{\partial \bar{e}} < 0$ .

Proposition 2 is rich in content. It demonstrates that the agent's optimal effort when facing a fair principal decreases with the fixed wage that he receives. The reason is as follows. The principal is willing to share output with the agent up to the point at which both receive the same payoff. If the fixed wage increases, the principal gives a lower amount of the output to the agent since this is sufficient to equalize the payoffs. In turn, the agent's reward for producing high output and his incentive for effort decrease.

As a higher fixed wage leads to higher wage costs and lowered incentives for the agent, the inequality-averse principal suffers from increasing the fixed wage. Hence, if this wage becomes too high, he would gain from imitating the selfish type and paying a lower fixed wage at the cost of being unable to induce implicit incentives. As part (v) of the proposition shows, the incentive to deviate becomes higher with a better monitoring technology. This is intuitive. If explicit incentives are more effective, it is less costly for the inequality-averse principal not to be able to use implicit incentives.

**Proposition 3** (i) *There exists a unique  $\hat{w}_{ia}'' \in (c(\bar{e}), \frac{y_H}{2} - c'(\bar{e}))$  such that  $(IC_s)$  is fulfilled for  $\hat{w}_{ia}$  if and only if  $\hat{w}_{ia} > \hat{w}_{ia}''$ .*

(ii)  $\hat{w}_{ia} = \hat{w}_{ia}''$  is implicitly defined by

$$\frac{y_H}{2} - \max \left\{ \frac{y_L}{2}, \hat{w}_{ia} \right\} = c' \left( \bar{e} + \frac{\hat{w}_{ia} - c(\bar{e})}{y_H - y_L} \right).$$

(iii)  $\frac{\partial \hat{w}_{ia}''}{\partial \bar{e}} < 0$ .

The intuition for Proposition 3 is similar to that for Proposition 2. If  $\hat{w}_{ia}$  increases, the selfish principal finds it less beneficial to imitate the inequality-averse type (because wage costs increase and the agent decreases his effort). Therefore, there is a cutoff value for  $\hat{w}_{ia}$  such that the selfish principal mimics the inequality-averse type only if the actual wage is less than the cutoff. The cutoff decreases with  $\bar{e}$  since the selfish principal has a higher incentive to reveal his type if he can rely on effective explicit incentives.

The two propositions allow us to determine the circumstances under which a separating equilibrium exists.

**Proposition 4** (i) *A separating equilibrium exists if and only if  $\hat{w}_{ia}'' < \hat{w}_{ia}'$ .*

(ii) *For all  $\bar{e} \in (0, 1]$  and for all  $y_H - y_L > 2 \cdot c'(\bar{e})$ , condition  $\hat{w}_{ia}'' < \hat{w}_{ia}'$  is fulfilled if*

$$\frac{y_L}{2} \geq \left( c'^{-1} \left( \frac{y_H - y_L}{2} \right) - \bar{e} \right) \cdot (y_H - y_L) + c(\bar{e}).$$

*In this case  $\hat{w}_{ia}''$  corresponds to the right-hand side of the latter inequality.*

The proposition indicates that under certain parameter constellations a separating equilibrium exists in which both types of principal offer different contracts to the agent so that the agent learns the principal type from the contract offer. In the separating equilibrium, explicit and implicit incentives are substitutes. Since the agent learns the principal type, a selfish principal cannot credibly promise to honor an implicit contract. Instead, he has to rely on monitoring. An inequality-averse principal, by contrast, can effectively induce implicit incentives and hence does not need to monitor the agent.<sup>9</sup>

---

<sup>9</sup>Note that in the separating equilibrium  $e^*(0) > \bar{e}$ .

Separation of the two types is feasible only under certain conditions. Different effects are at work. On one hand, the offer of a high fixed wage is less costly to the inequality-averse principal than to the selfish principal. This is because in some situations the inequality-averse principal pays a lower bonus in response to a higher fixed wage. The selfish principal never pays a bonus, so that a higher fixed wage unambiguously leads to higher wage costs. On the other hand, the inequality-averse principal gains relatively little from inducing a high level of effort because he shares the gains (i.e. high output levels) with the agent. Accordingly, the choice of a contract that induces the agent to expend greater effort is more attractive to the selfish principal. If  $y_L$  is high enough, the optimal bonus is always positive and characterized by the condition  $b_i = \frac{y_i}{2} - \hat{w}_{ia}$ . Then an increase in the fixed wage is completely costless for the inequality-averse principal since an increase in the fixed wage by one unit leads to a decrease in the bonus by one unit, leaving total wage costs unchanged. In this case the inequality-averse principal finds it always in his interest to set the fixed wage so high that the selfish principal does not want to mimic his contract offer.

Finally, we investigate how the existence of a separating equilibrium is affected by the effectiveness of explicit incentives. We tackle this question in two propositions. Since some of the effects are opposing, we impose more structure in the second proposition by assuming a specific effort cost function.

**Proposition 5** *There exists  $y^* \in (2 \cdot y_L, \infty) \cup \{\infty\}$  such that for all  $y_H < y^*$  we can find  $e^* \in (0, c'^{-1}(\min\{\frac{y_L}{2}, \frac{y_H - y_L}{2}\}))$ , implying the existence of a separating equilibrium for all  $\bar{e} > e^*$ .*

**Proposition 6** *Let  $c$  be a quadratic cost function with constant second derivative  $c''$ , and let  $(1 + \frac{1}{2\alpha}) \cdot c'' \cdot (\frac{1+\bar{e}}{2}) < \frac{y_H - y_L}{2}$ . Then we have  $\frac{\partial(\hat{w}'_{ia} - \hat{w}''_{ia})}{\partial \bar{e}} > 0$ .*

The two propositions have similar implications. Proposition 5 indicates that a separating equilibrium exists if  $\bar{e}$  becomes sufficiently high (and  $y_H$  is not too large). Proposition 6 implies that the separating equilibrium may become more likely with increasing  $\bar{e}$ . Intuitively, if a selfish principal can rely on more effective explicit incentives, it becomes easier to separate

an inequality-averse principal from the selfish type. This means that the possibility of offering explicit incentives for one type of principal may allow the other type of principal to induce more effective implicit incentives. In this sense, there is a strategic complementarity between explicit and implicit incentives.

## 5 Conclusion

We analyzed the interaction of explicit and implicit contracts in a model with selfish and fair principals. We investigated a separating equilibrium, in which principals reveal their type through a contract offer to an agent. We found that in this type of equilibrium, explicit and implicit contracts are substitutes. Since the agent learns the principal's type, a selfish principal has to rely on explicit incentives. A fair principal, by contrast, can effectively induce implicit incentives and hence does not need to use explicit incentives. Interestingly, if a selfish principal can rely on more effective explicit incentives, it becomes more likely that a fair principal can be separated from the selfish type and hence can make better use of implicit incentives. In this sense there is a strategic complementarity between explicit and implicit incentives.

Although we have confined attention to analysis of a separating equilibrium, a pooling equilibrium in which both types of principal offer the same contract typically exists as well. The proof is similar to the argument in the first part of the Appendix. Since Bayes' rule does not restrict beliefs off the equilibrium path, beliefs could be structured such that adherence to the pooling equilibrium is optimal. In our view, investigation of the separating equilibrium is more interesting. It involves transmission of information and enables us to understand in which way an inequality-averse principal can reveal his type and enter a more effective implicit contract with an agent.

# Appendix

## Beliefs off the equilibrium path

The conditions in Section 4 guaranteed that each type of principal prefers the own contract to the contract offered by the other type. Of course, the principal must prefer the own contract to any other feasible contract as well. Since contracts other than  $(\hat{w}_s, \hat{e}_s)$  and  $(\hat{w}_{ia}, \hat{e}_{ia})$  are not offered in equilibrium, Bayes' rule puts no restriction on the agent's belief about the principal's type in case he would observe such a different contract offer. In this section, we demonstrate that there exist off-equilibrium beliefs such that each type of principal prefers the own contract to any other feasible contract so that the separating equilibrium characterized before indeed exists.

To demonstrate this, assume that the agent believes the principal to be selfish for sure if a contract different from  $(\hat{w}_s, \hat{e}_s)$  or  $(\hat{w}_{ia}, \hat{e}_{ia})$  is offered. Obviously, the selfish type of principal then does not want to deviate from  $(\hat{w}_s, \hat{e}_s)$ . A deviation to a contract other than  $(\hat{w}_{ia}, \hat{e}_{ia})$  would not change the belief about his type and, given this belief,  $(\hat{w}_s, \hat{e}_s)$  is optimal for him.

It remains to show that the inequality-averse type of principal does not want to deviate from his contract offer either. In the case of a deviation to  $(\hat{w}, \hat{e})$ , the agent would believe that he is dealing with a selfish principal. Hence, he would choose effort equal to  $\hat{e}$ . Moreover, his individual rationality constraint would require  $\hat{w} \geq c(\hat{e})$ . If  $\hat{e} = 0$ , the principal's expected utility is bounded from above by

$$y_L - b_L(0) - \alpha \cdot (2 \cdot (b_L(0)) - y_L) = \frac{y_L}{2}$$

Similarly, if  $\hat{e} = \bar{e}$ , expected utility is bounded from above by

$$\begin{aligned} & \bar{e} \cdot [y_H - b_H(c(\bar{e})) - \alpha \cdot (2 \cdot (b_H(c(\bar{e})) + c(\bar{e})) - y_H)] \\ & + (1 - \bar{e}) \cdot [y_L - b_L(c(\bar{e})) - \alpha \cdot (2 \cdot (b_L(c(\bar{e})) + c(\bar{e})) - y_L)] - c(\bar{e}) \\ & = \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2} \end{aligned}$$

From (A8) (see section "Proofs of Lemmas and Propositions" in the Appendix) we observe that both these utilities are lower than the principal's equilibrium utility. This implies that



a deviation to any contract offer would always make the inequality-averse type of principal worse off.

## Proofs of lemmas and propositions

### Proof of Proposition 1.

In a separating equilibrium, the agent learns the principal's type. Accordingly, he infers that he does not receive any bonus when facing a selfish principal. Thus, in this case his effort is equal to  $\hat{e}_s$ . The individual rationality constraint requires  $\hat{w}_s \geq c(\hat{e}_s)$ . In the optimum, of course, the selfish principal sets  $\hat{w}_s = c(\hat{e}_s)$ . His expected profit is therefore given by  $\max \{y_L - c(0), y_L + \bar{e} \cdot (y_H - y_L) - c(\bar{e})\}$ . Since  $c(0) = 0$ ,  $c'(\bar{e}) \leq \frac{y_H - y_L}{2}$ ,  $c'' > 0$ , it follows

$$c(e) < e \cdot \frac{y_H - y_L}{2} < e \cdot (y_H - y_L) \quad \text{for all } e \in (0, \bar{e}]. \quad (\text{A1})$$

This in turn implies monitoring to be optimal for the selfish principal and thus  $\hat{e}_s = \bar{e}$ . Since, in addition,  $c'(\bar{e}) \leq \frac{y_L}{2}$  we also get

$$c(e) < e \cdot \frac{y_L}{2} \leq \frac{y_L}{2} \quad \text{for all } e \in (0, \bar{e}] \quad (\text{A2})$$

which completes the proof. ■

### Proof of Lemma 1.

Using  $b_i(\hat{w}) = \max \left\{ \frac{y_i}{2} - \hat{w}, 0 \right\}$ ,  $(IR_A)$  can be rewritten as

$$e^*(0) \cdot \max \left\{ \frac{y_H}{2}, \hat{w}_{ia} \right\} + (1 - e^*(0)) \cdot \max \left\{ \frac{y_L}{2}, \hat{w}_{ia} \right\} - c(e^*(0)) \geq 0.$$

The optimality of  $e^*(0)$  implies that the agent's expected payoff from choosing  $e^*(0)$  is not lower than the expected payoff from choosing  $\bar{e}$ . Hence, we have

$$\begin{aligned} & e^*(0) \cdot \max \left\{ \frac{y_H}{2}, \hat{w}_{ia} \right\} + (1 - e^*(0)) \cdot \max \left\{ \frac{y_L}{2}, \hat{w}_{ia} \right\} - c(e^*(0)) \\ & \geq \bar{e} \cdot \max \left\{ \frac{y_H}{2}, \hat{w}_{ia} \right\} + (1 - \bar{e}) \cdot \max \left\{ \frac{y_L}{2}, \hat{w}_{ia} \right\} - c(\bar{e}) \\ & \geq \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2} - c(\bar{e}). \end{aligned}$$

The latter term is obviously positive. ■

**Proof of Proposition 2.**

(i)

Under consideration of

$$\begin{aligned}
e^*(0) &= \max\{c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia})), \hat{e}_{ia}\} \\
&= (c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia})) - \hat{e}_{ia}) \cdot \chi_{(\hat{e}_{ia}, \infty)}(c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia}))) + \hat{e}_{ia}
\end{aligned}$$

we get for all  $\hat{w}_{ia}$  with  $c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia})) \neq \hat{e}_{ia}$ :

$$\begin{aligned}
\frac{\partial e^*(0)}{\partial \hat{w}_{ia}} &= \frac{1}{c''(c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia})))} \cdot \left( \frac{\partial b_H}{\partial \hat{w}_{ia}} - \frac{\partial b_L}{\partial \hat{w}_{ia}} \right) \\
&\quad \cdot \chi_{(\hat{e}_{ia}, \infty)}(c'^{-1}(b_H(\hat{w}_{ia}) - b_L(\hat{w}_{ia}))).
\end{aligned} \tag{A3}$$

Since

$$\frac{\partial b_H}{\partial \hat{w}_{ia}} - \frac{\partial b_L}{\partial \hat{w}_{ia}} = -\chi_{(-\infty, \frac{y_H}{2})}(\hat{w}_{ia}) + \chi_{(-\infty, \frac{y_L}{2})}(\hat{w}_{ia}) \tag{A4}$$

and  $c'' > 0$ , part (i) has been proven.

(ii)

The objective function of the inequality averse principal can be transformed as follows:

$$\begin{aligned}
\Phi_{ia}(\hat{w}_{ia}, \hat{e}_{ia}) &= e^*(0) \cdot [y_H - b_H(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_H(\hat{w}_{ia}) + \hat{w}_{ia}) - y_H)] \\
&\quad + (1 - e^*(0)) \cdot [y_L - b_L(\hat{w}_{ia}) - \alpha \cdot (2 \cdot (b_L(\hat{w}_{ia}) + \hat{w}_{ia}) - y_L)] - \hat{w}_{ia} \\
&= e^*(0) \cdot \left[ y_H - \max\left\{\frac{y_H}{2}, \hat{w}_{ia}\right\} + \alpha \cdot \left( y_H - 2 \cdot \max\left\{\frac{y_H}{2}, \hat{w}_{ia}\right\} \right) \right] \\
&\quad + (1 - e^*(0)) \cdot \left[ y_L - \max\left\{\frac{y_L}{2}, \hat{w}_{ia}\right\} + \alpha \cdot \left( y_L - 2 \cdot \max\left\{\frac{y_L}{2}, \hat{w}_{ia}\right\} \right) \right] \\
&= e^*(0) \cdot \left[ \frac{y_H}{2} - (1 + 2 \cdot \alpha) \cdot \max\left\{\hat{w}_{ia} - \frac{y_H}{2}, 0\right\} \right] \\
&\quad + (1 - e^*(0)) \cdot \left[ \frac{y_L}{2} - (1 + 2 \cdot \alpha) \cdot \max\left\{\hat{w}_{ia} - \frac{y_L}{2}, 0\right\} \right]
\end{aligned} \tag{A5}$$

Consequently, the derivative of  $\Phi_{ia}$  with respect to  $\hat{w}_{ia}$  (and  $\hat{w}_{ia} \notin \{\frac{y_L}{2}, \frac{y_H}{2}\}$ ) corresponds to

$$\begin{aligned}
& \frac{\partial \Phi_{ia}}{\partial \hat{w}_{ia}}(\hat{w}_{ia}, \hat{e}_{ia}) \\
&= \frac{\partial e^*(0)}{\partial \hat{w}_{ia}} \cdot \left[ \frac{y_H}{2} - (1 + 2 \cdot \alpha) \cdot \max \left\{ \hat{w}_{ia} - \frac{y_H}{2}, 0 \right\} \right] \\
&\quad - e^*(0) \cdot (1 + 2 \cdot \alpha) \cdot \chi_{(\frac{y_H}{2}, \infty)}(\hat{w}_{ia}) \\
&\quad - \frac{\partial e^*(0)}{\partial \hat{w}_{ia}} \cdot \left[ \frac{y_L}{2} - (1 + 2 \cdot \alpha) \cdot \max \left\{ \hat{w}_{ia} - \frac{y_L}{2}, 0 \right\} \right] \\
&\quad - (1 - e^*(0)) \cdot (1 + 2 \cdot \alpha) \cdot \chi_{(\frac{y_L}{2}, \infty)}(\hat{w}_{ia}) \\
&= -(1 + 2 \cdot \alpha) \cdot \left[ e^*(0) \cdot \chi_{(\frac{y_H}{2}, \infty)}(\hat{w}_{ia}) + (1 - e^*(0)) \cdot \chi_{(\frac{y_L}{2}, \infty)}(\hat{w}_{ia}) \right] \\
&\quad + \frac{\partial e^*(0)}{\partial \hat{w}_{ia}} \cdot \left[ \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left[ \max \left\{ \hat{w}_{ia} - \frac{y_L}{2}, 0 \right\} - \max \left\{ \hat{w}_{ia} - \frac{y_H}{2}, 0 \right\} \right] \right]. \tag{A6}
\end{aligned}$$

Since the first summand of the latter term obviously is negative if  $\hat{w}_{ia} > \frac{y_L}{2}$  and zero otherwise and the factor in the square brackets of the second summand is positive, part (ii) follows immediately from part (i).

(iii)

From Proposition 1 we know that  $b_i(\hat{w}_s) = \frac{y_i}{2} - \hat{w}_s$  and the right-hand side of condition  $(IC_{ia})$  can be transformed as follows:

$$\begin{aligned}
& \bar{e} \cdot [y_H - b_H(\hat{w}_s) - \alpha \cdot (2 \cdot (b_H(\hat{w}_s) + \hat{w}_s) - y_H)] \\
&\quad + (1 - \bar{e}) \cdot [y_L - b_L(\hat{w}_s) - \alpha \cdot (2 \cdot (b_L(\hat{w}_s) + \hat{w}_s) - y_L)] - \hat{w}_s \\
&= \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2}. \tag{A7}
\end{aligned}$$

Thus, under consideration of (A5) and the fact that the left-hand side of  $(IC_{ia})$  corresponds to the objective function  $\Phi_{ia}$ ,  $(IC_{ia})$  is equivalent to

$$\begin{aligned}
& e^*(0) \cdot \left[ \frac{y_H}{2} - (1 + 2 \cdot \alpha) \cdot \max \left\{ \hat{w}_{ia} - \frac{y_H}{2}, 0 \right\} \right] \\
&\quad + (1 - e^*(0)) \cdot \left[ \frac{y_L}{2} - (1 + 2 \cdot \alpha) \cdot \max \left\{ \hat{w}_{ia} - \frac{y_L}{2}, 0 \right\} \right] \\
&\geq \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2}. \tag{A8}
\end{aligned}$$

If we apply  $\hat{w}_{ia} = \frac{y_L}{2}$ , the latter inequality is equivalent to

$$e^*(0) \cdot \frac{y_H}{2} + (1 - e^*(0)) \cdot \frac{y_L}{2} \geq \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2}$$

which is fulfilled since

$$e^*(0) = \max \left\{ c'^{-1} \left( \frac{y_H - y_L}{2} \right), \hat{e} \right\} \stackrel{c'(\bar{e}) \leq \frac{y_H - y_L}{2}}{=} c'^{-1} \left( \frac{y_H - y_L}{2} \right) \geq \bar{e}.$$

On the other hand, if we apply  $\hat{w}_{ia} = \frac{y_H}{2} - c'(\bar{e})$ , we get  $e^*(0) = \bar{e}$  and the left-hand side of (A8) corresponds to

$$\bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \left( \frac{y_L}{2} - (1 + 2 \cdot \alpha) \cdot \frac{y_H - y_L - 2 \cdot c'(\bar{e})}{2} \right) < \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2}.$$

Consequently, condition  $(IC_{ia})$  is not fulfilled for  $\hat{w}_{ia} = \frac{y_H}{2} - c'(\bar{e})$ . Finally, since the left-hand side of (A8) is strictly decreasing in  $\hat{w}_{ia}$  as long as  $\hat{w}_{ia} \geq \frac{y_L}{2}$  and constant otherwise and the right-hand side of (A8) is independent of  $\hat{w}_{ia}$ , part (iii) has been shown.

(iv)

According to part (iii) and (A8) the critical parameter  $\hat{w}_{ia} = \hat{w}'_{ia} \in \left( \frac{y_L}{2}, \frac{y_H}{2} - c'(\bar{e}) \right)$  is implicitly defined by

$$\begin{aligned} e^*(0) \cdot \frac{y_H}{2} + (1 - e^*(0)) \cdot \left[ \frac{y_L}{2} - (1 + 2 \cdot \alpha) \cdot \left( \hat{w}_{ia} - \frac{y_L}{2} \right) \right] \\ &= \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2} \\ \Leftrightarrow e^*(0) \cdot \left( \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}_{ia} - \frac{y_L}{2} \right) \right) \\ &= \bar{e} \cdot \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}_{ia} - \frac{y_L}{2} \right) \tag{A9} \\ \Leftrightarrow 1 - e^*(0) &= (1 - \bar{e}) \cdot \frac{\frac{y_H - y_L}{2}}{\frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}_{ia} - \frac{y_L}{2} \right)} \end{aligned}$$

Since  $\hat{w}_{ia} < \frac{y_H}{2} - c'(\bar{e})$ , it follows  $e^*(0) = c'^{-1} \left( \frac{y_H}{2} - \hat{w}_{ia} \right)$  which in turn implies the following identity to be fulfilled by  $\hat{w}_{ia} = \hat{w}'_{ia}$ :

$$\begin{aligned} 1 - c'^{-1} \left( \frac{y_H}{2} - \hat{w}_{ia} \right) &= (1 - \bar{e}) \cdot \frac{\frac{y_H - y_L}{2}}{\frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}_{ia} - \frac{y_L}{2} \right)} \\ \Leftrightarrow \left( 1 - c'^{-1} \left( \frac{y_H}{2} - \hat{w}_{ia} \right) \right) \cdot \left( 1 + (1 + 2 \cdot \alpha) \cdot \left( \frac{2 \cdot \hat{w}_{ia} - y_L}{y_H - y_L} \right) \right) &= 1 - \bar{e}. \end{aligned}$$

(v)

According to (A9),  $\hat{w}'_{ia}$  is implicitly defined by

$$\begin{aligned} f_1(\hat{w}'_{ia}, \bar{e}) := & e^*(0) \cdot \left( \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}'_{ia} - \frac{y_L}{2} \right) \right) \\ & - \bar{e} \cdot \frac{y_H - y_L}{2} - (1 + 2 \cdot \alpha) \cdot \left( \hat{w}'_{ia} - \frac{y_L}{2} \right) = 0. \end{aligned}$$

To determine the monotonic behavior of  $\hat{w}'_{ia}$ , we apply the Implicit function theorem and need the following partial derivatives:

$$\begin{aligned} \frac{\partial f_1}{\partial \hat{w}'_{ia}}(\hat{w}'_{ia}, \bar{e}) &= \frac{\partial e^*(0)}{\partial \hat{w}'_{ia}} \cdot \left( \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}'_{ia} - \frac{y_L}{2} \right) \right) \\ &\quad + (e^*(0) - 1) \cdot (1 + 2 \cdot \alpha) < 0, \\ \frac{\partial f_1}{\partial \bar{e}}(\hat{w}'_{ia}, \bar{e}) &= -\frac{y_H - y_L}{2} < 0. \end{aligned}$$

The Implicit function theorem then gives

$$\frac{\partial \hat{w}'_{ia}}{\partial \bar{e}} = -\frac{\frac{\partial f_1}{\partial \bar{e}}}{\frac{\partial f_1}{\partial \hat{w}'_{ia}}} < 0.$$

■

### Proof of Proposition 3.

(i)

If we set  $\hat{w}_{ia} = \frac{y_H}{2} - c'(\bar{e})$  we get  $e^*(0) = \bar{e}$  and condition  $(IC_s)$  is equivalent to

$$0 \leq \frac{y_H}{2} - c'(\bar{e}) - c(\bar{e}). \quad (\text{A10})$$

This inequality is fulfilled since  $c'(\bar{e}) \leq \frac{y_L}{2}$  and  $c(\bar{e}) \leq \frac{y_H - y_L}{2}$ . However, if  $\hat{w}_{ia} = c(\bar{e})$  condition  $(IC_s)$  is equivalent to

$$\left( c'^{-1} \left( \frac{y_H - y_L}{2} \right) - \bar{e} \right) \leq 0 \quad (\text{A11})$$

which is obviously not true. Thus, part (i) has been proven since the right-hand side of  $(IC_s)$  is strictly increasing in  $\hat{w}_{ia}$  and (according to Proposition 2 (i)) the left-hand side of  $(IC_s)$  is (weakly) decreasing in  $\hat{w}_{ia}$ .

(ii)

The result immediately follows from part (i) and condition  $(IC_s)$ .

(iii)

According to  $(IC_s)$ ,  $\hat{w}_{ia}''$  is implicitly defined by

$$f_2(\hat{w}_{ia}'', \bar{e}) := (e^*(0) - \bar{e}) \cdot (y_H - y_L) - \hat{w}_{ia}'' + c(\bar{e}) = 0.$$

Again, we apply the Implicit function theorem. We need the following partial derivatives:

$$\begin{aligned} \frac{\partial f_2}{\partial \hat{w}_{ia}''}(\hat{w}_{ia}'', \bar{e}) &= \frac{\partial e^*(0)}{\partial \hat{w}_{ia}''} \cdot (y_H - y_L) - 1 < 0, \\ \frac{\partial f_2}{\partial \bar{e}}(\hat{w}_{ia}'', \bar{e}) &= -(y_H - y_L) + c'(\bar{e}) < 0. \end{aligned}$$

The Implicit function theorem gives

$$\frac{\partial \hat{w}_{ia}''}{\partial \bar{e}} = -\frac{\frac{\partial f_2}{\partial \bar{e}}}{\frac{\partial f_2}{\partial \hat{w}_{ia}''}} < 0.$$

■

#### Proof of Proposition 4.

The statement of the first part of the proposition is obvious. The second part follows since in the case  $\hat{w}_{ia} = \frac{y_L}{2}$  condition  $(IC_s)$  is equivalent to

$$\frac{y_L}{2} \geq \left( c'^{-1} \left( \frac{y_H - y_L}{2} \right) - \bar{e} \right) \cdot (y_H - y_L) + c(\bar{e})$$

and condition  $(IC_{ia})$  is fulfilled as a result of Proposition 2 (iii). It remains to show that the right-hand side of the latter inequality corresponds to  $\hat{w}_{ia}''$ . Since this statement is equivalent to the equation of Proposition 3 (ii) in the case  $\hat{w}_{ia} \leq \frac{y_L}{2}$  the proof is complete. ■

#### Proof of Proposition 5.

First, suppose  $y^* = 2 \cdot y_L$ , i.e.  $\frac{y_L}{2} \geq \frac{y_H - y_L}{2}$  for all  $y_H \leq y^*$ . In this case it is possible to consider  $\bar{e}$  approaching  $c'^{-1} \left( \frac{y_H - y_L}{2} \right)$  which leads to the following limit of the right-hand side of the sufficient condition from Proposition 4 (ii):

$$\lim_{\bar{e} \rightarrow c'^{-1} \left( \frac{y_H - y_L}{2} \right)} \left( \left( c'^{-1} \left( \frac{y_H - y_L}{2} \right) - \bar{e} \right) \cdot (y_H - y_L) + c(\bar{e}) \right) = c \left( c'^{-1} \left( \frac{y_H - y_L}{2} \right) \right) < \frac{y_L}{2}. \quad (\text{A12})$$

Note that the latter inequality results from (A2). Since the right-hand side of the sufficient condition from Proposition 4 (ii) is continuous in  $y_H$ , there exists  $\delta > 0$  such that inequality (A12) also holds true for all  $y_H$  with  $y_H \leq y^* = 2 \cdot y_L + \delta$  which completes the proof. ■

### Proof of Proposition 6.

As a result of part (v) of Proposition 2 and part (iii) of Proposition 3 we have to show that  $\left| \frac{\partial \hat{w}'_{ia}}{\partial \bar{e}} \right| > \left| \frac{\partial \hat{w}'_{ia}}{\partial \bar{e}} \right|$ . If we consider the functions  $f_1$  and  $f_2$  of the proofs of Propositions 2 and 3, this condition is fulfilled if

$$\left| \frac{\partial f_2}{\partial \bar{e}} \right| > \left| \frac{\partial f_1}{\partial \bar{e}} \right| \quad \text{and} \quad \left| \frac{\partial f_2}{\partial \hat{w}''_{ia}} \right| < \left| \frac{\partial f_1}{\partial \hat{w}'_{ia}} \right|.$$

The first inequality is equivalent to

$$y_H - y_L - c'(\bar{e}) > \frac{y_H - y_L}{2} \quad \Leftrightarrow \quad c'(\bar{e}) < \frac{y_H - y_L}{2}$$

and consequently true. Let  $e_1^*(0)$  and  $e_2^*(0)$  characterize  $e^*(0)$  if  $e^*(0)$  is calculated on the basis of  $\hat{w}_{ia} = \hat{w}'_{ia}$  and  $\hat{w}_{ia} = \hat{w}''_{ia}$ , respectively. Then the second inequality is equivalent to

$$\begin{aligned} & \left| \frac{\partial e_2^*(0)}{\partial \hat{w}''_{ia}} \right| \cdot (y_H - y_L) + 1 \\ & < \left| \frac{\partial e_1^*(0)}{\partial \hat{w}'_{ia}} \right| \cdot \left( \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}'_{ia} - \frac{y_L}{2} \right) \right) + (1 - e_1^*(0)) \cdot (1 + 2 \cdot \alpha). \end{aligned}$$

Notice that  $\hat{w}'_{ia} \in \left( \frac{y_L}{2}, \frac{y_H}{2} \right)$ . Under consideration of (A3), (A4), and the assumption of a quadratic cost function, this means that  $\left| \frac{\partial e_2^*(0)}{\partial \hat{w}''_{ia}} \right| \leq \left| \frac{\partial e_1^*(0)}{\partial \hat{w}'_{ia}} \right|$ . As a result, the above inequality is fulfilled if

$$(1 - e_1^*(0)) \cdot (1 + 2 \cdot \alpha) > 1 \quad \Leftrightarrow \quad e_1^*(0) < 1 - \frac{1}{1 + 2 \cdot \alpha} \quad \text{and} \quad (\text{A13})$$

$$\begin{aligned} & \left( \frac{y_H - y_L}{2} + (1 + 2 \cdot \alpha) \cdot \left( \hat{w}'_{ia} - \frac{y_L}{2} \right) \right) > y_H - y_L \\ \Leftrightarrow & \hat{w}'_{ia} > \frac{1}{1 + 2 \cdot \alpha} \cdot \frac{y_H - y_L}{2} + \frac{y_L}{2}. \end{aligned} \quad (\text{A14})$$

Inequality (A14) is valid since  $\hat{w}_{ia} = \frac{1}{1+2\alpha} \cdot \frac{y_H - y_L}{2} + \frac{y_L}{2}$  fulfills  $(IC_{ia})$  which can be seen by inserting  $\hat{w}_{ia}$  into (A8) (note that  $\frac{y_H}{2} > \frac{1}{1+2\alpha} \cdot \frac{y_H - y_L}{2} + \frac{y_L}{2}$ ):

$$\begin{aligned} & e^*(0) \cdot \frac{y_H}{2} + (1 - e^*(0)) \cdot \left[ y_L - \frac{y_H}{2} \right] \geq \bar{e} \cdot \frac{y_H}{2} + (1 - \bar{e}) \cdot \frac{y_L}{2} \\ \Leftrightarrow & e^*(0) \cdot (y_H - y_L) > (1 + \bar{e}) \cdot \frac{y_H - y_L}{2} \\ \Leftrightarrow & e^*(0) > \frac{1 + \bar{e}}{2}. \end{aligned}$$



Under consideration of  $e^*(0) = c'^{-1}\left(\frac{y_H - y_L}{2} \cdot \left(1 - \frac{1}{1+2\alpha}\right)\right)$  the latter inequality is equivalent to (since  $c'(0) = 0$ )

$$\frac{y_H - y_L}{2} > c' \left( \frac{1 + \bar{e}}{2} \right) \cdot \left( 1 + \frac{1}{2 \cdot \alpha} \right) = c'' \cdot \left( \frac{1 + \bar{e}}{2} \right) \cdot \left( 1 + \frac{1}{2 \cdot \alpha} \right),$$

which is result of the assumption.

Finally, we have to verify inequality (A13). Again, we make use of (A14) according to which  $e_1^*(0) < c'^{-1}\left(\frac{y_H - y_L}{2} \cdot \left(1 - \frac{1}{1+2\alpha}\right)\right)$ . Consequently, it is sufficient to show that

$$\begin{aligned} c'^{-1} \left( \frac{y_H - y_L}{2} \cdot \left( 1 - \frac{1}{1+2\alpha} \right) \right) &< 1 - \frac{1}{1+2\alpha} \\ \Leftrightarrow \frac{y_H - y_L}{2} &< \frac{c' \left( 1 - \frac{1}{1+2\alpha} \right)}{1 - \frac{1}{1+2\alpha}} = c'' = c'(1). \end{aligned}$$

This statement immediately follows from the initially assumed weakened Inada conditions.

■

## References

- Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts," 109 *Quarterly Journal of Economics* 1125-1156.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 1999. "Informal Authority in Organizations," 15 *Journal of Law, Economics, & Organization* 56-73.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm," 117 *Quarterly Journal of Economics* 39-84.
- Beaudry, Paul. 1994. "Why an Informed Principal may Leave Rents to an Agent," 35 *International Economic Review* 821-832.
- Bulow, Jeremy, John D. Geanakoplos, and Paul Klemperer. 1985. "Multimarket Oligopoly: Strategic Substitutes and Strategic Complements," 93 *Journal of Political Economy* 488-511.
- Che, Yeon-Koo, and Seung-Weon Yoo. 2001. "Optimal Incentives for Teams," 91 *American Economic Review* 525-541.
- Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999a. "The Economics of Career Concerns, Part I: Comparing Information Structures," 66 *Review of Economic Studies* 183-198.
- Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999a. "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies," 66 *Review of Economic Studies* 199-207.
- Englmaier, Florian, and Achim Wambach. 2010. "Optimal Incentive Contracts under Inequity Aversion," 69 *Games and Economic Behavior* 312-328.
- Falk, Armin, and Michael Kosfeld. 2006. "The Hidden Costs of Control," 96 *American Economic Review* 1611-1630.

- Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." 114 *Quarterly Journal of Economics* 817-868.
- Giebe, Thomas, and Oliver Gürtler. 2012. "Optimal Contracts for Lenient Supervisors." 81 *Journal of Economic Behavior & Organization* 403-420.
- Grund, Christian, and Dirk Sliwka. 2005. "Envy and Compassion in Tournaments," 14 *Journal of Economics and Management Strategy* 187-207.
- Halonen, Maija. 2002. "Reputation and the Allocation of Ownership," 112 *The Economic Journal* 539-558.
- Holmström, Bengt. 1982. "Managerial Incentive Problems – A Dynamic Perspective," *Essays in Economics and Management in Honor of Lars Wahlbeck*, Helsinki: Swedish School of Economics. (Reprinted in January 1999 in 66 *Review of Economic Studies* 169-82).
- Irlenbusch, Bernd, and Dirk Sliwka. 2006. "Career Concerns in a Simple Experimental Labour Market," *European Economic Review* 50, 147-170.
- James Jr., Harvey S. 2002. "The Trust Paradox: a Survey of Economic Inquiries into the Nature of Trust and Trustworthiness," 47 *Journal of Economic Behavior and Organization* 291-307.
- Jost, Peter-Jürgen. 1996. "On the Role of Commitment in a Principal-Agent Relationship with an Informed Principal," 68 *Journal of Economic Theory* 510-530.
- Levin, Jonathan. 2002. "Multilateral Contracting and the Employment Relationship," 117 *Quarterly Journal of Economics* 1075-1103.
- Levin, Jonathan. 2003. "Relational Incentive Contracts," 93 *American Economic Review* 835-847.
- MacLeod, W. Bentley, and James M. Malcomson. 1989. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," 57 *Econometrica* 447-80.

- Maskin, Eric, and Jean Tirole. 1990. "The Principal-Agent Relationship with an Informed Principal: the Case of Private Values," 58 *Econometrica* 379-409.
- Maskin, Eric, and Jean Tirole. 1992. "The Principal-Agent Relationship with an Informed Principal, II: Common Values," 60 *Econometrica* 1-42.
- Mohnen, Alwine, Kathrin Pokorny, and Dirk Sliwka. 2008. "Transparency, Inequity Aversion, and the Dynamics of Peer Pressure in Teams: Theory and Evidence," 26 *Journal of Labor Economics* 693-720.
- Neilson, William S., and Jill Stowe. 2010. "Piece-Rate Contracts for Other-Regarding Workers," 48 *Economic Inquiry* 575-586.
- Schmidt, Klaus M., and Monika Schnitzer. 1995. "The Interaction of Explicit and Implicit Contracts," 48 *Economics Letters* 193-199.
- Schöttner, Anja. 2008. "Relational Contracts, Multitasking, and Job Design," 24 *Journal of Law, Economics, & Organization* 138-162.
- Sliwka, Dirk. 2007. "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes," 97 *American Economic Review* 999-1012.
- Spier, Kathryn E. 1992. "Incomplete Contracts and Signalling," 23 *RAND Journal of Economics* 432-443.
- von Siemens, Ferdinand A. 2011. "Heterogeneous Social Preferences, Screening, and Employment Contracts," 63 *Oxford Economic Papers* 499-522.
- von Siemens, Ferdinand A. 2012. "Social Preferences, Sorting, and Competition," *Scandinavian Journal of Economics* forthcoming.