The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

Delegation and Interim Performance Evaluation

Miriam Schütte
Philipp C. Wichardt

CESIFO WORKING PAPER NO. 4193
CATEGORY 4: LABOUR MARKETS
APRIL 2013

Presented at CESifo Area Conference on Applied Microeconomics, March 2013

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the RePEc website: www.RePEc.org*
- *from the CESifo website: www.CESifo-group.org/wp*

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 4193*

Delegation and Interim Performance Evaluation

Abstract

This paper considers the effects of an interim performance evaluation on the decision of a principal to delegate authority to a potentially biased but better informed agent. Assuming the agents' outside option to be determined by market beliefs about their type, interim evaluations (a) provide a possibility for the principal to potentially separate biased agents from unbiased agents and (b) induce an incentive for biased agents to imitate unbiased ones in order to retain the decision authority and to increase their wages in later periods (in case of public evaluation). We show that the principal always profits from a private evaluation while a public evaluation is only beneficial if the corresponding wage effects are not too costly. Nevertheless, the principal prefers public over private evaluation if the imitation incentive for the biased type is high enough. Finally, regarding implications for economic policy, we show that in view of aggregate welfare any evaluation conducted ought to be disclosed to the public.

JEL-Code: C720, C730, D820, D860, L220, M540.

Keywords: delegation, signaling, reputation.

Miriam Schütte
Munich Graduate School of Economics
LMU Munich
Kaulbachstraße 45
Germany – 80539 Munich
miriam.schuette@lrz.uni-muenchen.de

Philipp C. Wichardt
Department of Economics
University of Rostock & Bonn
Germany
philipp.wichardt@uni-rostock.de

This Version: March 22, 2013
We are grateful to Florian Englmaier, Matthias Fahn, Fabian Herweg, Ehud Kalai, Sandra Ludwig, Niko Matouschek, Andrea Prat, Alvaro Sandroni, Klaus Schmidt and Peter Norman Sørensen as well as to the participants of the Munich Micro Workshop, the Management & Strategy Brown Bag Seminar at the Kellogg School of Management, the Spring Meeting of Young Economists 2012, the North American Summer Meeting of the Econometric Society 2012, the 2012 World Congress of the Game Theory Society, the European Meeting of the Econometric Society 2012 and the CESifo Conference on Applied Microeconomics 2013 for helpful comments and discussions. Moreover, Schütte thanks the Kellogg School of Management for its hospitality. Financial support of the German Research Foundation (DFG), GRK 801, is gratefully acknowledged.

1 Introduction

A central question in the design of organisations, which has been much discussed in the literature over the last decades, is how to optimally allocate decision rights to subordinate agents (e.g. Holmström, 1977, 1984, Aghion and Tirole, 1997, Alonso and Matouschek, 2008). The general problem is that while agents often may have better information about the profitability of certain projects – or at least be able to obtain such information – this does not necessarily imply that they will always opt for the projects which are most preferred by the principal.[1] A possible reason for this is that agents may be biased and therefore disagree with the principal on what project to choose.[2]

In order to illustrate the problem, consider a situation where a principal, e.g. a firm's headquarter, has to implement a certain project. Moreover, assume that the success of the project depends on its fit with some noisy variable, and that the realisation of this variable is not observable to the principal but to some subordinate agent, e.g. a division manager.[3] If the preferences of the principal and the agent are perfectly aligned, the problem can obviously be solved simply by either delegating the decision to the agent or asking him about his information. However, if there is a possibility that the agent's preferences differ from those of the principal, things change and the principal may rather prefer to take an uninformed decision himself.[4]

A possible way to improve matters for the principal in such a situation, which we focus on in the present paper, is to introduce performance evaluations. In practice, such evaluations typically serve multiple goals, e.g. providing the agent with incentives (e.g. MacLeod and Malcolmson, 1989, or Baker et al., 1994) or with performance feedback (see Zabojnik, 2011). What is more relevant in the present setting, however, is that performance appraisals are also used to determine assignments (cf. Cleveland et al., 1989), thereby taking an authority allocating function.

In fact, analysing a two period interaction with asymmetric information about the type of the agent, loyal or biased, we show that if the project can be subdivided into different stages, evaluating the agent on the way is always beneficial for the principal.

[1] For a discussion of how the expected profit from delegation decreases due to too frequent choices of suboptimal projects see, for example, Jensen and Meckling (1992).

[2] Alternatively, agents may differ in their ability to interpret incoming signals about the state of the world so that delegation will become very costly if the probability of unable agents letting opportunities pass, or choosing the wrong projects, is too high (e.g. Levy, 2005).

[3] Here one might think of certain aspects of market demand, the quality of some input, etc.

[4] There is, for example, anecdotal evidence that many managers engage in empire building, i.e. they invest more than the optimal investment from the perspective of their principals (e.g. Jensen, 1986).

First of all, the evaluation naturally grants the manager a possibility to learn the type of the agent and to condition at least the later delegation decision on this information. And, what is more, if the agent's wage is tied to the market's belief about his type, the principal may even want to make the result of the evaluation public. The reason for this is that disclosure of the evaluation creates an incentive for the biased agent to align his behaviour with that of the unbiased one in period one in order to gain from a higher wage in period two.[5] Thus, if the compensation of the agent in terms of an increased wage is not too high relative to the benefit of the principal from an aligned behaviour in the first period, public evaluation is preferred by the principal.[6] In any case, with an interim performance evaluation, there is more delegation in period one as the evaluation effectively reduces the economic risk of delegation for the principal.

Moreover, it is worth noting that the different outcomes in case of evaluation imply different effects on aggregate welfare compared to the situation without evaluation. As the exact nature of the (unique) outcomes depends the prior belief about the agent's type and the relation between wage schedule and utility, the details of the results are somewhat involved; see Section 2.5 for a discussion. Yet, at least two general observations can be made which are important from the perspective of a social planner. First of all, we find that, while wages have no direct impact on aggregate welfare,[7] they matter if the evaluation is public, namely through their (indirect) influence on outcomes. And, with high priors, giving all bargaining power to the agents maximises aggregate welfare. Moreover, we show that, in terms of aggregate welfare, any evaluation conducted ought to be disclosed to the public. Thus, inasmuch as allowed by the stylised argument given in the sequel, strengthening the agent's rights regarding bargaining power and accessibility of evaluations is welfare enhancing in situations like the one considered here.

Related Literature

As already indicated above, we are, of course, by no means the first to anaylse delegation of authority in organisations. Since the seminal work by Holmström (1977,

[5]For the case with more than two periods, which we consider as an extension, the effect persists for all but the last period (see Section 3.2 for details).

[6]Although the assumption of (partly) endogenous wages slightly complicates the analysis, the resulting overall two-period game is amenable to a common backward induction argument. Relying on standard equilibrium selection arguments (Cho and Kreps, 1987), we characterise the equilibria of the resulting signalling game, depending on the observability of the evaluation. These equilibria are uniquely determined by the prior belief about the agent being loyal and the specific relation between wages and utilities.

[7]Note that wages affect only the redistribution of payoffs and, hence, have no direct effect on aggregate welfare.

1984), a stream of literature has been concerned with the question of how to optimally allocate decision rights to potentially biased but better informed subordinate agents. And optimal mechanisms to delegate decisions have been studied, for example, by Melumad and Shibano (1991), or Alonso and Matouschek (2007, 2008). Moreover, granting authority to the agent may not only serve as a means to optimally use available information. As demonstrated by Aghion and Tirole (1997), granting authority also provides an incentive to the agent to engage in information acquisition as formal control rights are vacuous without knowing the state of the world. In a similar vein, Bester and Kraehmer (2008) show that delegating a project decision to the agent may also provide incentives to exert implementation effort. Adding to this literature, the present paper focuses on the incentive effects of delegation in a context of asymmetric information about the type of the agent where delegation in early rounds induces reputation concerns due to the evaluation.

While reputation concerns themselves have already been studied in a delegation context – for example, Englmaier et al. (2010) or Blanes i Vidal (2007) consider situations with the principal and the agent being equally uninformed about the agent's ability and analyse a model of career concerns (see Holmström, 1999)[8] – the present paper assumes that agents know their type ex ante in such a setting. However, inasmuch as optimal delegation mechanisms with asymmetric information about the agent's type have been studied before (e.g. Armstrong, 1995; Frankel, 2010; or Kovac and Kraehmer, 2012), reputation concerns are not central to the respective arguments.

Closer to the present discussion, Aghion et al. (2004) consider the optimal allocation of control rights contingent on announcing a type. They find that, if control is contractible, the loyal type will not be delegated any decision in order to induce truth-telling by the biased type. Different from the present setting, however, Aghion et al. consider a two stage (but one-shot) interaction where the principal benefits from learning the agent's type but where there are no incentives for pooling derived from later periods.

Results that are more similar to the ones derived in the present paper, however, have been obtained for situations in which the principal cannot commit to a delegation mechanism so that transfering control is cheap talk (Crawford and Sobel, 1982). In particular, a common finding in these models is that agents shade their reports in order to gain the principal's trust in the future (e.g. Benabou and Laroque, 1992, Morris,

[8]Ottaviani and Sorensen (2006 a and b) and Prat (2005) analyse career concerns when the principal can not commit to delegate authority.

2001, or Sobel, 1985).[9] While considering a different environment, these results are similar in spirit to the present ones, which show that, in a context with delegation, interim evaluations provide incentives for biased agents to align their choices with the principal's preferences.

A second important strand of literature, which is naturally related to the present discussion, is the literature on performance evaluations. In this area, most work is concerned with subjective performance evaluations and, thus, focuses on the problems caused by the non-verifiability of performance (e.g. Fuchs, 2007; or Prendergast and Topel, 1996).[10] In the present paper, by contrast, the performance measure is perfectly observable (although not verifiable by courts) and the – strategic – disclosure of performance information together with the resulting incentives for the agent are central to our results.

Another related type of argument is provided by Goltsman and Mukherjee (2011), who consider a multi-stage tournament where the strategic disclosure of performance information in the first period can increase effort incentives in the second period. Different from our setting, however, they analyse information disclosure to peers in a tournament who subsequently update winning probabilities. In contrast to this, we are concerned with the influence of signalling on outside option wages determined by market beliefs. Outside option wages, in turn, play a crucial role in Mukherjee (2008), who investigates strategic information disclosure to raiders in order to strengthen incentives for workers and shows that full disclosure is optimal. Although the main idea is related to the present one, Mukherjee assumes that principal and agent are equally uninformed about the agent's type, i.e. ex ante asymmetric information about types is not an issue.[11]

Finally, it is worth mentioning that the methodology applied in this paper goes back to the seminal paper by Spence (1973), who analyses signalling in a job market context. More specifically, the signalling game analysed in the sequel can be viewed as a modified version of the beer-quiche game first discussed by Cho and Kreps (1997)

[9] While repeated games arguments do not lie at the heart of the present discussion, it is worth noting that similar arguments have been given in the literature on repeated games and reputation (see, for example, Kreps and Wilson, 1982, Milgrom and Roberts, 1982, and Mailath and Samuelson, 2001, for other models with incomplete information, or Fudenberg and Levine 1989, 1992 and Cripps et al. 2004, 2007, for examples where the asymmetry of information between players stems from a short-lived player playing against a long-lived one).

[10] To give one example, Baker et al. (1994) show that supervisors may be tempted to underreport worker's performance and implicit incentives can be crowded out by explicit ones.

[11] Further works considering outside option wages in other contexts are Acemoglu and Pischke (1999) or Blanes i Vidal (2007).

except that we allow payoffs to be endogenous in that they may depend on posterior beliefs.

The rest of the paper is structured as follows: In Section 2, we introduce and analyse the model. In Section 3, we briefly discuss two possible extensions: a flexible timing of the evaluation and repeated evaluations. Section 4 concludes. All proofs are gathered in the appendix.

2 Model and Results

In the sequel, we describe the underlying delegation problem (Section 2.1). After briefly considering the benchmark case without evaluation in Section 2.2, we proceed to analyse the equilibria of the delegation game with private and with public intermediate evaluation (Section 2.3). As a next step, we compare the principal's optimal delegation decision in these three cases and analyse the principal's corresponding profit (Section 2.4). The analysis concludes with a discussion of welfare effects (Section 2.5).

2.1 The Underlying Delegation Problem

Consider the following standard delegation problem: A firm has to implement a project where the principal has the formal authority to decide what project is chosen, but he needs an agent to implement it.

Setup.
The set of possible projects, X, is given by a subset of the real line, i.e. $X \subseteq \mathbb{R}$. Both the principal and the agent enjoy a personal benefit from a project choice $x \in X$.

However, by assumption, the principal himself is unable to implement his preferred project, e.g. because he lacks some important information about local conditions which only the agent can acquire. Thus, in order to circumvent his lack of information, the principal can delegate the project decision to the agent.[12] The agent can be either of two types, $\tau \in T := \{b, l\}$. If the agent is loyal, $\tau = l$, his preferences are aligned with those of the principal. If the agent is biased, $\tau = b$, the preferences of the principal and the agent do not match. A priori, the principal and the market share a common prior that the agent is loyal with probability $p \in (0, 1)$, i.e. $p = Prob(\tau = l)$.

[12] In line with the delegation literature, this lack of information can be modelled by the realisation of some state of nature, θ, which is observed only by the agent. For the present purposes, we omit an explicit reference to θ as focus of the argument lies primarily on the agent's signalling motive (and not on the effects of different realisations of θ).

Remark 1 *For the purposes of the present discussion, we think of biases as reflecting differences in preferences. However, as should become clear from the subsequent argument, it is also possible to think of the agent's and the principal's preferences as being generally aligned. In that case, being biased would reflect a high cost when trying to satisfy these preferences for a given state of the world (and being unbiased would reflect a low cost of doing so). While some details of the modelling would of course change, the general thrust of the argument does not hinge on the interpretation.*

Utility

The players' utility is given by $u_i(x)$, $i \in \{P, b, l\}$. For the sake of argument, we assume that there always exists a unique optimal project x_i, $i \in \{P, b, l\}$, which maximises utility for player i, $i \in \{P, b, l\}$,[13] i.e.

$$x_i := arg \max_{x \in X} u_i(x) \ .$$

Note that, as the preferences of the biased agent differ from those of the principal, we have $x_b \neq x_P$, while the loyal agent prefers the same project, i.e. $x_l = x_P$.

The corresponding utility is referred to as u_i^+, i.e. for $i \in \{P, b, l\}$ we define

$$u_i^+ := u_i(x_i).$$

Similarly, utilities in case an undesired project is implemented are defined as follows:

$$\begin{aligned} u_b^- &:= u_b(x_P) \\ u_P^- &:= u_P(x_b) \\ u_l^- &:= u_l(x_b). \end{aligned}$$

Furthermore, if the principal retains authority, he chooses a default project x_P^0, which can be interpreted as the best choice according to the principal's knowledge. The *default utilities* realised in this case are denoted by

$$u_i^0 := u_i(x_P^0) \ , \ \textit{for all } i \in \{P, b, l\} \ .$$

In order to make the problem interesting, we assume that delegating to a loyal

[13] Here and in the following, we slightly abuse notation by sometimes referring to the agent's different types as player b and player l.

agent is profitable for the principal while delegating to a biased agent is harmful, i.e.

$$u_P^- < u_P^0 < u_P^+ \,.$$

The biased (loyal) agent's default utility u_b^0 (u_l^0), in turn, is always smaller than his optimal choice, i.e. $u_b^0 < u_b^+$ ($u_l^0 < u_l^+$); whether u_b^0 (u_l^0) is smaller or larger than u_b^- (u_l^-) need not be specified here.

Contracts and Wages.
As usual, the principal allocates authority, i.e. he determines whether he or the agent decides on the project x. The actual choice of x is not contractible, though. Only the allocation of authority can be contractually fixed. In either case, the principal pays a wage $w(\mu)$ to the agent, which depends on the market's belief about the agent being loyal, μ. Thus, while not restricting attention to common competitive wages, we still assume that the wage paid by the principal correctly matches the agent's expected outside option. In that sense, wages are considered endogenous, and depend on the observability of the agent's project decision.

Note that the principal benefits from compensating the agent for potential upward adjustments in the market's belief whenever the payoff-surplus from the increased probability of employing the loyal agent outweighs the loss through increased wages. For high priors, this is generally the case if outside wages do not increase too fast.[14] In order to ensure that this is also the case for small priors, we assume that the agent is needed for the implementation of the project and that employing a new or no agent is costly for the principal, e.g. because a new agent has to acquire firm-specific knowledge.[15]

Moreover, although the project choice is observable to the principal (and to the market), we assume that it is not enforceable by the courts. Thus, the principal cannot reward the agent for a loyal decision, but instead pays a flat wage which can be adjusted after a public evaluation.[16]

[14]Formally, we need $w(\mu)-w(p) < \min\left\{(\mu-p)(u_P^+-u_P^-), \mu(u_P^+-u_P^-)-(u_P^0-u_P^-)\right\}$ for all $\mu > p$.

[15]The point to note here is that, for small priors p and certain parameters, the only incentive for the biased agent to choose the loyal action, x_l, with positive probability (if the decision is delegated) is the increased wage in Period 2 because the principal may centralise in Period 2 no matter which project the agent chooses. In that case, the principal would be better off paying only $w(p)$ – such that the agent leaves – and employing a new agent. However, assuming the agent to be needed for the implementation of the project and that employing a new or no agent is costly, with cost being larger than $w(\eta) - w(p)$ (where η is the delegation threshold without evaluation; see Section 2.2 for details), ensures that also these special cases remain tractable.

[16]In case of private evaluations, we assume that downward adjustments of wages require publicly

As mentioned before, the principal and the market share a common prior that the agent is loyal with probability $p \in (0,1)$. Accordingly, without any further (verifiable) information about the agent, i.e. if no evaluation takes place or evaluation is private, the common market wage for the agent is $w(p) \geq 0$.

Finally, we normalise $w(0) = 0$ and assume that w is (weakly) increasing in the market belief that the agent is loyal, μ.[17]

It is important to note that, if it is commonly known that the agent is loyal, the maximal profit the principal expects to earn, net of what he is able to generate himself, is given by $u_P^+ - u_P^0$. In the following, we therefore assume that the wage cannot exceed this threshold, i.e.

$$w(\mu) < u_P^+ - u_P^0 \qquad \forall \mu \in (0,1)$$

and $w(1) \leq u_P^+ - u_P^0$.[18]

Furthermore, we assume that w is continuous and weakly convex, i.e. for all $x, y \in [0,1]$ we have:[19]

$$w(tx + (1-t)y) \leq tw(x) + (1-t)w(y) \qquad \forall t \in [0,1].$$

In the sequel, we divide the agent's employment with the principal into two equal periods – before and after an evaluation takes place – and refer to these as Period 1 and Period 2, respectively. In the baseline case, where the agent is not evaluated, these two periods are equivalent and the respective payoffs are considered to be generated in each period.

Summing up, with the above specifications, total per period payoffs for the principal can be specified as

$$U_P(x, \mu) = u_P(x) - w(\mu) \ ,$$

verifiable information (public evaluation) and upward wage adjustments are never profitable for the principal as the agent's outside option remains fix.

[17] Assuming the agent to be liability constrained does not change the results of the paper which essentially rely on the monotonicity of wages.

[18] Here we implicitly assume that all principals face the same constraints so that this threshold will not exclude any of them from the market.

[19] From a technical point of view, if wages are assumed to be competitive and thus reflecting the expected marginal product of a worker given the belief about his loyalty, one would have to pay negative wages in case $p \leq \eta$. If agents are protected by limited liability, principals are forced to pay a constant wage of zero for low prior beliefs, and a linearly increasing wage for $p > \eta$. This case provides an example for a weakly convex wage schedule. Besides lacking a real-world justification, concave wage functions do not provide any new results, which is why they are not considered.

and the corresponding expression for the agent is given by

$$U_\tau(x, \mu) = u_\tau(x) + w(\mu) \ .$$

2.2 No Evaluation

To begin with, consider the case where the principal does not evaluate the agent so that reputation is not an issue. In this case, the principal has to decide whether or not to delegate decision authority to the agent at the beginning of the employment phase.

Obviously, without interim evaluation, the principal cannot observe the agent's choice after Period 1 so that the agent always chooses his preferred project x_τ if the decision is delegated to him. Accordingly, the principal's expected per period benefit from project choice in case of delegation is given by:

$$E[u_P(x_\tau)] = pu_P^+ + (1-p)u_P^-.$$

This immediately leads to Proposition 1.

Proposition 1 *Without intermediate evaluation, the principal prefers delegation to centralisation if*

$$p \geq \eta := \frac{u_P^0 - u_P^-}{u_P^+ - u_P^-}.$$

2.3 Equilibria With Evaluation

As a next step, consider the case where the principal evaluates the agent in the middle of his employment phase.

Private Evaluation.

Recall that, by assumption neither the project choice nor the allocation of authority in either period can be contractually fixed ex ante. However, we now assume that the principal privately observes the agent's project choice after the first half of the employment (given that it has been delegated to the agent). In this case, the agent's outside option remains unchanged as the market remains ignorant about the outcome of the evaluation so that the agent's wage in Period 2 is again given by $w(p)$. The principal, however, might reconsider his delegation decision at the beginning of Period 2.

More specifically, in Period 1, the principal decides whether or not to delegate the decision to the agent. The action space of the principal, thus, is given by $A_P = \{D, C\}$, consisting of delegation, D, and centralisation, C. If the principal retains his decision authority, C, he always takes the default decision and his belief about the agent's loyalty is not updated. If the principal delegates his decision authority to the agent, D, the agent decides which project to implement. At the end of Period 1, the principal's profit is realised and the principal is informed about the agent's project choice.

In Period 2, the principal faces essentially the same delegation problem as in the first except that he now can utilise the agent's observed behaviour from Period 1 to update his belief about the type of the agent. The updated beliefs are denoted as follows:

$$\begin{aligned} \mu_+ &:= prob(\tau = l | x = x_l) \\ \mu_- &:= prob(\tau = l | x = x_b). \end{aligned}$$

The resulting signalling game that arises after delegation in Period 1 is illustrated in Figure 1.

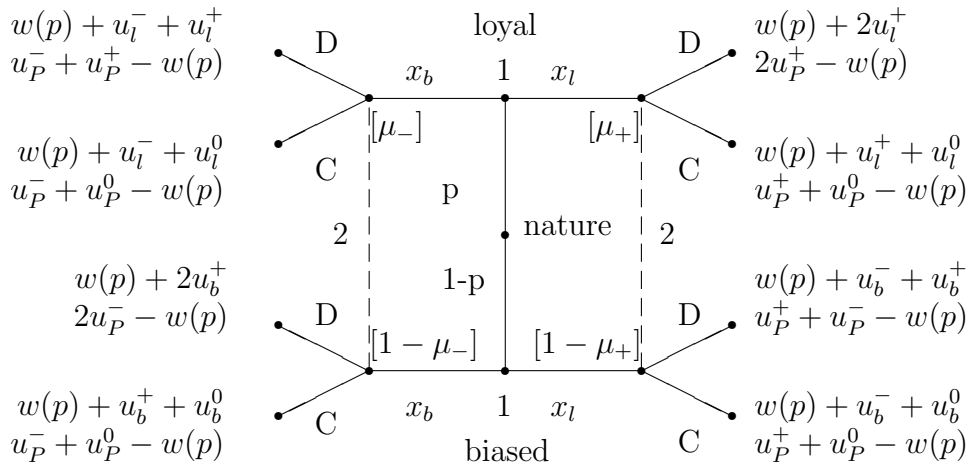

Figure 1: Graphical illustration of the strategic situation with private evaluation after a delegation decision of the principal in Period 1. Player 1, the sender, is the agent and player 2, the receiver, is the principal. First period wages are neglected as they are of no strategic relevance.

In the sequel, we analyse the subgame that arises after delegation by the principal

in Period 1 while neglecting Period-1-wages (as these have no strategic impact). In doing so, we denote the action the principal chooses in Period 2 upon observing x by $a_P(x)$. A strategy of the principal in the subgame, thus, consists of a pair $s_P = (a_P(x_l), a_P(x_b))$. A strategy for the agent in the resulting subgame, in turn, comprises the choice of an action both in case of delegation in Period 1 (which has occurred by assumption) and in Period 2. As it is strictly dominant for the agent to choose his preferred action whenever the decision is delegated in Period 2, we focus the analysis on the agent's strategy in Period 1, denoted by s_τ, which is given by specifying $x \in \{x_b, x_l\}$.

Noting that it is dominant for the principal to choose D (C) upon observing x_l (x_b resp.) if $\mu_+ \geq \eta$ ($\mu_- \leq \eta$) and defining

$$\Delta_b := u_b^+ - u_b^-,$$

i.e. Δ_b is the biased agent's utility differential between choosing x_b and x_l, we obtain the following equilibria;[20] see Figure 2 for illustration.[21]

Lemma 1 *For the signalling game depicted in Figure 1, the following three types of equilibria are compatible with the Intuitive Criterion (Cho and Kreps, 1987):*

- *If $\Delta_b < u_b^+ - u_b^0$ and $p > \eta$:*
  Pooling (pool(D,C)): $s_P = (D, C)$, $s_l = s_b = x_l$

- *If $\Delta_b < u_b^+ - u_b^0$ and $p \leq \eta$:*
  Principal and biased agent randomise (mix2(λ_P,C)):
  $s_P = \lambda_P D + (1 - \lambda_P)C$, $s_l = x_l$, $s_b = \lambda x_l + (1 - \lambda)x_b$, ($\lambda = 1$ if $p = \eta$)

- *If $\Delta_b = u_b^+ - u_b^0$:*
  Biased agent randomises (mix1(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = \lambda x_l + (1-\lambda)x_b$

- *If $\Delta_b > u_b^+ - u_b^0$:*
  Separating (sep(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = x_b$

The main point to note here, is that the biased agent's incentive to choose his preferred project in Period 1 is stronger the larger his utility from centralisation in Period 2, i.e. u_b^0, is compared to his utility from taking an opportunistic choice in

[20] If $\mu = \eta$, the principal is indifferent which is why the dominance is not strict.

[21] The structure of these equilibria is robust to introducing some noise into the principal's observation of the agent's project choice.

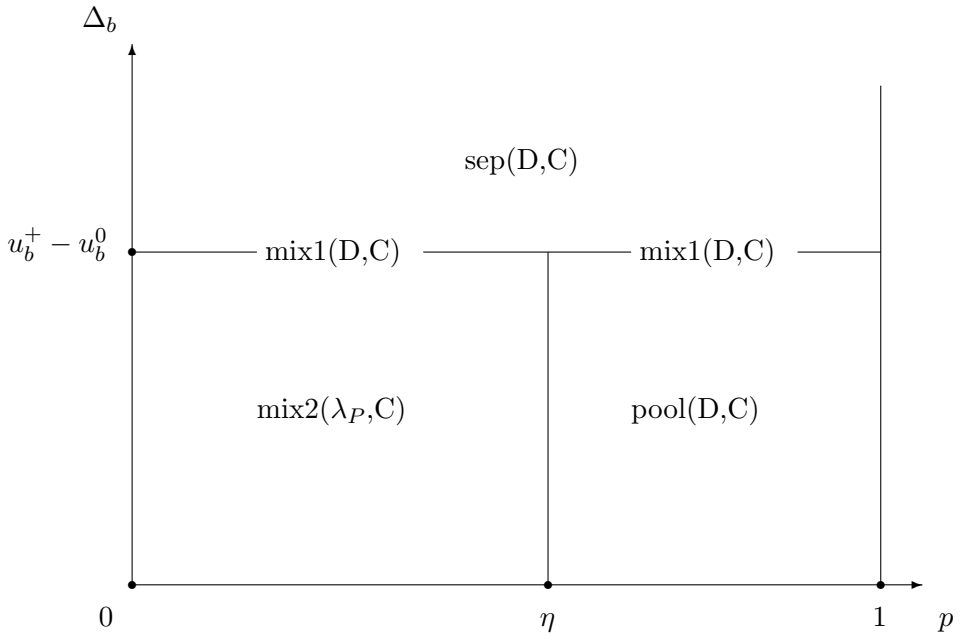

Figure 2: Equilibria of the signalling game with private evaluation depending on the prior belief p and the biased agent's utility differential, Δ_b. Note that outside of the null-set where the mix1(D,C) equilibrium or the mix2(λ_P,C) equilibrium with $p = \eta$ is played, all equilibria are unique for a given pair of parameters (p, Δ_b).

Period 1, u_b^-. Put differently, an equilibrium where the biased agent separates is more likely the larger the difference between the biased agent's utility differential Δ_b and $u_b^+ - u_b^0$. The principal on the other hand chooses to centralise in Period 2 whenever he observes a separating behaviour. Upon observing the loyal project, though, he again delegates the decision in Period 2 – if his prior belief about the agent being loyal is high enough.

Public Evaluation.
Finally, we consider the case where the principal evaluates the agent and reveals this information to the market. In this case, the agent's outside option might change as a function of μ, the principal's and the market's posterior belief about the agent's loyalty after observing his project choice. More specifically, with public evaluation, the agent's wage depends on the updated beliefs, μ_+ and μ_-; the signalling game which arises after delegation in Period 1 is illustrated in Figure 3.

The analysis of the resulting signalling game is analogous to the one in case of private evaluation. The results are provided in the following Lemmata.

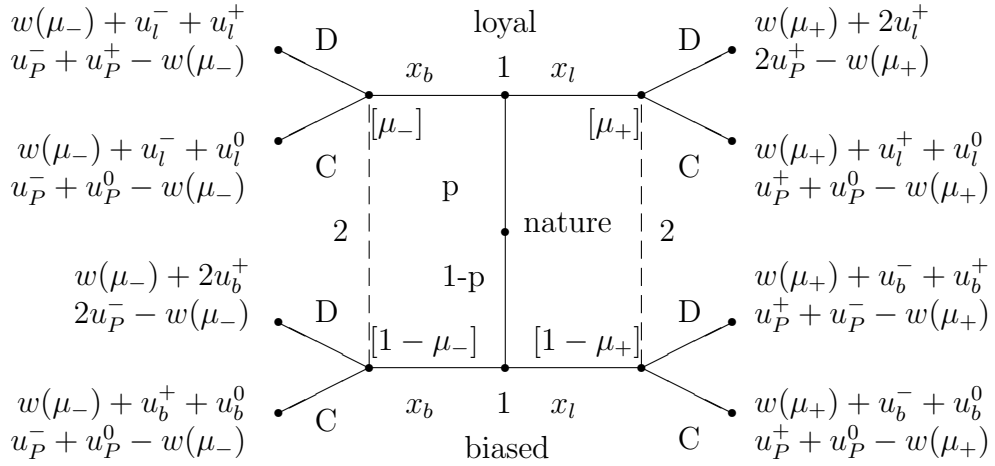

Figure 3: Graphical illustration of the strategic situation with public evaluation after a delegation decision of the principal in Period 1. Player 1, the sender, is the agent and player 2, the receiver, is the principal. First period wages are neglected as they are of no strategic relevance.

Lemma 2 *If $p > \eta$, the following three types of equilibria of the signalling game depicted in Figure 3 are compatible with the Intuitive Criterion:*

- *If $\Delta_b \leq w(p) + u_b^+ - u_b^0$:*
  Pooling: (pool(D,C)) $s_P = (D, C)$, $s_l = s_b = x_l$
- *If $w(p) + u_b^+ - u_b^0 < \Delta_b < w(1) + u_b^+ - u_b^0$:*
  Biased agent randomises (mix1(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = \lambda x_l + (1-\lambda) x_b$
- *If $\Delta_b \geq w(1) + u_b^+ - u_b^0$:*
  Separating (sep(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = x_b$

Thus, similar to the case with private evaluation, we find that, for $p > \eta$, the biased agent is more likely to imitate the loyal agent if his utility differential, Δ_b, is small. However, with public evaluation the biased agent benefits from a loyal project choice after delegation in Period 1 not only through additional delegation in Period 2 but also through an increased wage in Period 2. Thus, pooling incentives and, hence, the area where the biased agent imitates the loyal one are increased. Moreover, there is a non-degenerate area where the biased agent trades off wage and project-choice based incentives and therefore randomises between project choices (mix1(D,C)).

For the case of low prior beliefs, considered in the lemma below, we also observe a larger area for the mix1(D,C) equilibrium. Furthermore, if the biased agent's utility differential, Δ_b, is sufficiently small, his incentive to opt for his preferred project is dominated by an increased wage resulting from the loyal project choice, even if the principal centralises either way. The principal, in turn, is not willing to delegate in Period 2, even after observing his preferred project choice as his prior belief about the agent's loyalty is too small.

Lemma 3 *If $p \leq \eta$, the following three types of equilibria of the signalling game depicted in Figure 3 are compatible with the Intuitive Criterion:*

- *If $\Delta_b \leq w(p)$:*
  Pooling (pool(C,C)): $s_P = (C, C)$, $s_l = x_l$, $s_b = x_b$, (Principal randomises if $p = \eta$)
- *If $w(p) < \Delta_b \leq w(\eta)$:*
  Biased agent randomises (mix1(C,C)): $s_P = (C, C)$, $s_l = x_l$, $s_b = \lambda x_l + (1-\lambda)x_b$
- *If $w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0$:*
  Principal and biased agent randomise (mix2(λ_P,C)):
  $s_P = \lambda_P D + (1 - \lambda_P)C$, $s_l = x_l$, $s_b = \lambda x_l + (1 - \lambda)x_b$, ($\lambda = 1$ if $p = \eta$)
- *If $w(\eta) + u_b^+ - u_b^0 \leq \Delta_b < w(1) + u_b^+ - u_b^0$:*
  Biased agent randomises (mix1(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = \lambda x_l + (1-\lambda)x_b$
- *If $\Delta_b \geq w(1) + u_b^+ - u_b^0$:*
  Separating (sep(D,C)): $s_P = (D, C)$, $s_l = x_l$, $s_b = x_b$

Note that with public evaluation, all equilibria outside of the null-set $\{p = \eta\} \cap \left\{\Delta_b \leq w(\eta) + u_b^+ - u_b^0\right\}$ are unique for a given pair of parameters (p, Δ_b). Figure 4 illustrates the resulting equilibria and their relation to the underlying wage schedule.

In contrast to the situation with private evaluation, the biased agent now benefits from pooling with the loyal type not only because he may decide on the project in Period 2 but also because he receives a higher wage in the later period. Accordingly, the pooling equilibrium is played up to larger values of Δ_b if the prior belief is high. The mixed equilibrium, in turn, extends to a non-null set for $p \geq \eta$ because the biased agent's wage, given by $w(\mu_+) = u_b^0 - u_b^-$, endogenously adapts to changes in Δ_b, thereby keeping him indifferent for a continuum of values of Δ_b. Thus, if the prior

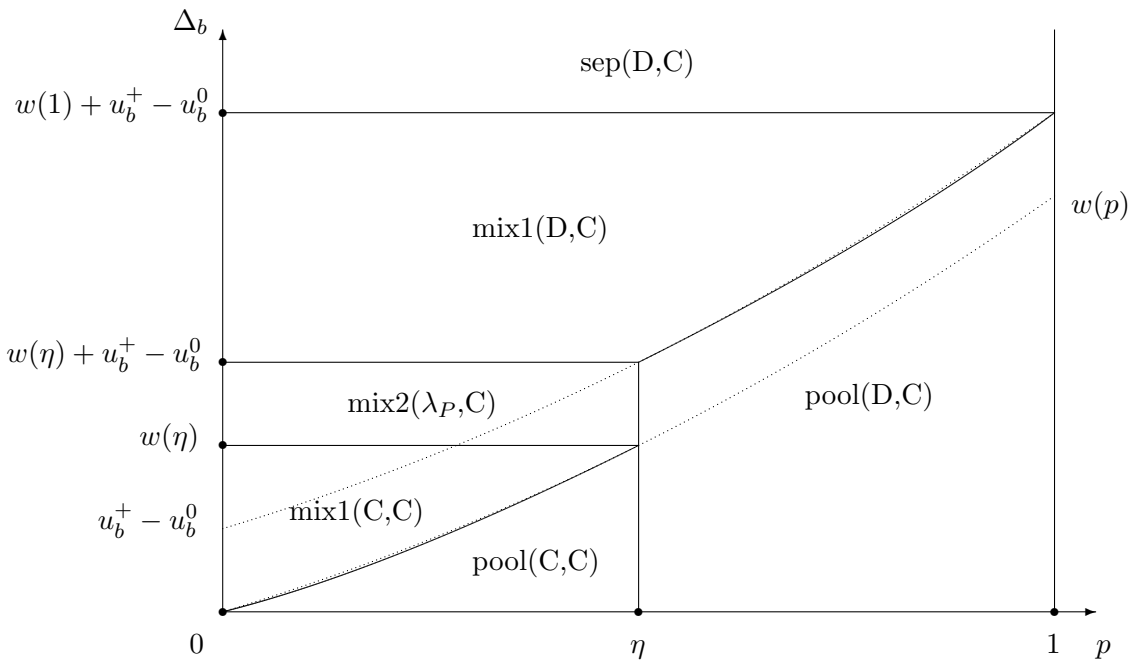

Figure 4: Equilibria of the signalling game with public evaluation depending on prior belief p and the biased agent's utility differential, Δ_b – for a given wage schedule (indicated by the lower curve). The upper curve depicts the wage schedule increased by $u_b^+ - u_b^0$. Solid lines determine the borders of the respective equilibrium area.

belief is low, the principal might even choose to centralise in Period 2 independent of the agent's decision as the wage schedule entails some incentive to pool for the biased agent – in contrast to the case where evaluation is private.

2.4 The Principal's Behavior

We proceed to analyse whether the principal prefers to delegate the project decision to the agent in Period 1, depending on his prior belief about the agent's loyalty and the utility parameters (Section 2.4.1). Having done so, we investigate if the principal benefits from an evaluation, given the policy of information disclosure, and whether he chooses to disclose an evaluation if he is allowed to do so (Section 2.4.2).

2.4.1 The Principal's Delegation Decision

In case the principal retains authority in Period 1, the belief about the agent's loyalty cannot be updated and the principal delegates in Period 2 if $p \geq \eta$. Thus, the overall

payoff of the principal in case of centralisation, depending on p, is given by

$$E[U_P|C] = \begin{cases} 2u_P^0 - w(p) & \text{if } p < \eta, \\ u_P^0 + pu_P^+ + (1-p)u_P^- - w(p) & \text{else,} \end{cases}$$

where the first-period wage $w(p)$ is again left out.

In the sequel, we compare the principal's payoffs with and without evaluation depending on whether the evaluation is public or private.

Private Evaluation.
In case of a private evaluation, it is straightforward to show that the principal is more likely to delegate in the first period compared to the situation without evaluation.

Proposition 2 *If the principal can privately evaluate the agent, he delegates more often in Period 1 than in an interaction without evaluation. In particular, the principal delegates for all $p \geq p_U^*(\Delta_b)$, where p_U^* is given by*

$$p_U^*(\Delta_b) = \begin{cases} \eta^2 & \textit{if } \Delta_b < u_b^+ - u_b^0, \\ \in \left[\eta^2, \frac{\eta}{2-\eta}\right] & \textit{if } \Delta_b = u_b^+ - u_b^0, \\ \frac{\eta}{2-\eta} & \textit{if } \Delta_b > u_b^+ - u_b^0. \end{cases}$$

Furthermore, $\eta^2 < \frac{\eta}{2-\eta} < \eta$, thus

- *the principal delegates for a larger range of prior beliefs than without evaluation, in which case he delegates only for $p \geq \eta$, and*
- *the principal delegates more the lower the biased agent's concerns about the decisions are.*

Note that this result is rather intuitive. The biased agent has an incentive to deviate from his preferred option in Period 1 in order to signal loyalty if he does not suffer too much from an opportunistic choice in the first period.[22] By delegating the project decision to the agent, the principal benefits either from learning the agent's type (if Δ_b is high), or from more aligned decision making (if Δ_b is low). Thus, in either case the principal is more likely to delegate the decision to the agent than in the case without evaluation.

[22]Cast in the ability interpretation indicated in Remark 1, the biased agent would invest a high effort/cost, which is individually suboptimal in the setting without evaluation, in order to imitate the more able agent.

Moreover, aligning the biased agent's preferences is in fact more beneficial to the principal than learning his type. Therefore, the principal is more likely to delegate in the pooling equilibrium than in the separating equilibrium – indicated by $\eta^2 < \frac{\eta}{2-\eta}$.

Public Evaluation.
If the evaluation is observable to the market, the principal has to pay a higher wage to the biased agent in a pooling equilibrium than in a separating one. Still the principal is more likely to delegate the project choice than in the case without evaluation.

Proposition 3 *If evaluations are public, the principal delegates more often in Period 1 if he evaluates the agent than in an interaction without evaluation. In particular, he delegates if $p \geq p_O^*(\Delta_b)$, where the threshold $p_O^*(\Delta_b)$ is implicitly defined via*

$$p_O^* = \delta_{w,\mu}(p_O^*) = \frac{[u_P^0 - u_P^- - w(p_O^*)]\mu}{(u_P^+ - u_P^-)\max\{\mu, \eta\} + (u_P^+ - u_P^0) - w(\mu)}$$

and $\mu(\Delta_b) \in [0,1]$ is given by

$$\mu(\Delta_b) = \begin{cases} w^{-1}(\Delta_b) & \text{if } \Delta_b \leq w(\eta), \\ \eta & \text{if } w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0, \\ w^{-1}(\Delta_b - (u_b^+ - u_b^0)) & \text{if } w(\eta) + u_b^+ - u_b^0 \leq \Delta_b < w(1) + u_b^+ - u_b^0, \\ 1 & \text{if } \Delta_b \geq w(1) + u_b^+ - u_b^0, \end{cases}$$

Furthermore, $\delta_{w,\mu}(p) \leq \eta$ for all $p \in [0, \eta]$, so the principal delegates for a larger range of prior beliefs than without evaluation, in which case he delegates the decision only for $p \geq \eta$.

Note that the notion of the inverse wage function is valid in the above definition even for weakly increasing wages, as the domain of Δ_b is defined by means of the image of w.

The main point to note here is that the biased agent again has an incentive to deviate from his preferred option in Period 1 in order to signal loyalty – essentially as in the case with private evaluation. However, in exchange for this, he now not only is more likely to be granted the right to decide over the project but also receives a higher wage in Period 2. In combination, this induces more aligned decisions in the first period of the relationship and, hence, more delegation.

Furthermore, even if wages matter, the principal still benefits more from aligning the biased agent's preferences than from learning his type and paying him less in the

later stage. Therefore, we obtain a similar result as in the case of private evaluation: The principal is more likely to delegate if the probability for the biased agent to choose the loyal project is higher.

Finally, we can show that the delegation threshold increases in Δ_b, that is, the principal benefits more from aligning the biased agent's preferences than from learning his type and paying him less in the later stage.

Proposition 4 *All other things equal, the principal delegates more the lower the biased agent's concerns about the decision are, i.e. the delegation threshold $p_O^*(\Delta_b)$ fulfilling $p_O^* = \delta_{w,\mu}(p_O^*)$ increases in Δ_b.*

Assuming wages to be linear in μ, the increasing delegation threshold can be determined explicitly, as stated in Corollary 1.

Corollary 1 *If wages are linear, i.e. $w(\mu) = \mu W$, the delegation threshold in Proposition 3 is given by*

$$p_O^*(\Delta_b) = \begin{cases} \frac{\eta}{W}\Delta_b & \text{if } \Delta_b \leq \eta W, \\ \eta^2 & \text{if } \eta W < \Delta_b < \eta W + u_b^+ - u_b^0, \\ \frac{(u_b^0 - u_b^-)\eta}{(u_b^0 - u_b^-) + (1-\eta)W} & \text{if } \eta W + u_b^+ - u_b^0 \leq \Delta_b < W + u_b^+ - u_b^0, \\ \frac{\eta}{2-\eta} & \text{if } W + u_b^+ - u_b^0 \leq \Delta_b. \end{cases}$$

A Comparison of Public and Private Evaluation.

A natural question that arises from the above analysis is how the observability of the evaluation influences the principal's delegation decision. At first sight, it might seem intuitive to assume that the principal is less likely to delegate if the evaluation is public in order not to reveal too much information to the market and thereby possibly increasing the agent's outside option. In the present setting, however, this may not be true. In fact, exposing the agent's behaviour to the market exposing the agent's behaviour to the market can be beneficial to the principal if this changes the biased agent's behaviour. However, as we will see below, this effect crucially depends on the shape of the wage schedule.

Proposition 5 shows that, if the biased agent's concern about the project decision, Δ_b, is high enough, the principal is more likely to delegate if the evaluation is not observed by the market.

Proposition 5 *If w is convex, there exists a threshold $\rho \in [0, w(1) + u_b^+ - u_b^0]$, such that for $\Delta_b \geq \rho$ delegation is more likely if the evaluation is private than if it is public.*

Note that the level of the threshold strongly depends on the shape of the wage function and decreases with its convexity. In fact, for "very convex" wage functions, ρ can be smaller than $w(\eta)$, such that public evaluation grants more discretion to the agent only if his concern about the project choice is negligible.

Figure 5 illustrates the results from Proposition 5; for the case of public evaluation, it distinguishes two cases:

- w is "not too convex", i.e. $\frac{\eta}{2-\eta}w(\eta) - \eta w(\frac{\eta}{2-\eta}) \leq (1-\eta)\frac{\eta}{2-\eta}(u_P^+ - u_P^0)$,
- w is "very convex", i.e. $\frac{\eta}{2-\eta}w(\eta) - \eta w(\frac{\eta}{2-\eta}) \geq (1-\eta)\frac{\eta}{2-\eta}(u_P^+ - u_P^0)$,

where p_{VC}^* refers to the "very convex" case and p_{NC}^* to the "not too convex" case. These two thresholds are shown as solid lines, whereas the delegation threshold in case of private evaluation is depicted as a dashed line. The benchmark threshold where no evaluation takes place is indicated by a dotted line.

Note that the shape of $w(\mu)$ does not allow for any conclusions about the shape of $p_O^*(\Delta_b)$ so that Figure 5 is just an illustration of the delegation thresholds.

Moreover, considering linear wages, there is weakly more delegation with public evaluation.

Corollary 2 *If wages are linear, delegation is (weakly) more likely if the evaluation is public than if it is private.*

Figure 6 illustrates the results from Corollary 2. In particular, it shows the delegation thresholds depending on the evaluation regime.

2.4.2 The Principal's Profit

In the following, we analyse the rent effects for the principal resulting from an intermediate evaluation. The main point to note here is that a separation of agents as well as an alignment of the agents' behaviour (in Period 1), which only occur when evaluation takes place, can be beneficial for the principal.

Private Evaluation.
For the principal, an intermediate evaluation can be profitable either because he learns the agent's type or because the biased agent imitates the loyal one by choosing the

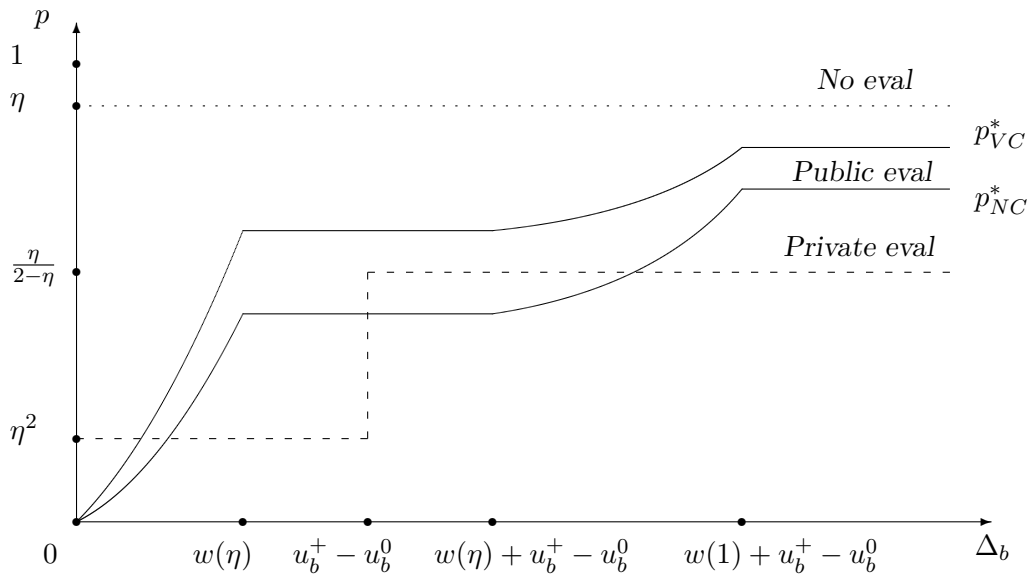

Figure 5: Illustration of the delegation threshold in case of public evaluation, p^* (solid lines). The upper graph, p^*_{VC}, refers to the "very convex" case, the lower graph, p^*_{NC}, to the "not too convex" case. The dashed (dotted) line illustrates the delegation threshold in case of private evaluation (no evaluation); here we assume $w(\eta) < u_b^+ - u_b^0$ (the opposite case looks similar). Except for the constant parts of p^*, the slope and curvature of the graph are only exemplary – the shape cannot be inferred from the shape of $w(\mu)$.

principal's preferred project in the first period. If the evaluation is not observable to the market, outside options do not change and the principal does not have to reimburse the agent for an opportunistic choice – as opposed to the case where the evaluation is public. Thus, the following Proposition is immediate.

Proposition 6 *If the interim evaluation is private, the principal is always (weakly) better off by evaluating the agent.*

In order to see this, consider any equilibrium where the principal plays (D, C) and the biased agent plays a mixed strategy $(\lambda, 1-\lambda)$ with $\lambda \in [0, 1]$. Then the principal's expected total profits with (unobservable) private evaluation, Π_U, net of the profits with no evaluation, Π_N, are given by

$$E[\Pi_U - \Pi_N] = (1-p)\lambda(u_P^+ - u_P^0) + (1-p)(u_P^0 - u_P^-).$$

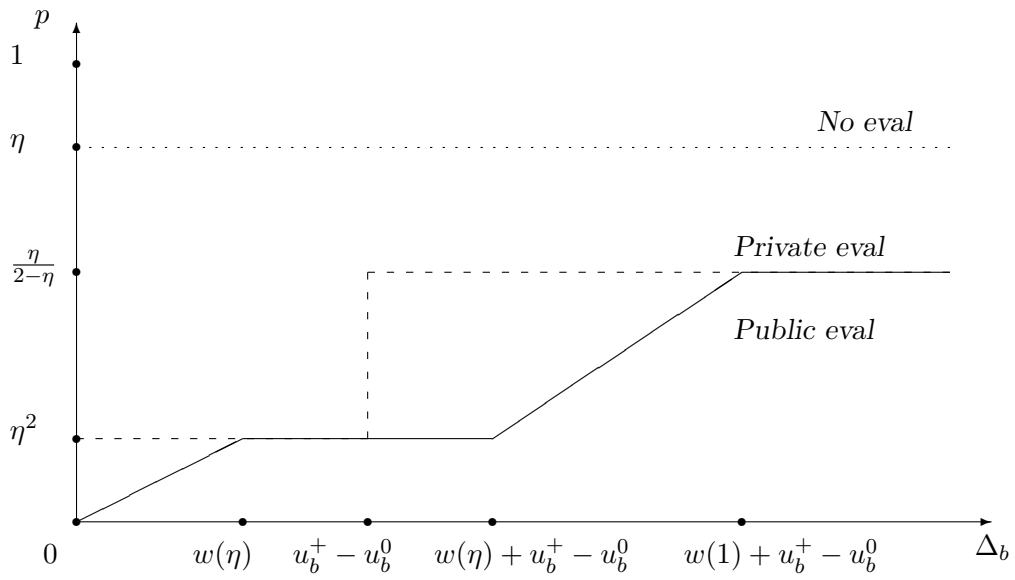

Figure 6: Illustration of the delegation threshold in the three cases of no evaluation (dotted line), private evaluation, p_U^* (dashed line), and public evaluation, p_O^* (solid line). Here, wages are linear and $w(\eta) < u_b^+ - u_b^0$ is assumed; the opposite case looks similar.

As this term is non-negative for all $\lambda \in [0, 1]$, it is immediate that the principal always prefers to evaluate for $p > \eta$ if the result is not observable to the market. A similar argument holds for intermediate prior beliefs where the principal still decides to delegate in Period 1 if a private evaluation takes place. If the prior belief is too low for the principal to delegate in Period 1, he is indifferent between evaluating or not.

Public Evaluation.
If the interim evaluation is observable to the market, the principal has to later compensate the biased agent for an earlier opportunistic behaviour. Thus, it is no longer true in general that evaluating the agent benefits the principal. In fact, whether the biased agent's aligned behaviour in Period 1 outweighs the possible losses in terms of higher Period-2-wages depends on which equilibrium is played in the respective case.

In particular, in case of high prior beliefs, the principal plays (D, C) and the biased agent plays a mixed strategy $(\lambda, 1 - \lambda)$ with $\lambda \in [0, 1]$ in any equilibrium. Thus, the principal's expected total profits with (observable) public evaluation, Π_O, net of the

profits with no evaluation, Π_N, are given by

$$E[\Pi_O - \Pi_N] = (1-p)\left[\lambda(u_P^+ - u_P^0) + (u_P^0 - u_P^-)\right] - \left[[p + (1-p)\lambda]\, w(\mu_+) - w(p)\right],$$

where $\lambda \in [0,1]$ denotes the probability of the biased agent choosing x_l.

The above formula illustrates that the principal trades off his earnings from the agent's project choice in Period 1, which are increasing in the biased agent's probability to imitate the loyal one, against the wage payment to the agent, also increasing in λ.

In particular, in a pooling equilibrium, evaluating the agent publicly does not change his outside option, while the principal still profits from the biased agent imitating the loyal one. Thus, the principal decides to evaluate the agent in this case. However, as the probability of the biased agent choosing the principal's preferred project decreases, the expected wage payment to the agent increases faster than the corresponding gain from the agent's project choice due to convexity of wages. Thus, the principal prefers to evaluate the agent for $p > \eta$ whenever there is "enough pooling" and his scope of evaluation depends on the shape of the wage function.

For low prior beliefs, the principal centralises authority in Period 1, and the principal prefers to evaluate the agent if and only if he delegates the project choice in case of a public evaluation.

Proposition 7 summarises these findings.

Proposition 7 *If $p > \eta$ and the interim evaluation is observable to the market, the principal is better off by evaluating the agent if and only if*

$$pw(\mu_+) - \mu_+ w(p) \leq p(1-\mu_+)(u_P^+ - u_P^0) + (1-p)\mu_+(u_P^0 - u_P^-).$$

In particular, for all $p > \eta$, there exists a threshold $\rho(p) > 0$ such that an evaluation is preferred for all $\Delta_b \leq \rho(p)$, i.e. if the biased agent's concern about the project choice is low enough.
If $p \leq \eta$, the principal weakly prefers to evaluate the agent if the evaluation is public.

Remark 2 *The posterior belief $\mu_+ \in [0,1]$, in case a public evaluation takes place, is endogenously given by the model parameters. In particular, if $p \leq \eta$, we have*

$$\mu_+(\Delta_b) = \begin{cases} p & \text{if } \Delta_b \leq w(p), \\ w^{-1}(\Delta_b) & \text{if } w(p) < \Delta_b \leq w(\eta), \\ \eta & \text{if } w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0, \\ w^{-1}(\Delta_b - (u_b^+ - u_b^0)) & \text{if } w(\eta) + u_b^+ - u_b^0 \leq \Delta_b < w(1) + u_b^+ - u_b^0, \\ 1 & \text{if } \Delta_b \geq w(1) + u_b^+ - u_b^0, \end{cases}$$

and if $p > \eta$, this yields

$$\mu_+(\Delta_b) = \begin{cases} p & \text{if } \Delta_b \leq w(p) + u_b^+ - u_b^0, \\ w^{-1}(\Delta_b - (u_b^+ - u_b^0)) & \text{if } w(p) + u_b^+ - u_b^0 < \Delta_b < w(1) + u_b^+ - u_b^0, \\ 1 & \text{if } \Delta_b \geq w(1) + u_b^+ - u_b^0. \end{cases}$$

Moreover, again assuming wages to be linear, it is easy to see that $pw(\mu_+) - \mu_+ w(p) = 0$. Furthermore, as $p(1-\mu_+)(u_P^+ - u_P^0) + (1-p)\mu_+(u_P^0 - u_P^-) \geq 0$, the following corollary is immediate.

Corollary 3 *If wages are linear and the interim evaluation is observable to the market, the principal is (weakly) better off by evaluating the agent.*

A Comparison of Public and Private Evaluation.

Although we have seen that the principal always prefers to evaluate the agent for high prior beliefs if the evaluation is private – which is not necessarily the case of public evaluation – this does not allow us to conclude that the principal would always prefer to conceal the evaluation if he had the choice. In fact, for a given prior $p > \eta$ and a given utility differential Δ_b, the behaviour of the biased agent might change if his evaluation is public because wage considerations then play a role. Thus, equilibria played in the two regimes might be different. Figure 7 illustrates this change of the biased agent's behaviour for $p > \eta$.

For example, if the biased agent pools in the public regime in order not to lose his wage, he separates in a private regime for the same parameter values. In this case, the principal is better off in the public regime where he gets a profit of

$$E[U_P|O, pool(D,C)] = 2pu_P^+ + (1-p)(u_P^+ + u_P^-) - w(p),$$

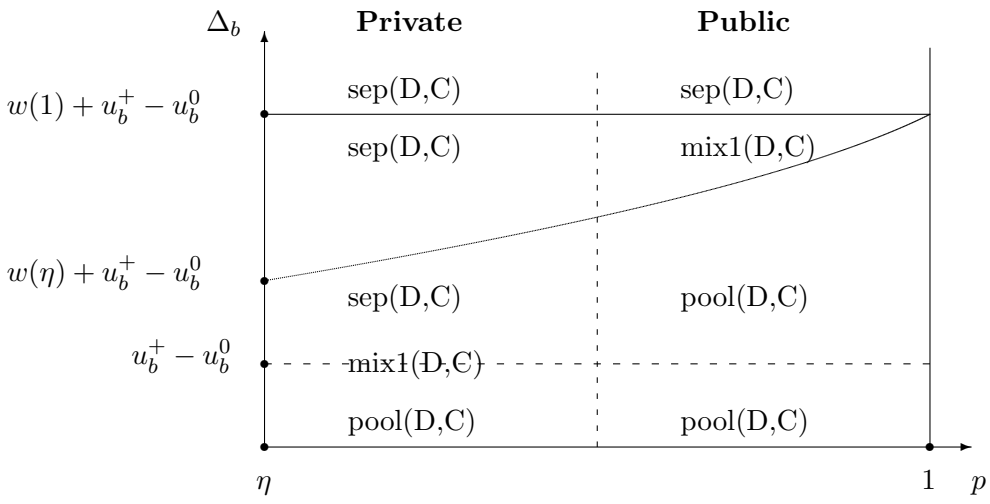

Figure 7: Illustration of the different equilibria played in case $p \geq \eta$. The left-hand side states the equilibria played in a private regime, whereas the right-hand side indicates which equilibrium is played in the same area of the graph if the evaluation is public.

compared to

$$E[U_P|U, sep(D,C)] = 2pu_P^+ + (1-p)(u_P^- + u_P^0) - w(p)$$

in the private case; here O and U again indicate public (observable) and private (unobservable) evaluation, respectively. Thus, whenever the separating equilibrium is played in the case of private evaluation, the principal is better off with public evaluation as long as this results in "enough pooling." And this occurs if the biased agent's utility differential is relatively low as stated in Proposition 8.

Proposition 8 *If $p > \eta$, there exists a threshold $\rho(p) \in [w(p)+u_b^+-u_b^0, w(1)+u_b^+-u_b^0]$, such that the principal reveals the evaluation for $\Delta_b < \rho$ and conceals it for $\Delta_b > \rho$. If $p \leq \eta$ and Δ_b is such that the decision is delegated both in case of private and public evaluation, then there exist thresholds $\rho_1(p) \in [w(p), w(\eta)]$ and $\rho_2(p) \in [\rho_1(p), w(1) + u_b^+ - u_b^0]$, such that the principal reveals the evaluation if $\Delta_b < \rho_1(p)$ and conceals it if $\Delta_b > \rho_2(p)$. Furthermore, if $w(\eta) \leq u_b^+ - u_b^0$ and w is "very convex", i.e.*

$$pw(\eta) - \eta w(p) \geq p(1-\eta)(u_P^+ - u_P^0),$$

we have $\rho_1(p) = \rho_2(p)$.

However, if wages are assumed to be linear, the principal pays an expected wage of $w(p)$ to the agent in any case. Hence, the biased agent's increased incentive to imitate the loyal agent in case of public evaluation is no longer costly for the principal.

Corollary 4 *For any prior belief p, the principal weakly prefers to reveal the evaluation to the market if wages are linear.*

2.5 Welfare

In the following, we consider the players' welfare obtained with evaluation and compare the results to the interaction without evaluation. As the results differ depending on whether delegation occurs without evaluation, we split the subsequent discussion in two parts: to begin with, we analyse the case of high priors, $p > \eta$, where we see delegation if no evaluation takes place, and then move on to the case of low priors, $p \leq \eta$, where there is no delegation without evaluation.

2.5.1 The Case of High Priors: $p > \eta$

First, we analyse aggregate welfare, followed by a discussion of comparative statics with regard to utility parameters and the wage schedule. As the main analysis is essentially the same whether evaluation is public or not – as transfers from principal to agent cancel out – the following propositions hold for both regimes unless otherwise stated.

Aggregate Welfare.
As we have seen in Section 2.4, introducing an interim evaluation may be beneficial for the principal, because both a separation of agents and an alignment of the agents' behaviour (in Period 1), which only occur with evaluation, can increase his profit.

Regarding the agent, the appreciation of the evaluation depends on his type. In particular, the loyal agent weakly prefers being evaluated: he always chooses the principal's preferred project in Period 1; and with public evaluation, he may not only gain discretion but also a higher wage in Period 2 (except a pooling equilibrium is played). The biased agent, by contrast, is always better off without evaluation: with evaluation, choosing his preferred project is always punished by centralisation (and a wage drop to zero if the evaluation is public), while without evaluation the decision is always delegated.

Formally, accounting for both players and both types of the agent, expected aggregate welfare with evaluation, W_E, net of expected aggregate welfare with no evaluation, W_N, is given by

$$E[W_E - W_N] = (1-p)\lambda[(u_P^+ - u_P^0) - (u_b^0 - u_b^-)] - (1-p)[(u_b^+ - u_b^0) - (u_P^0 - u_P^-)],$$

with $\lambda = 0$ if $u_b^0 - u_b^- \geq w(1)$ (separating equilibrium) and $\lambda = 1$ if $u_b^0 - u_b^- \leq w(p)$ (pooling equilibrium).

Thus, assuming that the principal's benefits from an optimal project decision outweigh the respective benefits of the biased agent, i.e. assuming

$$u_P^+ - u_P^- =: \Delta_P \geq \Delta_b = u_b^+ - u_b^- \ ,$$

welfare is increased by an evaluation whenever a pooling equilibrium is played. Whether evaluating an agent is efficient in general, though, depends on the relation of the principal's profit from centralisation compared to a biased project choice, $u_P^0 - u_P^-$, to the biased agent's net gain from deciding over the project, $u_b^+ - u_b^0$.

Proposition 9 *If $p > \eta$, the comparison of expected welfare with and without evaluation yields the following results:*

- *If $u_P^0 - u_P^- \geq u_b^+ - u_b^0$, an interim evaluation increases aggregate welfare.*
- *If $u_P^0 - u_P^- < u_b^+ - u_b^0$, there exists a threshold $\rho \in [w(\eta) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0]$ ($\rho = u_b^+ - u_b^0$) if the evaluation is public (private), such that an interim evaluation increases aggregate welfare if and only if $\Delta_b < \rho$.*

The intuition for the first result is straightforward. With evaluation, the principal has the opportunity to centralise in the second period when observing an inappropriate project choice whereas without evaluation the decision is always delegated if the prior is high. If this opportunity to centralise is more valuable to the principal than it harms the biased agent, evaluation increases the aggregate welfare.

Regarding the second result, if centralisation harms the biased agent more than it benefits the principal, welfare is no longer increased if a separating equilibrium is played. Nevertheless, aggregate payoffs are larger if the biased agent imitates the loyal agent in Period 1 – compared to the situation where the principal delegates twice without evaluation. Thus, welfare is still increased if (and only if) the biased agent's utility differential, Δ_b, is low enough as this increases pooling incentives.

A Comparison of Public and Private Evaluation.
Considering each type of equilibrium separately, aggregate welfare is the same whether the evaluation is public or private. However, the highest aggregate welfare is achieved in the pooling equilibrium and it decreases in the probability that the biased agent separates. Furthermore, for any given prior $p > \eta$, the probability that the biased agent separates is at least as high with private evaluation than it is with public evaluation. Thus, Lemma 4 is immediate.

Lemma 4 *If $p > \eta$, the aggregate welfare generated in case of public evaluation is at least as high as if the evaluation is private.*

Welfare Effects of the Wage Schedule.
Another determinant of aggregate welfare is the wage schedule. Of course, total welfare never changes under a simple redistribution of payoffs. Thus, in a usual context, the wage schedule does not influence total welfare. In the present setting with public evaluation, however, the wage schedule, together with $u_b^0 - u_b^-$, determines the equilibrium played in the signalling game which, in turn, influences total welfare. In particular, high wages (weakly) increase the area where a pooling equilibrium is played and (weakly) decrease the area where a separating equilibrium is played, which has a positive effect on aggregate welfare. Hence, focusing on the effect of the wage schedule, paying competitive wages weakly increases welfare when intermediate evaluation takes place if prior beliefs are high.

Proposition 10 *If $p > \eta$, welfare in case of public evaluation is maximised if $w(\mu)$ is maximal within the principal's budget constraints for $\mu \in [p, 1]$.*

Intuitively, increasing wages for a given belief μ strengthens the incentive for the biased agent to mirror the behaviour of the loyal agent in Period 1 (so as to obtain the high wage and the discretion to choose in Period 2). In particular, the higher the wages, the more likely is the biased agent to choose the project which is aligned with the principal's preferences. Accordingly, for higher wages, the region where a pooling equilibrium is played is increased (while the region for a separating equilibrium is reduced). As the pooling equilibrium is welfare enhancing if $u_b^+ - u_b^- = \Delta_b \leq \Delta_P = u_P^+ - u_P^-$, this improves aggregate welfare.

If the evaluation is private, the wage schedule has no influence on aggregate welfare.

2.5.2 The Case of Low Priors: $p \leq \eta$.

Finally, we briefly turn to the case where the prior belief is such that the principal would not delegate without an evaluation, i.e. $p \leq \eta$.

Unfortunately, for such small priors, it is difficult to determine clear-cut welfare results. However, if the principal's benefit from centralisation compared to letting the biased agent decide about the project, $u_P^0 - u_P^-$, is smaller than the biased agent's respective gain when choosing the project, $u_b^+ - u_b^0$, welfare is enhanced with evaluation if Δ_b is large enough.

Still assuming that the principal's benefits from an optimal project decision outweigh the respective benefits of the biased agent, i.e. $u_P^+ - u_P^- \Delta_P \geq \Delta_b = u_b^+ - u_b^-$, we obtain the following result.

Proposition 11 *If $p \leq \eta$ and $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$ if the evaluation is public ($\Delta_b \geq u_b^+ - u_b^0$ in case of private evaluation), total expected welfare is (weakly) increased with evaluation if*

$$u_b^+ - u_b^0 \geq u_P^0 - u_P^-.$$

Note that the intuition for this result is in line with the interpretation in case $p > \eta$. If the prior is low, the principal centralises if no evaluation takes place. By contrast, if the agent is evaluated, he delegates in the first period and thereby induces a benefit of u_b^+ for the biased agent – instead of u_b^0 without evaluation. If this relative gain of the biased agent is larger than the principal's benefit from centralising compared to letting the biased agent decide, aggregate welfare is increased.

Arguing along the same lines as in the case $p > \eta$, we conclude that public evaluation yields a higher aggregate welfare than private evaluation whenever the former regime yields more pooling by the biased agent.

Lemma 5 *If $p \leq \eta$ and $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$, public evaluation generates a higher aggregate welfare than private evaluation.*

3 Extensions

In this section, we briefly consider two natural extensions of our model. First, in Section 3.1, we relax the assumption of equally long/important periods. Finally, Section 3.2 shows that our results are robust to adding more evaluations and intertemporal discounting.

3.1 Flexible Timing of the Evaluation

Instead of presupposing that the principal might evaluate the agent in the middle of his employment period, we now consider a situation where the principal can choose to evaluate the agent after a fraction $\alpha \in [0, 1]$ of his employment period. In particular, we analyse at which point in time it is optimal for the principal to evaluate the agent depending on the observability of the evaluation.

For the sake of argument, we focus on the case of high prior beliefs, i.e. $p > \eta$. Moreover, we define

$$\alpha_b(\mu) := \frac{w(\mu) + (u_b^+ - u_b^0)}{w(\mu) + (u_b^+ - u_b^0) + (u_b^+ - u_b^-)}.$$

Note that α_b is increasing in μ.

3.1.1 Private Evaluation

Consider first the case of private evaluation. In this case, the biased agent chooses x_l in the first period if the importance of the first period is sufficiently low, i.e. $\alpha < \alpha_b(0) =: \alpha_b$. Otherwise, a separating equilibrium is played and the biased agent randomises between strategies if $\alpha = \alpha_b$. Furthermore, the principal's payoff in the separating equilibrium decreases in α, whereas it increases in α for the pooling equilibrium. Thus, Lemma 6 is straightforward.

Lemma 6 *For $p > \eta$, the principal maximises his expected utility by choosing $\alpha = \alpha_b(0) =: \alpha_b$ for all p. His expected payoff in this case, given that the biased agent plays the mixed strategy $(\lambda, 1 - \lambda)$, is then given by*

$$\begin{aligned} E[\Pi_U|\alpha = \alpha_b] &= pu_P^+ + (1-p)\lambda[\alpha_b u_P^+ + (1-\alpha_b)u_P^-] \\ &+ (1-p)(1-\lambda)[\alpha_b u_P^- + (1-\alpha_b)u_P^0] - w(p). \end{aligned}$$

3.1.2 Public Evaluation

Similarly, in case of public evaluation, a pooling equilibrium is played if the weight on the first period is low, i.e. $\alpha < \alpha_b(p)$, and the biased agent separates for $\alpha > \alpha_b(1)$.

However, different from the private case, a mixed equilibrium is played for $\alpha \in [\alpha_b(p), \alpha_b(1)]$. Moreover, assuming $pw(1) - w(p) \leq (1-p)(u_P^0 - u_P^-)$ for all $p \geq \eta$, the same dynamics as in the previous case apply. Yet, the optimal α in case the

mix1(D,C) equilibrium is played cannot be derived without further assumptions on the wage schedule.

Lemma 7 *For $p > \eta$ and $pw(1) - w(p) \leq (1-p)(u_P^0 - u_P^-)$, the principal maximises his expected utility by choosing some $\alpha \in [\alpha_b(p), \alpha_b(1)]$. In this case, the biased player mixes between strategies and the principal's expected payoff is given by*

$$\begin{aligned} E[\Pi_O|\alpha] &= pu_P^+ + (1-p)\lambda[\alpha u_P^+ + (1-\alpha)u_P^-] + (1-p)(1-\lambda)[\alpha u_P^- + (1-\alpha)u_P^0] \\ &- \alpha w(p) - (1-\alpha)[p + (1-p)\lambda]w(\mu_+). \end{aligned}$$

Further,

$$w(\mu_+) = \frac{\alpha}{1-\alpha}(u_b^+ - u_b^-) - (u_b^+ - u_b^0).$$

Thus, the principal decides to evaluate the agent at an earlier stage if the evaluation is private in order to increase the importance of the second period. If the evaluation is public, the second period already provides higher incentives because of wage concerns and the principal evaluates the agent at a later stage, thereby emphasising the first period.[23]

3.2 More Evaluations

Finally, we consider the case where $N-1$ evaluations take place thereby dividing the employment period into N intervals. For reasons of tractability, we focus on the case $p > \eta$. For small prior beliefs, however, the dynamics of the model are the same and similar results apply.

The game we analyse in the following is a repeated game where the principal is uncertain about the type of the agent. Both players are long-lived, as for example in Mailath and Samuelson (2006, ch. 16) or Schmidt (1993). However, our subsequent argument does not strictly follow these lines in the literature but rather gives an inductive argument for what happens if further periods are added.

As we consider a repeated game with incomplete information, the one-deviation principle does not hold. Nonetheless, we can show that for high prior beliefs an equilibrium strategy for the principal is to delegate upon observing the loyal outcome and to centralise up to the final period otherwise – regardless of the observability of the evaluations. This grim-trigger strategy turns out to be the principal's unique subgame perfect strategy consistent with the Intuitive Criterion (Cho and Kreps, 1987).

[23] This result hinges on the assumption that wages are non-negative and not "too convex," i.e. $pw(1) - w(p) \leq (1-p)(u_P^0 - u_P^-)$.

3.2.1 Private Evaluation

In order to reach a solution for the multi-period game, we apply a backward induction argument. If the evaluations are private, the biased agent chooses x_b in the penultimate period if $u_b^0 > u_b^-$, he pools if $u_b^0 < u_b^-$ and he plays a mixed strategy if $u_b^0 = u_b^-$.

The Case $u_b^0 \geq u_b^-$.
If $u_b^0 > u_b^-$, the biased agent chooses the biased project in the penultimate period, $N-1$, and the principal centralises. Thus, if the biased agent chooses x_l in Period $N-2$, the principal's belief is unchanged and the separating equilibrium is played in the penultimate period.[24] However, if he chooses x_b, the principal's belief regarding the agent being loyal immediately drops to zero and, hence, the principal centralises from then on. Thus, choosing x_l instead of x_b, the biased agent receives only u_b^- in Period $N-2$ but avoids getting u_b^0 in Period $N-1$.

In the borderline case, if $u_b^0 = u_b^-$, the biased agent plays a mixed strategy in the penultimate period. However, if we analyse prior periods it turns out that this behaviour cannot be sustained and the biased agent chooses x_b in prior periods, followed by centralisation by the principal (see the Appendix for details).

Lemma 8 shows that these results hold even if we allow for arbitrary discounting by the biased agent, where $\delta \in [0,1]$ denotes the biased agent's discount factor.

Lemma 8 *If $u_b^0 \geq u_b^-$, there is an equilibrium in the finitely repeated game where the biased agent chooses x_b if the decision is delegated to him for all $\delta \in [0,1]$. The loyal agent chooses x_l, and the principal delegates unless he observes x_b, in which case he centralises in all remaining periods.*

The Case $u_b^0 < u_b^-$.
If $u_b^0 < u_b^-$, a pooling equilibrium is played in Period $N-1$ and the principal delegates the decision to the agent. Thus, in Period $N-2$, the biased agent gets u_b^- by choosing x_l until in the last period he chooses his preferred project and gains a profit of u_b^+. On the other hand, if he chooses x_b at some point, he momentarily earns u_b^+ but is stuck with a payoff of u_b^0 for all remaining periods. Lemma 9 shows that it is therefore indeed optimal for the biased agent to pool up to the last period if he is patient enough.

[24] The belief does not change if the biased agent pools because the loyal agent always chooses x_l.

Lemma 9 *If $u_b^0 < u_b^-$ and $\delta > \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0}$, there is an equilibrium in the finitely repeated game where the biased agent chooses x_l in all but the last period if the decision is delegated to him, and x_b in the last period. The loyal agent chooses x_l, and the principal delegates unless he observes x_b, in which case he centralises in all remaining periods.*

3.2.2 Public Evaluation

If the evaluations are public, the equilibria also depend on the wage schedule and the area where a mixed equilibrium is played with one evaluation is not a null-set (as opposed to the case of private evaluation).

The Case $\Delta_b > w(p) + u_b^+ - u_b^0$.
If $\Delta_b \geq w(1) + u_b^+ - u_b^0$, that is, $u_b^0 - u_b^- \geq w(1)$, the biased agent chooses x_b in Period $N-1$ and the principal centralises in Period N, the last period. Moving back to Period $N-2$, an argument similar to the one in the case of private evaluation can be applied. The difference is that by choosing x_l the biased agent not only gains discretion in the following period but also earns a higher wage. However, as wages cannot exceed $u_b^0 - u_b^-$ in the separating equilibrium, this effect is too small to make the biased agent pool in any prior period.

If $\Delta_b \in (w(p) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0)$, the biased agent plays a mixed strategy in Period $N-1$. However, in this case it turns out that he chooses x_b in previous periods and for all discount factors.

Lemma 10 *If $\Delta_b > w(p) + u_b^+ - u_b^0$, there is an equilibrium in the finitely repeated game where the biased agent chooses x_b if the decision is delegated to him for all $\delta \in [0,1]$. The loyal agent chooses x_l, and the principal delegates unless he observes x_b, in which case he centralises in all remaining periods.*

The Case $\Delta_b \leq w(p) + u_b^+ - u_b^0$.
Finally, if Δ_b is low, a pooling equilibrium is played in the penultimate period, $N-1$, and the decision is delegated to the agent. Therefore, the biased agent chooses x_l in any previous period if he is sufficiently patient.

Lemma 11 *If $\Delta_b \leq w(p) + u_b^+ - u_b^0$ and $\delta > \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)}$, there is an equilibrium in the finitely repeated game where the biased agent chooses x_l in all but the last period if the decision is delegated to him, and x_b in the last period. The loyal agent chooses*

x_l, and the principal delegates unless he observes x_b, in which case he centralises in all remaining periods.

To conclude, the pure strategy equilibria observed in the two-period version of the repeated game in case $p > \eta$ extend to an arbitrary finitely repeated game, given that the biased agent is patient enough. The equilibria where the biased agent mixes between strategies however can not be sustained with more periods. In these cases, the biased agent chooses the non-cooperative strategy irrespective of his patience.

4 Concluding Remarks

In this paper, we have analysed a model of delegation between an uninformed principal and an informed but potentially biased agent. In particular, assuming generic utilities and discrete choice options, we have explored the effects of an interim evaluation of the agent compared to a situation where no evaluation takes place.

As we have shown, in the relationship with interim evaluation, concerns for reputation may lead the biased agent to misrepresent his preferences and choose the project which is preferred by the principal at an early stage of the interaction, thereby inducing increased levels of delegation by the principal in the first period. Moreover, we have seen that this effect is particularly strong if the evaluation is observable to the market, in which case the monetary compensation for the agent's outside option provides an additional incentive.

Regarding the principal, it turned out that, somewhat contrary to what one might think at first glance, he may even benefit from a public evaluation compared to a situation without evaluation, if the biased agent's imitating behaviour is not too costly in terms of wages. Thus, although the principal always benefits from a private (and costless) evaluation, revealing the evaluation to the market may even increase the profitability of the evaluation for the principal. This derives from the fact that, while the agent's expected wage may increase in case the evaluation is observable, the biased agent also has a stronger incentive to align his project choice with the principal's preferences which may overcompensate the wage effect.

In addition, we have argued that, from a welfare perspective, an interim evaluation may lead to an increase in aggregate welfare compared to the case without evaluation. Moreover, we have shown that if the agent is evaluated it is always welfare-enhancing to reveal this information to the market (as this increases the biased agent's incentive to imitate the loyal one). And, last but not least, considering the welfare effects of

the wage schedule, we have shown that, if delegation occurs both with and without evaluation, paying competitive wages is optimal. This is due to the fact that increases in welfare are essentially achieved at the expense of the biased agent, and high wages provide the strongest incentives for this type to align his first period behaviour with the principal's preferences.

Regarding possible policy implications, the analysis thus suggests that, in a setting as the one considered here, strengthening the agent's rights in terms of access to the results of his evaluation and wage claims improves welfare. However, while these results may indeed indicate some lines of thought for more general intuitions, we are fully aware that general claims are of course difficult to make based on the stylised theoretical analysis conducted here.

Appendix

Equilibrium Refinement

In this paper, we use the Intuitive Criterion (Cho and Kreps, 1987) as a refinement for equilibria in the signalling game.

In short, the Intuitive Criterion rules out all equilibria which are sustained by unreasonable beliefs off the equilibrium path. In particular, in a pooling equilibrium where both types choose x_b, the loyal agent is the first one to switch to x_l when payoffs are gradually changed. Thus, reasonable off-equilibrium beliefs have to satisfy $\mu_+ = 1$ in this case. But if $\mu_+ = 1$, the principal delegates after observing x_l and the agent's wage after choosing x_l is maximal. Therefore, choosing x_l is strictly dominant for the loyal agent if $\mu_+ = 1$. This leads us to the following Lemma.

Lemma 12 *All pooling equilibria of the signalling game where both types choose x_b are ruled out by the intuitive criterion.*

Accordingly, for a pooling equilibrium where both types choose x_l, the Intuitive Criterion requires $\mu_- = 0$. This requirement is not restrictive for the considered equilibria. See Cho and Kreps (1987) or Mas-Colell et al. (1995, pp. 467ff.) for further details on the Intuitive Criterion.

Proofs of Section 2

Proof of Proposition 1. The principal prefers delegation to centralisation when no evaluation takes place if

$$\begin{aligned} & pu_P^+ + (1-p)u_P^- \geq u_P^0 \\ \Leftrightarrow \quad & p \geq \frac{u_P^0 - u_P^-}{u_P^+ - u_P^-} \end{aligned}$$

■

Proof of Lemma 1. In order to analyse the equilibria of the signalling game in case of private evaluation, we have to conduct an analogous analysis to the one in the Proof of Lemma 2 and Lemma 3. However, the difference with private evaluation is that wages are omitted in the analysis of best replies as they have no strategic impact. Hence, the equilibrium analysis in case of private evaluation yields the same result as assuming $w \equiv 0$ in the statement of Lemma 2 and Lemma 3. Indeed, by applying Lemmata 2 and 3 to $w \equiv 0$, the mix1(D,C) equilibrium reduces to a null-set and the equilibria where the principal centralises in any case vanish. ■

Proof of Lemma 2 and Lemma 3. In order to analyse the equilibria of the signalling game in case of public evaluation, we consider all possible strategies of the principal separately. The pure strategies are analysed in cases 1 to 4, while case 5 includes all possible mixed strategies by the principal. Throughout the proof, we omit the analysis of pooling equilibria where both types choose x_b because of Lemma 12.

Case 1: $s_P = (D, C) \Rightarrow \mu_- \leq \eta$, $\mu_+ \geq \eta$
In this case, the loyal agent chooses x_l if $s_P = (D, C)$ and $\mu_+ \geq \mu_-$ irrespective of the biased agent's strategy.

In a separating equilibrium, we have $s_b = x_b$, $\mu_- = 0$ and $\mu_+ = 1$. Thus, $s_b = x_b$ is a best reply for the biased agent if

$$\begin{aligned} & w(0) + u_b^+ + u_b^0 \geq w(1) + u_b^- + u_b^+ \\ \Leftrightarrow \quad & \Delta_b \geq w(1) + u_b^+ - u_b^0. \end{aligned}$$

In a pooling equilibrium, we have $s_b = x_l$ and $\mu_+ = p$. As $\mu_+ \geq \eta$ is needed for the principal to choose $s_P = (D, C)$, this equilibrium can only exist if $p \geq \eta$.

Furthermore, this equilibrium can only be sustained if

$$\begin{aligned} & w(p) + u_b^- + u_b^+ \geq w(\mu_-) + u_b^+ + u_b^0 \\ \Leftrightarrow \quad & \Delta_b \leq w(p) + u_b^+ - u_b^0, \end{aligned}$$

where we restricted the analysis to rational beliefs and thus $\mu_- = 0$.

In an equilibrium where the biased agent mixes between x_l and x_b with probability λ, $1-\lambda$, respectively, we have $\mu_- = 0$ as the loyal agent always chooses x_l. The biased agent's indifference condition is given by

$$\begin{aligned} & w(\mu_+) + u_b^- + u_b^+ = w(\mu_-) + u_b^+ + u_b^0 \\ \Leftrightarrow \quad & \Delta_b = w(\mu_+) + u_b^+ - u_b^0. \end{aligned}$$

The randomisation parameter λ is then implicitly given by $\mu_+ = \frac{p}{p+(1-p)\lambda}$. Furthermore, if $p \leq \eta$, we always have $\mu_+ \in [\eta, 1]$ or, equivalently, $\lambda \in [0, \frac{p(1-\eta)}{\eta(1-p)}]$. If instead $p \geq \eta$, $\mu_+ \in [p, 1]$ or, equivalently, $\lambda \in [0, 1]$. Thus, there exists some mixed strategy $(\lambda, 1-\lambda)$ by the biased agent fulfilling his indifference condition whenever

- $w(\eta) \leq u_b^0 - u_b^- \leq w(1)$ if $p \leq \eta$ or
- $w(p) \leq u_b^0 - u_b^- \leq w(1)$ if $p \geq \eta$.

Hence, if $s_P = (D, C)$, there are three types of equilibria depending on Δ_b and $u_b^+ - u_b^0$: separating, pooling on x_l and one where the biased agent randomises.

Case 2: $s_P = (D, D) \Rightarrow \mu_- \geq \eta$**,** $\mu_+ \geq \eta$

In this case, we have $\mu_- \geq \eta$ and $\mu_+ \geq \eta$, so there cannot be a separating equilibrium. In a pooling equilibrium with $s_l = s_b = x_l$, rational beliefs are given if $\mu_- = 0$, which is not possible either.

Finally, an equilibrium where one of the types mixes and the other type plays a pure strategy can only occur if the loyal agent mixes (in which case either $\mu_+ = 1$ or $\mu_- = 1$), which occurs if and only if $u_l^+ - u_l^- = w(\mu_-) - w(\mu_+)$ and thus $\mu_- \geq \mu_+$. But in this case the biased agent chooses x_b, which yields $\mu_+ = 1$ and $\mu_- < 1$. Thus, there is no equilibrium where *one* of the types mixes.

Accordingly, the only possible mixed equilibrium in this case is the one where both types mix between their pure strategies. This equilibrium requires

$$u_b^+ - u_b^- = -(u_l^+ - u_l^-) = w(\mu_+) - w(\mu_-),$$

which is not possible as we assumed uniqueness of the preferred project and thus $u_i^+ > u_i^-$ for $i \in \{l, b\}$.

Hence, there is no equilibrium with $s_P = (D, D)$.

Case 3: $s_P = (C, C) \Rightarrow \mu_- \leq \eta$**,** $\mu_+ \leq \eta$

In this case, $\mu_- \leq \eta$ and $\mu_+ \leq \eta$ so there is no separating equilibrium.

Moreover, in any pooling equilibrium with $s_l = s_b = x_l$, we conclude $\mu_+ = p$ and thus $p \leq \eta$ is required. Then, the biased agent chooses x_l if

$$\begin{aligned} & w(p) + u_b^- + u_b^0 \geq w(\mu_-) + u_b^+ + u_b^0 \\ \Leftrightarrow \quad & u_b^+ - u_b^- \leq w(p), \end{aligned}$$

assuming rational beliefs, i.e. $\mu_- = 0$. The loyal type also prefers x_l in this case.

Furthermore, an equilibrium where one of the types mixes and the other type plays a pure strategy can only occur if the biased agent mixes (in which case either $\mu_+ = 0$ or $\mu_- = 0$). The biased agent is indifferent in this case if and only if

$$u_b^+ - u_b^- = w(\mu_+) - w(\mu_-).$$

Thus, $\mu_+ > \mu_-$, $\mu_- = 0$ and $s_l = x_l$. As $\lambda \leq 1$, we have that $\mu_+ \geq p$. On the other hand, $\mu_+ \leq \eta$ such that the principal centralises when observing x_l. Thus, this mixed equilibrium is only possible if $p \leq \eta$, in which case we find a randomisation λ whenever

$$w(p) \leq u_b^+ - u_b^- \leq w(\eta)$$

Finally, an equilibrium where both types play mixed strategies requires

$$u_b^+ - u_b^- = -(u_l^+ - u_l^-) = w(\mu_+) - w(\mu_-),$$

which is again not possible.

Hence, if $p \leq \eta$, there are two types of equilibria in this case: A pooling equilibrium if $u_b^+ - u_b^- \leq w(p)$ and a mixed equilibrium if $w(p) \leq u_b^+ - u_b^- \leq w(\eta)$.

Case 4: $s_P = (C, D) \Rightarrow \mu_- \geq \eta$**,** $\mu_+ \leq \eta$

In this case, $\mu_- \geq \eta$ and $\mu_+ \leq \eta$. Thus, the only possible separating equilibrium is one in which $s_l = x_b$ and $s_b = x_l$ is played. However, if $\mu_- = 1$ and $\mu_+ = 0$, the biased agent strictly prefers x_b. So there is no separating equilibrium in this case. In any pooling equilibrium with $s_l = s_b = x_l$ and reational beliefs, we have $\mu_- = 0$, which is

not possible. Hence, there is no pooling equilibrium.

Furthermore, an equilibrium where the biased agent mixes is only possible if he is indifferent, i.e.

$$\begin{aligned} & w(\mu_-) + 2u_b^+ = w(\mu_+) + u_b^- + u_b^0 \\ \Leftrightarrow \quad & w(\mu_-) - w(\mu_+) = u_b^- + u_b^0 - 2u_b^+ < 0, \end{aligned}$$

which is not possible if $\mu_- \geq \mu_+$. Thus, we can exclude such an equilibrium where the biased agent randomises.

On the other hand, if only the loyal agent mixes, the requirement $\mu_- \geq \eta$ can only be achieved if the biased agent chooses x_l and thus, $\mu_- = 1$. But if $\mu_- = 1$, the biased agent strictly prefers x_b. Hence, there is no equilibrium in case $s_P = (C, D)$ where one of the types mixes.

Finally, an equilibrium where both types mix is only possible if $u_i^+ = u_i^-$ for $i \in \{l, b\}$, which is ruled out by assumption.

Thus, there is no equilibrium if $s_P = (C, D)$.

Case 5: $s_P = \lambda_P D + (1 - \lambda_P)C$

The principal randomises between D and C with probabilities $(\lambda_P, 1 - \lambda_P)$ after observing x_l if and only if $\eta = \mu_+$.

Consider the case where the biased agent plays x_l with probability λ and x_b with probability $1 - \lambda$ and the loyal agent chooses x_l. Then, $\mu_- = 0$ and $s_P(x_b) = C$. The principal randomises in order to make the biased agent indifferent, which occurs if

$$\begin{aligned} & w(\mu_-) + u_b^+ + u_b^0 = w(\mu_+) + \lambda_P(u_b^- + u_b^+) + (1 - \lambda_P)(u_b^- + u_b^0) \\ \Leftrightarrow \quad & \lambda_P = \frac{u_b^+ - u_b^- - w(\eta)}{u_b^+ - u_b^0}, \end{aligned}$$

using $\eta = \mu_+$ and $\mu_- = 0$. Then, $\lambda_P \in [0, 1]$ if and only if $w(\eta) \leq \Delta_b \leq w(\eta) + u_b^+ - u_b^0$. By contrast, the principal is indifferent between C and D if and only if $\mu_+ = \eta$. However, if the biased agent randomises with probabilities $(\lambda, 1 - \lambda)$, μ_+ is given by $\mu_+ = \frac{p}{p+(1-p)\lambda}$, which is equivalent to

$$\lambda = \frac{p(1 - \eta)}{\eta(1 - p)}.$$

Thus, $\lambda \leq 1$ if and only if $p \leq \eta$ and there is no equilibrium for $p > \eta$.

All other strategies by the biased and the loyal agent are ruled out in equilibrium

applying the same arguments as in the proof of Lemma 1.

Hence, in an equilibrium where the principal randomises between C and D, the loyal agent chooses x_l, while the biased agent randomises between x_l and x_b. This equilibrium is possible only if $p \leq \eta$. ■

Proof of Proposition 2. Here, we compare the principal's payoff in case he delegates in Period 1 with the payoff he could guarantee by centralisation, as given in Section 2.4.1.

If $p > \eta$ and the principal delegates in Period 1, the principal delegates in Period 2 upon observing x_l and centralises otherwise. The biased agent plays a mixed strategy $(\lambda, 1-\lambda)$ with $\lambda \in [0,1]$, i.e. also accounting for pure strategies. The principal's payoff if he delegates in Period 1 is given by

$$E[U_P|D] = p[2u_P^+ - w(p)] + (1-p)[\lambda(u_P^+ + u_P^- - w(p)) + (1-\lambda)(u_P^- + u_P^0 - w(p))],$$

where we omit first-period wages. Thus, the principal delegates in Period 1 if

$$\begin{aligned} & E[U_P|D] \geq u_P^0 + pu_P^+ + (1-p)u_P^- - w(p) \\ \Leftrightarrow \quad & [p + (1-p)\lambda](u_P^+ - u_P^0) \geq 0, \end{aligned}$$

which is always the case. Thus, the principal delegates at least for all $p > \eta$.

If $p \leq \eta$ and $\Delta_b \geq u_b^+ - u_b^0$, the biased agent mixes between strategies with $\lambda \leq \frac{1-\eta}{\eta}\frac{p}{1-p}$ and the principal delegates upon observing x_l and centralises otherwise (in case the decision was delegated in Period 1). This also covers the separating equilibrium, in which case $\lambda = 0$. The principal's payoff in case of delegation in Period 1 is then given by

$$\begin{aligned} E[U_P|D] &= p[2u_P^+ - w(p)] + (1-p)[\lambda(u_P^+ + u_P^-) + (1-\lambda)(u_P^- + u_P^0) - w(p)] \\ &= u_P^0 + u_P^- + p(u_P^+ - u_P^-) + \frac{p}{\mu_+}(u_P^+ - u_P^0) - w(p), \end{aligned}$$

using $\lambda = \frac{1-\mu_+}{\mu_+}\frac{p}{1-p}$ and omitting first-period wages.
The principal delegates in Period 1 if

$$\begin{aligned} & E[U_P|D] \geq 2u_P^0 - w(p) \\ \Leftrightarrow \quad & p \geq \frac{\mu_+(u_P^0 - u_P^-)}{\mu_+(u_P^+ - u_P^-) + (u_P^+ - u_P^0)} \in [\eta^2, \frac{\eta}{2-\eta}]. \end{aligned}$$

In particular, if $\Delta_b > u_b^+ - u_b^0$, $\mu_+ = 1$ and the threshold is given by $\frac{\eta}{2-\eta}$.

If $p \leq \eta$ and $\Delta_b < u_b^+ - u_b^0$, both the biased agent and the principal randomise, where the biased agent's strategy $(\lambda, 1-\lambda)$ is given by $\lambda = \frac{1-\eta}{\eta}\frac{p}{1-p}$ and the principal's delegates with probability λ_P upon observing x_l, and centralises if $x = x_b$. The principal's payoff in case of delegation in Period 1 is then given by

$$\begin{aligned} E[U_P|D] &= p[2\lambda_P u_P^+ + (1-\lambda_P)(u_P^+ + u_P^0)] + (1-p)\lambda[\lambda_P(u_P^+ + u_P^-) \\ &+ (1-\lambda_P)(u_P^+ + u_P^0)] + (1-p)(1-\lambda)(u_P^- + u_P^0) - w(p) \\ &= u_P^0 + u_P^- + p\lambda_P(u_P^+ - u_P^-) + \frac{p}{\eta}[(u_P^+ - u_P^-) - \lambda_P(u_p^0 - u_P^-)] - w(p), \end{aligned}$$

again omitting first-period wages. The principal delegates in Period 1 if

$$\begin{aligned} & E[U_P|D] \geq 2u_P^0 - w(p) \\ \Leftrightarrow \quad & p \geq \eta^2. \end{aligned}$$

■

Proof of Proposition 3. As in the Proof of Proposition 2, we compare the principal's payoff in case he delegates in Period 1 with the payoff he could guarantee by centralisation, as given in Section 2.4.1.

If $p > \eta$ and the principal delegates in Period 1, he delegates in Period 2 upon observing x_l and centralises otherwise. In turn, if he centralises in Period 1, he does not update his belief and delegates in Period 2 as $p > \eta$. Thus, delegation in Period 1 is preferred by the principal if

$$\begin{aligned} & p[2u_P^+ - w(\mu_+)] + (1-p)[\lambda(u_P^+ + u_P^- - w(\mu_+)) + (1-\lambda)(u_P^- + u_P^0)] \\ \geq \quad & u_P^0 + pu_P^+ + (1-p)u_P^- - w(p) \\ \Leftrightarrow \quad & w(p) + [p + (1-p)\lambda][u_P^+ - u_P^0 - w(\mu_+)] \geq 0, \end{aligned}$$

which holds for all $\lambda \in [0,1]$ as $w(\mu_+) \leq u_P^+ - u_P^0$. Thus, the principal delegates at least for all $p > \eta$, as in the case without evaluation.

If $p \leq \eta$ and the principal centralises in Period 1, he does not update his prior belief and centralises again in Period 2 as $p \leq \eta$. Now we consider all possible equilibria of the signalling game for the case $p \leq \eta$ and determine the delegation decision of the principal in Period 1.[25]

[25] Here, we neglect the null-set $\{p = \eta\} \cap \{\Delta_b \leq w(\eta) + u_b^+ - u_b^0\}$.

Case 1: $\Delta_b \leq w(\eta)$

In this case, the biased agent chooses x_l or a mixed strategy, while the principal centralises in any case. Thus, the principal delegates in this case if

$$\begin{aligned} & p[u_P^+ + u_P^0 - w(\mu_+)] + (1-p)[\lambda(u_P^+ + u_P^0 - w(\mu_+)) + (1-\lambda)(u_P^- + u_P^0)] \\ \geq \quad & 2u_P^0 - w(p) \\ \Leftrightarrow \quad & \frac{p}{\mu_+}[u_P^+ - u_P^- - w(\mu_+)] \geq u_P^0 - u_P^- - w(p), \end{aligned}$$

where we used

$$\lambda = \frac{p}{1-p}\frac{1-\mu_+}{\mu_+}.$$

As $u_P^+ - u_P^- \geq w(\mu_+)$, it follows that the principal delegates if

$$p \geq \frac{[u_P^0 - u_P^- - w(p)]\mu_+}{u_P^+ - u_P^- - w(\mu_+)} = \frac{[u_P^0 - u_P^- - w(p)]\mu_+}{(u_P^+ - u_P^-)\eta + (u_P^+ - u_P^0) - w(\mu_+)}.$$

The latter equals $\delta_{w,\mu_+}(p)$ because $\max\{\mu_+, \eta\} = \eta$.

In the pooling equilibrium, $\mu_+ = p$, and the principal delegates if

$$p \geq \frac{[u_P^0 - u_P^- - w(p)]p}{u_P^+ - u_P^- - w(p)},$$

which is always the case. In order not to have to distinguish the case of a pooling equilibrium from the other cases, it remains to show that $\delta_{w,\mu}(p) \leq p$ for all $\Delta_b \leq w(p)$, such that we can use the definition of $\delta_{w,\mu}(p)$ with $\mu(\Delta_b) := w^{-1}(\Delta_b)$ for all $\Delta \leq w(\eta)$. Indeed, for $\Delta_b \leq w(p)$ and thus $\mu \leq p$, we have

$$\delta_{w,\mu}(p) = \frac{[u_P^0 - u_P^- - w(p)]\mu}{u_P^+ - u_P^- - w(\mu)} \leq \frac{[u_P^0 - u_P^- - w(p)]p}{u_P^+ - u_P^- - w(p)} \leq p.$$

Case 2: $\Delta_b > w(\eta)$

In this case, three equilibria are possible: sep(D,C), mix1(D,C) or mix2(λ_P,C). In general, the principal delegates if

$$\begin{aligned} & p[\lambda_P(2u_P^+ - w(\mu_+)) + (1-\lambda_P)(u_P^+ + u_P^0 - w(\mu_+))] + (1-p)(1-\lambda)(u_P^- + u_P^0) \\ + \quad & (1-p)\lambda[\lambda_P(u_P^+ + u_P^- - w(\mu_+)) + (1-\lambda_P)(u_P^+ + u_P^0 - w(\mu_+))] \\ \geq \quad & 2u_P^0 - w(p) \\ \Leftrightarrow \quad & p[u_P^+ - u_P^- - w(\mu_+)] + p\lambda_P[\mu_+(u_P^+ - u_P^-) - (u_P^0 - u_P^-)] \geq [u_P^0 - u_P^- - w(p)]\mu_+, \end{aligned}$$

where we used $\lambda = \frac{p}{1-p}\frac{1-\mu_+}{\mu_+}$.
If $\lambda_P \neq 1$, i.e. in the mix2 equilibrium, it has to hold that $\mu_+ = \eta$ in order to make the principal indifferent between his choices. Thus, in this case, $\mu_+(u_P^+ - u_P^-) - (u_P^0 - u_P^-) = 0$ and the principal delegates if

$$p \geq \frac{[u_P^0 - u_P^- - w(p)]\eta}{u_P^+ - u_P^- - w(\eta)} = \frac{[u_P^0 - u_P^- - w(p)]\eta}{(u_P^+ - u_P^-)\eta + (u_P^+ - u_P^0) - w(\eta)} = \delta_{w,\eta}(p)$$

because $\mu_+ = \eta$.
In all other cases, $\lambda_P = 1$ and the principal delegates if

$$p \geq \frac{[u_P^0 - u_P^- - w(p)]\mu_+}{(u_P^+ - u_P^-)\mu_+ + (u_P^+ - u_P^0) - w(\mu_+)} = \delta_{w,\mu_+}(p)$$

because $\mu_+ \geq \eta$ in these cases. Note that, apart from the pooling equilibrium, $\mu(\Delta) = \mu_+$.

Furthermore, as $w(\mu_+) \leq u_P^+ - u_P^0$, we have

$$\begin{aligned} \delta_{w,\mu_+}(p) &\leq \frac{[u_P^0 - u_P^- - w(p)]\mu_+}{(u_P^+ - u_P^-)\max\{\mu_+, \eta\}} \\ &\leq \frac{[u_P^0 - u_P^- - w(p)]\mu_+}{(u_P^+ - u_P^-)\mu_+} = \frac{u_P^0 - u_P^- - w(p)}{u_P^+ - u_P^-} \leq \eta. \end{aligned}$$

Thus, we have shown that the principal delegates more often with public evaluation than if no evaluation takes place. ■

Proof of Proposition 4. We define

$$f(\mu, p) := \delta_{w,\mu}(p) - p.$$

Then, f is continuously differentiable in $p \in [0,1]$ and piecewise continuously differentiable on $\{\mu < \eta\}$ and $\{\mu > \eta\}$. Furthermore, $f(\mu, 0) = \delta_{w,\mu}(0) > 0$ and f is strictly decreasing in p:

$$\frac{\partial f}{\partial p}(\mu, p) = -\frac{\mu w'(p)}{(u_P^+ - u_P^-)\max\{\mu, \eta\} + (u_P^+ - u_P^0) - w(\mu)} - 1 < 0$$

Thus, for every $\mu \geq 0$ there exists a unique $p^*(\mu) > 0$ with

$$f(\mu, p^*(\mu)) = 0.$$

Fix such a point (μ^*, p^*) with $f(\mu^*, p^*) = 0$. Then we know that $\frac{\partial f}{\partial p}(\mu^*, p^*) \neq 0$ for all (μ^*, p^*).
Hence, by the Implicit Function Theorem, $p^*(\mu)$ can be locally represented by a continuously differentiable function. Accordingly, there exists a continuously differentiable function $g : U \to V$ from an environment U of μ^* to an environment V of p^* such that

$$g(\mu) = p^*(\mu) \qquad \forall \mu \in U.$$

Furthermore,

$$\frac{\partial g}{\partial \mu}(\mu^*) = -\left(\frac{\partial f}{\partial p}(\mu^*, p^*)\right)^{-1} \frac{\partial f}{\partial \mu}(\mu^*, p^*).$$

As we have seen, $\frac{\partial f}{\partial p}(\mu^*, p^*) < 0$ for all (μ^*, p^*), thus the sign of $\frac{\partial g}{\partial \mu}(\mu^*)$ is given by the sign of $\frac{\partial f}{\partial \mu}(\mu^*, p^*)$.

Let $\mu^* < \eta$. Then, $p^* = p(\mu^*)$ is given by

$$[u_P^0 - u_P^- - w(p^*)]\mu^* = p^*[u_P^+ - u_P^- - w(\mu^*)].$$

As the right-hand side is strictly positive for $p^* > 0$, we have $u_P^0 - u_P^- - w(p^*) > 0$ for all $p^* > 0$ in this case.

If $\mu^* > \eta$, $p^* = p(\mu^*)$ is given by

$$[u_P^0 - u_P^- - w(p^*)]\mu^* = p^*[(u_P^+ - u_P^-)\mu^* + (u_P^+ - u_P^0) - w(\mu^*)].$$

Applying the same reasoning as before, we conclude $u_P^0 - u_P^- - w(p^*) > 0$ for all $p^* > 0$.

Thus, for $\mu^* < \eta$ we have

$$\frac{\partial f}{\partial \mu}(\mu^*, p^*) = \frac{[u_P^0 - u_P^- - w(p^*)][u_P^+ - u_P^- - w(\mu^*) + \mu^* w'(\mu^*)]}{[u_P^+ - u_P^- - w(\mu^*)]^2}.$$

and for $\mu^* > \eta$

$$\frac{\partial f}{\partial \mu}(\mu^*, p^*) = \frac{[u_P^0 - u_P^- - w(p^*)][u_P^+ - u_P^0 - w(\mu^*) + \mu^* w'(\mu^*)]}{[(u_P^+ - u_P^-)\mu^* + (u_P^+ - u_P^0) - w(\mu^*)]^2}$$

In both cases, the derivative is non-negative and thus

$$\frac{\partial g}{\partial \mu}(\mu^*) > 0$$

for all $\mu^* \neq \eta$. As μ is (weakly) increasing in Δ_b, we conclude that the delegation threshold weakly increases in Δ_b. ■

Proof of Corollary 1. Corollary 1 represents a special case of Proposition 3. Hence, let's assume that the delegation threshold is implicitly defined by

$$p_O^* = \frac{[u_P^0 - u_P^- - w(p_O^*)]\mu}{(u_P^+ - u_P^-)\max\{\mu, \eta\} + (u_P^+ - u_P^0) - w(\mu)}.$$

Using $w(\mu) = \mu W$ for all $\mu \in [0, 1]$, the above expression is equivalent to

$$[(u_P^+ - u_P^-)\max\{\mu, \eta\} + (u_P^+ - u_P^0) - \mu W]p_O^* = (u_P^0 - u_P^-)\mu - p_O^*\mu W$$

and thus

$$p_O^*(\mu) = \frac{(u_P^0 - u_P^-)\mu}{(u_P^+ - u_P^-)\max\{\mu, \eta\} + (u_P^+ - u_P^0)}.$$

Inserting $\mu(\Delta_b)$ as defined in Proposition 3 yields the result. ■

Proof of Proposition 5. Now we compare the delegation decision for public and private evaluation. In order to do so, we determine conditions such that p_O^* lies above or below the step function p_U^* for those Δ_b where p_O^* is constant. In the following considerations, we use that for all $x \in [0, 1]$

$$p_O^*(\Delta_b) \leq x \quad \Leftrightarrow \quad \delta_{w,\mu(\Delta_b)}(x) \leq x.$$

This equivalence is illustrated in Figure 8.

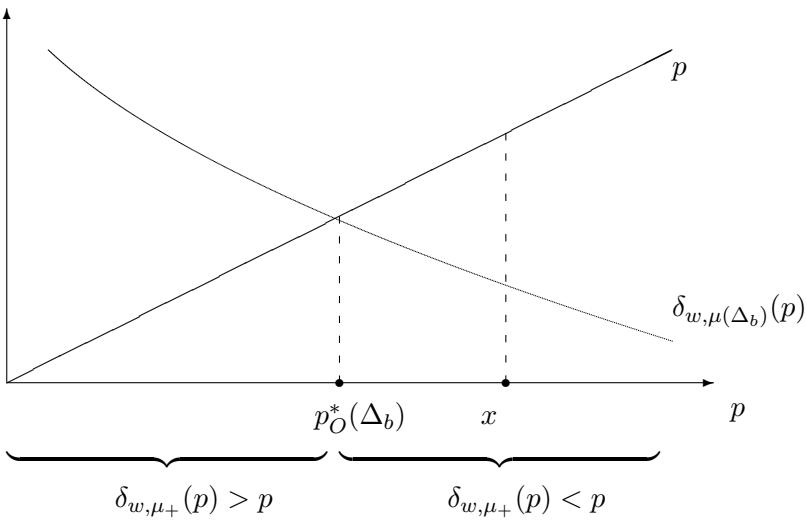

Figure 8: Illustration of the equivalence of $p_O^*(\Delta_b) \leq x$ and $\delta_{w,\mu(\Delta_b)}(x) \leq x$.

Case 1: $\Delta_b \geq w(1) + u_b^+ - u_b^0$

In this case, a separating equilibrium is played, both with public and private evaluation. Delegation is more likely when the evaluation is public if $p_O^*(\Delta_b)$, with $\mu(\Delta_b) = 1$, lies below $\frac{\eta}{2-\eta} = p_U^*(\Delta_b)$.

$$\begin{aligned}
& p_O^*(\Delta_b) \leq \frac{\eta}{2-\eta} \\
\Leftrightarrow \quad & \delta_{w,1}\left(\frac{\eta}{2-\eta}\right) \leq \frac{\eta}{2-\eta} \\
\Leftrightarrow \quad & \frac{u_P^0 - u_P^- - w(\frac{\eta}{2-\eta})}{(u_P^+ - u_P^-) + (u_P^+ - u_P^0) - w(1)} \leq \frac{u_P^0 - u_P^-}{(u_P^+ - u_P^-) + (u_P^+ - u_P^0)} \\
\Leftrightarrow \quad & \frac{\eta}{2-\eta} w(1) - w\left(\frac{\eta}{2-\eta}\right) \leq 0.
\end{aligned}$$

As w is (weakly) convex, the above inequality does not hold and $p_O^*(\Delta_b)$ lies above $\frac{\eta}{2-\eta} = p_U^*(\Delta_b)$ for $\Delta_b \geq w(1) + u_b^+ - u_b^0$. Thus, delegation is weakly more likely with private evaluation. As $p_O^*(\Delta_b)$ is continuous and weakly increasing from zero, this result implies that $p_O^*(\Delta_b)$ cuts $p_U^*(\Delta_b)$ at some point $\rho \leq w(1) + u_b^+ - u_b^0$. Thus, the Proposition follows immediately. In case 2, we describe the relation between $p_O^*(\Delta_b)$ and $p_U^*(\Delta_b)$ in further detail.

Case 2: $w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0$

In this case, $\mu(\Delta_b) = \eta$ if the evaluation is public, and $p_U^*(\Delta_b)$ is given by η^2 for $\Delta_b < u_b^+ - u_b^0$, and by $\frac{\eta}{2-\eta}$ if $\Delta_b > u_b^+ - u_b^0$. Now, three cases are possible: either $p_O^*(\Delta_b)$ lies below or above any of these thresholds, or in between.

Firstly, $p_O^*(\Delta_b) \leq \eta^2$ if and only if

$$\begin{aligned}
& \delta_{w,\eta}(\eta^2) \leq \eta^2 \\
\Leftrightarrow \quad & \frac{[u_P^0 - u_P^- - w(\eta^2)]\eta}{(u_P^+ - u_P^-) - w(\eta)} \leq \frac{(u_P^0 - u_P^-)^2}{(u_P^+ - u_P^-)^2} \\
\Leftrightarrow \quad & \eta w(\eta) - w(\eta^2) \leq 0,
\end{aligned}$$

which is not the case if w is convex.

On the other hand, $p_O^*(\Delta_b) \leq \frac{\eta}{2-\eta}$ if and only if

$$\begin{aligned}
& \delta_{w,\eta}\left(\frac{\eta}{2-\eta}\right) \leq \frac{\eta}{2-\eta} \\
\Leftrightarrow \quad & \frac{[u_P^0 - u_P^- - w(\frac{\eta}{2-\eta})]\eta}{(u_P^+ - u_P^-) - w(\eta)} \leq \frac{u_P^0 - u_P^-}{(u_P^+ - u_P^-) + (u_P^+ - u_P^0)} \\
\Leftrightarrow \quad & \frac{\eta}{2-\eta} w(\eta) - \eta w\left(\frac{\eta}{2-\eta}\right) \leq (1-\eta)\frac{\eta}{2-\eta}(u_P^+ - u_P^0).
\end{aligned}$$

Hence, $p_O^*(\Delta_b)$ lies between η^2 and $\frac{\eta}{2-\eta}$ for $w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0$ if w is "not too convex", and above $\frac{\eta}{2-\eta}$ if w is "very convex". ■

Proof of Corollary 2. In order to prove Corollary 2, we use the proof of Proposition 3.

A linear wage schedule implies that $p_O^*(\Delta_b) = \frac{\eta}{2-\eta} = p_U^*(\Delta_b)$ for $\Delta_b \geq w(1) + u_b^+ - u_b^0$. Furthermore, $p_O^*(\Delta_b) = \eta^2 \leq p_U^*(\Delta_b)$ if $w(\eta) < \Delta_b < w(\eta) + u_b^+ - u_b^0$. Taken together, this implies that $p_O^*(\Delta_b)$ lies weakly below the step function $p_U^*(\Delta_b)$ and thus there is weakly more delegation when the evaluation is public. ■

Proof of Proposition 6. If $p > \eta$ and the principal privately evaluates the agent, he receives a payoff of

$$E[\Pi_U] = 2pu_P^+ + (1-p)\lambda(u_P^+ + u_P^-) + (1-p)(1-\lambda)(u_P^- + u_P^0) - w(p),$$

where $\lambda \in [0,1]$, covering both the pooling and the separating equilibrium.

On the other hand, if $p > \eta$, the principal delegates in Period 1 if there is no evaluation. In this case, his expected payoff is given by

$$E[\Pi_N] = 2[pu_P^+ + (1-p)u_P^-] - w(p),$$

where we omit the first-period wage $w(p)$ in all cases.

Thus, evaluating the agent is preferred whenever

$$\begin{aligned}
& E[\Pi_U] \geq E[\Pi_N] = 2[pu_P^+ + (1-p)u_P^-] - w(p) \\
\Leftrightarrow \quad & (1-p)\lambda(u_P^+ - u_P^0) + (1-p)(1-\lambda)(u_P^0 - u_P^-) \geq 0,
\end{aligned}$$

which is true for all $\lambda \in [0,1]$.

On the other hand, if $p \leq \eta$, the principal does not delegate if no evaluation takes

place. Hence, comparing the principal's payoffs with and without evaluation yields the same analysis as comparing his payoffs after a delegation decision in Period 1 and after centralisation, given that a private evaluation takes place. Hence, the principal prefers to evaluate the agent whenever p lies above the delegation threshold, and he is indifferent if there is no delegation in any case. ■

Proof of Proposition 7. As in the Proof of Proposition 6, if $p > \eta$, the principal's expected payoff without evaluation is given by

$$E[\Pi_N] = 2[pu_P^+ + (1-p)u_P^-] - w(p),$$

omitting the first-period wage.

With public evaluation, the principal's payoff is given by

$$E[\Pi_O] = p[2u_P^+ - w(\mu_+)] + (1-p)\lambda[u_P^+ + u_P^- - w(\mu_+)] + (1-p)(1-\lambda)(u_P^- + u_P^0),$$

where $\lambda \in [0,1]$ (including the pooling and the separating equilibrium) and again omitting first-period wages. Using $\mu_+ = \frac{p}{p+(1-p)\lambda}$, this yields for the net expected payoff

$$E[\Pi_O - \Pi_N] = p\frac{1-\mu_+}{\mu_+}(u_P^+ - u_P^0) + (1-p)(u_P^0 - u_P^-) + w(p) - \frac{p}{\mu_+}w(\mu_+),$$

which is non-negative if and only if

$$pw(\mu_+) - \mu_+ w(p) \leq p(1-\mu_+)(u_P^+ - u_P^0) + (1-p)\mu_+(u_P^0 - u_P^-).$$

The above inequality holds for $\mu_+ = p$. Furthermore, for all $\mu_+ \geq p$, the non-negative left-hand side, $pw(\mu_+) - \mu_+ w(p)$, increases because w is convex, while the right-hand side, $p(1-\mu_+)(u_P^+ - u_P^0) + (1-p)\mu_+(u_P^0 - u_P^-)$, is decreasing in μ_+ as $p > \eta$. Knowing that μ_+ weakly increases in Δ_b, we conclude that there is a threshold $\rho > 0$ (possibly infinity) such that a public evaluation is preferred for all $\Delta_b \leq \rho$.

If $p \leq \eta$, the principal centralises without evaluation which yields an expected payoff of $E[\Pi_N] = 2u_P^0 - w(p)$ (omitting first-period wages). Comparing the principal's expected payoffs in case of evaluation vs. no evaluation yields the same analysis as comparing delegation in Period 1 to centralisation in Period 1 in case of public evaluation. Hence, public evaluation is preferred by the principal whenever p lies above the delegation threshold, and he is indifferent if there is no delegation in any case. ■

Proof of Proposition 8. Let's first assume $p > \eta$. If $\Delta_b \leq w(p) + u_b^+ - u_b^0$, the biased agent pools in case the evaluation is public, whereas any mixed strategy by the biased agent is possible in the private case. Thus, the principle prefers to reveal the evaluation whenever

$$\begin{aligned} & 2pu_P^+ + (1-p)(u_P^+ + u_P^-) - w(p) \\ \geq \; & 2pu_P^+ + (1-p)\lambda(u_P^+ + u_P^-) + (1-p)(1-\lambda)(u_P^- + u_P^0) - w(p), \end{aligned}$$

which is true for all $\lambda \in [0,1]$.

If $w(p) + u_b^+ - u_b^0 \leq \Delta_b$, the biased agent randomises between strategies in case the evaluation is public, with $\lambda \in [0,1]$, covering the pooling and the separating case. On the other hand, a separating equilibrium is played in the private case. Thus, the principle prefers to reveal the evaluation whenever

$$\begin{aligned} & 2pu_P^+ + p\frac{1-\mu_+}{\mu_+}(u_P^+ + u_P^-) + \frac{\mu_+ - p}{\mu_+}(u_P^- + u_P^0) - \frac{p}{\mu_+}w(\mu_+) \\ \geq \; & 2pu_P^+ + (1-p)(u_P^- + u_P^0) - w(p), \end{aligned}$$

using $\lambda = \frac{p}{1-p}\frac{1-\mu_+}{\mu_+}$. This is equivalent to

$$pw(\mu_+) - \mu_+ w(p) \leq p(1-\mu_+)(u_P^+ - u_P^0).$$

Note that the above inequality is fulfilled for $\Delta_b = w(p) + u_b^+ - u_b^0$, where $\mu_+ = p$. On the other hand, it does not hold for $\mu_+ = 1$ because w is convex. As the left-hand side is increasing in $\mu_+ \geq p$, while the right-hand side decreases in μ_+, we conclude that for all $p > \eta$, there exists a threshold $\rho(p) \in [w(p) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0]$ such that the principal prefers to reveal the evaluation for $\Delta_b < \rho(p)$ and to conceal it if $\Delta_b > \rho(p)$.

Now we assume $p \leq \eta$ and consider the case where the decision is delegated both in case of private and public evaluation.

If $\Delta_b \leq w(\eta) \leq u_b^+ - u_b^0$, a mix2($\lambda_P$,C) equilibrium is played in the private case, whereas pool(D,C) or mix1(D,C) occurs with public evaluation. Hence, public evaluation is preferred if

$$E[\Pi_O] = \frac{p}{\mu_+}(u_P^+ - u_P^-) + u_P^0 + u_P^- - \frac{p}{\mu_+}w(\mu_+) \geq \frac{p}{\eta}(u_P^+ - u_P^-) + u_P^0 + u_P^- - w(p) = E[\Pi_U],$$

where $\mu_+ = p$ for $\Delta_b \leq w(p)$ and $\mu_+ = w^{-1}(\Delta_b)$ if $\Delta_b \in [w(p), w(\eta)]$. This is equivalent to

$$pw(\mu_+) - \mu_+ w(p) \leq p\left(1 - \frac{\mu_+}{\eta}\right)(u_P^+ - u_P^-).$$

The above inequality holds for $\mu_+ = p$, but not for $\mu_+ = \eta$. As the left-hand side increases in $\mu_+ \geq p$, while the right-hand side decreases in μ_+, there is a threshold $\rho_1(p) \in [w(p), w(\eta)]$ such that the principal reveals the evaluation for $\Delta_b < \rho_1(p)$. On the other hand, if $\Delta_b \leq u_b^+ - u_b^0 \leq w(\eta)$, we face the same situation. Hence, for the case $u_b^+ - u_b^0 \leq w(\eta)$, we see that public evaluation is preferred at least up to some threshold $\rho_1(p) \in [w(p), w(\eta)]$.

In order to determine $\rho_2(p)$, we distinguish two cases: w being "not too convex" and "very convex".

Case 1: $pw(\eta) - \eta w(p) \leq p(1-\eta)(u_P^+ - u_P^0)$
If $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$, a mix1(D,C) or sep(D,C) equilibrium is played if the evaluation is public, while the biased agent always separates in the private case. Some algebra shows that the principal reveals the evaluation if

$$pw(\mu_+) - \mu_+ w(p) \leq p(1-\mu_+)(u_P^+ - u_P^0)$$

for $\mu_+ \in [\eta, 1]$. While the above inequality never holds if $\mu_+ = 1$, it holds for $\mu_+ = \eta$ given that w is "not too convex". Applying the same reasoning as above, we conclude that there is a threshold $\rho_2(p) \in [w(\eta) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0]$ such that the principal reveals the evaluation for $\Delta_b \in [w(\eta) + u_b^+ - u_b^0, \rho_2(p))$ and conceals it for $\Delta_b > \rho_2(p)$.

Case 2: $pw(\eta) - \eta w(p) \geq p(1-\eta)(u_P^+ - u_P^0)$
From case 1 we conclude that the principal prefers to conceal the evaluation for all $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$ if w is "very convex". Furthermore, if $\Delta_b \in [\max\left\{w(\eta), u_b^+ - u_b^0\right\}, w(\eta) + u_b^+ - u_b^0]$, both the principal and the biased agent randomise between strategies if the evaluation is public, while a separating equilibrium is played in the private case. Comparing the equilibria, it turns out that the principal prefers to conceal the evaluation if and only if w is "very convex", i.e. $pw(\eta) - \eta w(p) \geq p(1-\eta)(u_P^+ - u_P^0)$.

If $u_b^+ - u_b^0 \leq \Delta_b \leq w(\eta)$, a pool(C,C) or mix1(C,C) equilibrium is played in the public case, while the biased agent separates if the evaluation is private. Here, the

principal prefers to reveal the evaluation if

$$pw(\mu_+) - \mu_+ w(p) \leq p[1 - \mu_+(2-\eta)](u_P^+ - u_P^-)$$

with $\mu_+ \in [\max\left\{w^{-1}(u_b^+ - u_b^0), p\right\}, \eta]$. The above inequality does not hold for $\mu_+ = \eta$, as w is "very convex", and it holds for $\mu_+ = p$ because $p \leq \eta < \frac{1}{2-\eta}$. Applying the same arguments as before, there exists a threshold $\rho_2(p) \in [\rho_1(p), w(\eta)]$ such that the principal conceals the evaluation for $\Delta_b > \rho_2(p)$.

On the other hand, if $w(\eta) \leq \Delta_b \leq u_b^+ - u_b^0$, a mix2($\lambda_P$,C) equilibrium is played in both cases and the principal prefers to conceal the evaluation as wages are convex. Together with the result on $\rho_1(p)$, we can conclude that $\rho_2(p) = \rho_1(p)$ in this case and the principal conceals the evaluation if and only if $\Delta_b > \rho_2(p) = \rho_1(p) \in [w(p), w(\eta)]$. ■

Proof of Corollary 4. If wages are linear, the expected wage if the evaluation is public equals the one in the private case. Hence, replacing $pw(\mu_+) - \mu_+ w(p) = 0$ in the Proof of Proposition 8 yields that the principal prefers to reveal the evaluation to the market for any of the considered cases. ■

Proof of Proposition 9. If $p > \eta$, expected aggregate welfare without evaluation is given by

$$E[W_N] = 2p(u_P^+ + u_l^+) + 2(1-p)(u_b^+ + u_P^-).$$

When evaluation takes place and $p > \eta$ (regardless of whether it is public or private), the principal's strategy is $s_P = (D, C)$ in all equilibria and the loyal agent always chooses $s_l = x_l$. The biased agent's strategy is given by $(\lambda, 1-\lambda)$ with $\lambda \in [0,1]$. Thus, we can calculate the expected aggregate welfare with evaluation,

$$\begin{aligned} E[W_E] &= 2p(u_P^+ + u_l^+) + (1-p)\lambda[(u_P^+ + u_P^-) + (u_b^- + u_b^+)] \\ &+ (1-p)(1-\lambda)[(u_P^- + u_P^0) + (u_b^+ + u_b^0)]. \end{aligned}$$

Accordingly, expected net aggregate welfare is given by

$$E[W_E - W_N] = (1-p)\lambda[(u_P^+ - u_P^0) - (u_b^0 - u_b^-)] + (1-p)[(u_P^0 - u_P^-) - (u_b^+ - u_b^0)].$$

If $u_P^0 - u_P^- \geq u_b^+ - u_b^0$, the second summand is non-negative. Hence, if $\lambda = 0$ (separating equilibrium), we have $E[W_E - W_N] \geq 0$. Second, whenever $\lambda \in (0,1)$, it has to hold that $u_b^0 - u_b^- \leq w(1) \leq u_P^+ - u_P^0$ in the public case, or $u_b^0 - u_b^- = 0$ if the evaluation is private. Thus, in this case, the first summand is also non-negative which

again yields $E[W_E - W_N] \geq 0$. Finally, if $\lambda = 1$ (pooling equilibrium), $E[W_E - W_N] \geq 0$ if and only if $u_P^+ - u_P^- \geq u_b^+ - u_b^-$, which holds by assumption.

On the other hand, if $u_P^0 - u_P^- \leq u_b^+ - u_b^0$, aggregate welfare is weakly increased if and only if

$$\lambda[(u_P^+ - u_P^0) - (u_b^0 - u_b^-)] \geq (u_b^+ - u_b^0) - (u_P^0 - u_P^-) \geq 0.$$

Obviously, in a separating equilibrium ($\lambda = 0$), this inequality does not hold, while for $\lambda = 1$, it is fulfilled as $u_P^+ - u_P^- \geq u_b^+ - u_b^-$. In intermediate cases, we know that $u_P^+ - u_P^0 \geq u_b^0 - u_b^-$ and thus welfare is weakly increased if and only if

$$\lambda \geq \frac{(u_b^+ - u_b^0) - (u_P^0 - u_P^-)}{(u_P^+ - u_P^0) - (u_b^0 - u_b^-)} \in [0, 1].$$

Hence, there exists a threshold $\rho \in [w(\eta) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0]$ if the evaluation is public, or $\rho = u_b^+ - u_b^0$ in the private case, such that welfare is increased for $\Delta_b < \rho$ and decreased for $\Delta_b > \rho$. ■

Proof of Lemma 4. If $p > \eta$, the derivative of the expected aggregate welfare in case of evaluation with regard to λ is given by

$$\frac{\partial E[W_E]}{\partial \lambda} = (1 - p)[(u_P^+ - u_P^0) - (u_b^0 - u_b^-)].$$

If the biased agent does not separate in the public case, we have $u_b^0 - u_b^- \leq w(1) \leq u_P^+ - u_P^0$, and thus

$$\frac{\partial E[W_E]}{\partial \lambda} \geq 0$$

for all $\lambda > 0$.

Furthermore, in this case we have that the randomisation factor λ is always at least as large in the public case as in the private case, for a given (p, Δ_b). Thus, public evaluation yields a (weakly) higher aggregate welfare.

If the biased agent separates in the public case, he also does so with private evaluation and aggregate welfare is the same in both cases. ■

Proof of Proposition 10.

As $E[W_N]$ is constant in λ, we conclude from the Proof of Lemma 4 that

$$\frac{\partial E[W_E - W_N]}{\partial \lambda} = \frac{\partial E[W_E]}{\partial \lambda} \geq 0.$$

As the wage schedule influences the played equilibrium, increasing $w(p)$ increases the region for values of Δ_b where a pooling equilibrium is played if the evaluation is public. In turn, rising $w(1)$ reduces the region for values of Δ_b where a separating equilibrium is played. Furthermore, if $w(p) + u_b^+ - u_b^0 < \Delta_b < w(1) + u_b^+ - u_b^0$, λ decreases in μ_+, where $\mu_+ = w^{-1}(\Delta_b - (u_b^+ - u_b^0))$. Thus, for a given $\Delta_b \in (w(p) + u_b^+ - u_b^0, w(1) + u_b^+ - u_b^0)$, welfare is maximal if $w(\mu_+(\Delta_b))$ is maximal. Hence, high wages for $\mu \in [p, 1]$ (weakly) increase expected welfare in case of public evaluation. ■

Proof of Proposition 11. If $p \leq \eta$, expected aggregate welfare without evaluation is given by

$$E[W_N] = 2[u_P^0 + pu_l^0 + (1-p)u_b^0].$$

If $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$ if the evaluation is public, or $\Delta_b \geq u_b^+ - u_b^0$ in the private case, either a mix1(D,C) or a sep(D,C) equilibrium is played. Using the result for $E[W_E]$ from Proposition 9 and $p + (1-p)\lambda = \frac{p}{\mu_+}$, we conclude that the net expected welfare is given by

$$\begin{aligned} E[W_E - W_N] &= [(u_b^+ - u_b^0) - (u_P^0 - u_P^-)] + \frac{p}{\mu_+}[(u_P^+ - u_P^0) - (u_b^0 - u_b^-)] \\ &+ p[2(u_l^+ - u_l^0) + (u_b^0 - u_b^-) + (u_P^+ - u_P^-) - (u_b^+ - u_b^0)], \end{aligned}$$

where $\mu_+ \in (\eta, 1]$. We assume throughout the proof that

$$u_b^+ - u_b^0 \geq u_P^0 - u_P^-. \tag{1}$$

Condition 1 ensures that the first of the three summands in $E[W_E - W_N]$ is non-negative.
If $\mu_+ < 1$, we have $u_b^0 - u_b^- = 0$ in the private case and $u_b^0 - u_b^- \leq w(1) \leq u_P^+ - u_P^0$ if the evaluation is public. Hence, also the second summand is non-negative in this case. Finally, we have $u_b^+ - u_b^0 \leq \Delta_b < u_b^+ - u_b^- \leq u_P^+ - u_P^-$, which yields non-negativity for the third summand.
If $\mu_+ = 1$, net expected welfare is given by

$$E[W_E - W_N] = [(u_b^+ - u_b^0) - (u_P^0 - u_P^-)] + p[2(u_l^+ - u_l^0) + (u_P^+ - u_P^-) + (u_P^+ - u_P^0) - (u_b^+ - u_b^0)].$$

The second summand is non-negative because $u_b^+ - u_b^0 \leq u_P^+ - u_P^-$ as above. ■

Proof of Lemma 5. If $p \leq \eta$ and $\Delta_b \geq w(\eta) + u_b^+ - u_b^0$, the mix1(D,C) or the sep(D,C) equilibrium is played in case of public evaluation, whereas the separating

equilibrium is played if the evaluation is private. Following the same argument as in the Proof of Lemma 4, it is immediate that the aggregate welfare is weakly increased by revealing the evaluation. ■

Proofs of Section 3

Proof of Lemma 6. If the evaluation is private and $p > \eta$, the equilibrium analysis is analogous to Lemma 1. Here, we only briefly discuss the resulting equilibria.
In all equilibria, the principal delegates after observing x_l and centralises otherwise. The biased agent then strictly prefers x_l if

$$\begin{aligned} & \alpha u_b^- + (1-\alpha)u_b^+ > \alpha u_b^+ + (1-\alpha)u_b^0 \\ \Leftrightarrow \quad & \alpha < \frac{u_b^+ - u_b^0}{(u_b^+ - u_b^0) + (u_b^+ - u_b^-)} = \alpha_b(0) =: \alpha_b. \end{aligned}$$

Otherwise, he plays a separating equilibrium, and he mixes between strategies in case of equality.

The principal's payoff in the pooling equilibrium is given by

$$E[\Pi_U | \alpha < \alpha_b] = pu_P^+ + (1-p)[\alpha u_P^+ + (1-\alpha)u_P^-] - w(p),$$

which increases in α.

On the other hand, the principal's payoff in the separating equilibrium is given by

$$E[\Pi_U | \alpha > \alpha_b] = pu_P^+ + (1-p)[\alpha u_P^- + (1-\alpha)u_P^0] - w(p),$$

which decreases in α. Thus, it is optimal for the principal to choose $\alpha = \alpha_b$, in which case his payoff is given by

$$\begin{aligned} E[\Pi_U | \alpha = \alpha_b] &= pu_P^+ + (1-p)\lambda[\alpha_b u_P^+ + (1-\alpha_b)u_P^-] \\ &+ (1-p)(1-\lambda)[\alpha_b u_P^- + (1-\alpha_b)u_P^0] - w(p). \end{aligned}$$

■

Proof of Lemma 7. If the evaluation is public and $p > \eta$, the equilibrium analysis is analogous to Lemma 2. Here, we only briefly discuss the resulting equilibria.
In all equilibria, the principal delegates after observing x_l and centralises otherwise.

In a pooling equilibrium, for the biased agent to strictly prefer x_l, it has to hold that

$$\begin{aligned} & \alpha u_b^- + (1-\alpha)u_b^+ + (1-\alpha)w(p) > \alpha u_b^+ + (1-\alpha)u_b^0 \\ \Leftrightarrow \quad & \alpha < \frac{w(p) + (u_b^+ - u_b^0)}{w(p) + (u_b^+ - u_b^0) + (u_b^+ - u_b^-)} = \alpha_b(p). \end{aligned}$$

On the other hand, a separating equilibrium is played if

$$\begin{aligned} & \alpha u_b^- + (1-\alpha)u_b^+ + (1-\alpha)w(1) < \alpha u_b^+ + (1-\alpha)u_b^0 \\ \Leftrightarrow \quad & \alpha > \frac{w(1) + (u_b^+ - u_b^0)}{w(1) + (u_b^+ - u_b^0) + (u_b^+ - u_b^-)} = \alpha_b(1). \end{aligned}$$

Finally, the biased agent randomises between strategies if

$$\begin{aligned} & \alpha u_b^- + (1-\alpha)u_b^+ + (1-\alpha)w(\mu_+) = \alpha u_b^+ + (1-\alpha)u_b^0 \\ \Leftrightarrow \quad & \alpha = \frac{w(\mu_+) + (u_b^+ - u_b^0)}{w(\mu_+) + (u_b^+ - u_b^0) + (u_b^+ - u_b^-)} = \alpha_b(\mu_+) \end{aligned} \tag{2}$$

and $\mu_+ \in [p, 1]$ if and only if $\alpha \in [\alpha_b(p), \alpha_b(1)]$.

The principal's payoff in the pooling equilibrium is given by

$$E[\Pi_O | \alpha < \alpha_b(p)] = pu_P^+ + (1-p)[\alpha u_P^+ + (1-\alpha)u_P^-] - w(p),$$

which increases in α.

On the other hand, the principal's payoff in the separating equilibrium is given by

$$E[\Pi_O | \alpha > \alpha_b(1)] = pu_P^+ + (1-p)[\alpha u_P^- + (1-\alpha)u_P^0] - \alpha w(p) - (1-\alpha)pw(1),$$

which decreases in α if $pw(1) - w(p) \leq (1-p)(u_P^0 - u_P^-)$. Thus, in this case it is optimal for the principal to choose $\alpha \in [\alpha_b(p), \alpha_b(1)]$, in which case his payoff is given by

$$\begin{aligned} E[\Pi_O|\alpha] &= pu_P^+ + (1-p)\lambda[\alpha u_P^+ + (1-\alpha)u_P^-] + (1-p)(1-\lambda)[\alpha u_P^- + (1-\alpha)u_P^0] \\ &- \alpha w(p) - (1-\alpha)[p + (1-p)\lambda]w(\mu_+). \end{aligned}$$

Furthermore, from (2) we conclude that

$$w(\mu_+) = \frac{\alpha}{1-\alpha}(u_b^+ - u_b^-) - (u_b^+ - u_b^0).$$

■

Proof of Lemma 8.
If $u_b^0 > u_b^-$, the biased agent chooses x_b in the penultimate period. He chooses x_b again in the period before if

$$\begin{aligned} & u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \geq u_b^- + \delta u_b^+ + \delta^2 u_b^0 \\ \Leftrightarrow \quad & \delta \leq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0}, \end{aligned}$$

which is always the case, as $u_b^0 > u_b^-$ and thus $\frac{u_b^+ - u_b^-}{u_b^+ - u_b^0} > 1$. So the biased agent chooses x_b again in period $N-2$. By induction, we see that the biased agent chooses x_b in period $N-n$, given that he has chosen x_b in period $N-(n-1)$ if

$$\begin{aligned} & u_b^+ + \underbrace{\sum_{k=1}^{n} \delta^k u_b^0}_{\text{C after } x_b} \geq u_b^- + \underbrace{\delta u_b^+ + \sum_{k=2}^{n} \delta^k u_b^0}_{x_b \text{ in } N-(n-1)\text{, then C}} \\ \Leftrightarrow \quad & \delta \leq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0}. \end{aligned}$$

If $u_b^0 = u_b^-$, the biased agent plays a mixed strategy in the penultimate period. He chooses x_b in the period before if

$$\begin{aligned} & u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \geq u_b^- + \delta[\lambda(u_b^- + \delta u_b^+) + (1-\lambda)(u_b^+ + \delta u_b^0)] \\ \Leftrightarrow \quad & \delta(1 - \lambda + \delta\lambda)(u_b^+ - u_b^-) \leq u_b^+ - u_b^- \\ \Leftrightarrow \quad & \delta \leq 1, \end{aligned}$$

where we used $u_b^- = u_b^0$. For all prior periods, the same argument applies as in case $u_b^0 > u_b^-$.
Furthermore, if the principal observes x_l, he knows that the agent is loyal and thus his best reply is to delegate the decision again. However, if he observes x_b, the agent is biased with certainty and the principal maximises his future payoff by centralising until the last period. ■

Proof of Lemma 9.
If $u_b^0 < u_b^-$, the biased agent chooses x_l in the penultimate period. If we add a period,

the biased agent again chooses x_l if and only if

$$\begin{aligned} & u_b^- + \delta u_b^- + \delta^2 u_b^+ \geq u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \\ \Leftrightarrow \quad & (1+\delta)(u_b^- - u_b^0) \geq (1-\delta^2)(u_b^+ - u_b^0) \\ \Leftrightarrow \quad & \delta \geq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0}. \end{aligned}$$

Now $\frac{u_b^+ - u_b^-}{u_b^+ - u_b^0} < 1$ as $u_b^0 < u_b^-$ and the biased agent pools if δ is large enough. By induction, the biased agent pools in period $N-n$ given that he chooses x_l in period $N-(n-1)$ if and only if

$$\begin{aligned} & u_b^- + \underbrace{\sum_{k=1}^{n-1} \delta^k u_b^- + \delta^n u_b^+}_{x_l \text{ until last period}} \geq u_b^+ + \underbrace{\sum_{k=1}^{n} \delta^k u_b^0}_{\text{C after } x_b} \\ \Leftrightarrow \quad & \frac{1-\delta^n}{1-\delta}(u_b^- - u_b^0) \geq (1-\delta^n)(u_b^+ - u_b^-) \\ \Leftrightarrow \quad & \delta \geq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0}, \end{aligned}$$

which gives us the same constraint as before.
Moreover, if the principal observes x_l, he does not update his prior belief. But as we assume $p \geq \eta$, his best reply is to delegate the decision. If he observed the off-equilibrium outcome x_b, the rational beliefs assumption requests that the principal infers that the agent is biased. Thus, his best reply is to centralise in all future periods if he observed x_b. ■

Proof of Lemma 10.
If $\Delta_b \geq w(1) + u_b^+ - u_b^0$, that is, $u_b^0 - u_b^- \geq w(1)$, the biased agent chooses x_b in the penultimate period. He chooses x_b again in the period before if

$$\begin{aligned} & u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \geq u_b^- + \delta(u_b^+ + w(p)) + \delta^2 u_b^0 \\ \Leftrightarrow \quad & \delta \leq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)}, \end{aligned}$$

which is always the case, as $w(p) < w(1) \leq u_b^0 - u_b^-$ in the separating equilibrium and thus $\frac{u_b^+ - u_b^-}{u_b^+ - u_b^0} > 1$. So the biased agent chooses x_b again in period $t = N-2$ – recall that Period N is the last one. By induction, we see that the biased agent chooses x_b

in period $N-n$, given that he has chosen x_b in period $N-(n-1)$ if

$$\begin{aligned} & u_b^+ + \underbrace{\sum_{k=1}^{n} \delta^k u_b^0}_{\text{C after } x_b} \geq u_b^- + \underbrace{\delta(u_b^+ + w(p)) + \sum_{k=2}^{n} \delta^k u_b^0}_{x_b \text{ in } T-(n-1)\text{, then C}} \\ \Leftrightarrow \quad & \delta \leq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)}. \end{aligned}$$

If $\Delta_b \in (w(p) + u_b^+ - u_b^-, w(1) + u_b^+ - u_b^-)$, the biased agent plays a mixed strategy in the penultimate period. He chooses x_b before if

$$\begin{aligned} & u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \geq u_b^- + \delta[w(p) + \lambda(u_b^- + \delta(u_b^+ + w(\mu_+))) + (1-\lambda)(u_b^+ + \delta u_b^0)] \\ \Leftrightarrow \quad & \delta[w(p) + \delta\lambda w(\mu_+)] \leq (1+\delta\lambda)[(u_b^+ - u_b^-) - \delta(u_b^+ - u_b^0)]. \end{aligned}$$

The latter inequality holds because $w(p) \leq w(\mu_+) = u_b^0 - u_b^-$ and thus

$$\begin{aligned} \delta[w(p) + \delta\lambda w(\mu_+)] &\leq \delta(1+\delta\lambda)w(\mu_+) = \delta(1+\delta\lambda)(u_b^0 - u_b^-) \\ &= \delta(1+\delta\lambda)[(u_b^+ - u_b^-) - (u_b^+ - u_b^0)] \\ &\leq (1+\delta\lambda)[(u_b^+ - u_b^-) - \delta(u_b^+ - u_b^0)]. \end{aligned}$$

With regard to the principal's equilibrium strategy, the same argument as in the proof of Lemma 8 applies and the grim-trigger strategy is the principal's unique subgame perfect best reply. ■

Proof of Lemma 11.
If $\Delta_b \leq w(p) + u_b^+ - u_b^0$, the biased agent chooses x_l in the penultimate period. If we add a period, the biased agent again chooses x_l if and only if

$$\begin{aligned} & u_b^- + \delta(u_b^- + w(p)) + \delta^2(u_b^+ + w(p)) \geq u_b^+ + \delta u_b^0 + \delta^2 u_b^0 \\ \Leftrightarrow \quad & \delta(1+\delta)w(p) \geq (1-\delta^2)(u_b^+ - u_b^0) + (1+\delta)(u_b^0 - u_b^-) \\ \Leftrightarrow \quad & \delta w(p) \geq (1-\delta)(u_b^+ - u_b^0) + (u_b^0 - u_b^-) \\ \Leftrightarrow \quad & \delta \geq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)}. \end{aligned}$$

Now $\frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)} \leq 1$ as $w(p) \geq u_b^0 - u_b^-$ and the biased agent pools if δ is large enough. By induction, the biased agent pools in Period $N-n$ given that he chooses x_l in

Period $N-(n-1)$ if and only if

$$
\begin{aligned}
& u_b^- + \underbrace{\sum_{k=1}^{n-1} \delta^k (u_b^- + w(p)) + \delta^n (u_b^+ + w(p))}_{x_l \text{ until last period}} \geq u_b^+ + \underbrace{\sum_{k=1}^{n} \delta^k u_b^0}_{\text{C after } x_b} \\
\Leftrightarrow \quad & \frac{1-\delta^n}{1-\delta}(u_b^- - u_b^0 + w(p)) \geq (1-\delta^n)(u_b^+ - u_b^0 + w(p)) \\
\Leftrightarrow \quad & \delta \geq \frac{u_b^+ - u_b^-}{u_b^+ - u_b^0 + w(p)},
\end{aligned}
$$

which gives us the same constraint as before.

With the same argument as in Lemma 9, the grim-trigger strategy is the principal's unique subgame perfect best reply consistent with the Intuitive Criterion. ■

References

Acemoglu, D., Pischke, J.-S., 1999. The Structure of Wages and Investment in General Training. *Journal of Political Economy* 107, pp. 539-572.

Alonso, R., Matouschek, N., 2007. Rational Delegation. *RAND Journal of Economics* 38, pp. 1070-1089.

Alonso, R., Matouschek, N., 2008. Optimal Delegation. *Review of Economic Studies* 75, pp. 259-293.

Aghion, P., Tirole, J. 1997. Formal and Real Authority in Organizations. *Journal of Political Economy* 105, pp. 1-29.

Aghion, P., Rey, P., Dewatripont, M., 2004. Transferable Control. *Journal of the European Economic Association* 2, pp. 115-138.

Armstrong, M., 1995. Delegation and Discretion. Munich Personal RePEc Archive: http://mpra.ub.uni-muenchen.de/17069/ [22.3.2013]

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

Benabou, R., Laroque, G., 1992. Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility. *Quarterly Journal of Economics* 107, pp. 921-958.

Bester, H., Krähmer, D., 2008. Delegation and Incentives. *RAND Journal of Economics* 39, pp. 664-682.

Blanes i Vidal, J., 2007. Delegation of Decision Rights and the Winner's Curse. *Economics Letters* 94, pp. 163.169.

Cho, I.-K., Kreps, D., 1987. Signalling Games and Stable Equilibria. *Quarterly Journal of Economics* 102, pp. 179-222.

Cleveland, J.N., Murphy, K.R., Williams, R.E., 1989. Multiple Uses of Performance Appraisal: Prevalence and Correlates. *Journal of Applied Psychology* 74, pp. 130-135.

Crawford, V., Sobel, J., 1982. Strategic Information Transmission. *Econometrica* 50, pp. 1431-1451.

Cripps, M.W., Mailath, G.J., Samuelson, L., 2004. Imperfect Monitoring and Impermanent Reputations. *Econometrica* 72, pp. 407-432.

Cripps, M.W., Mailath, G.J., Samuelson, L., 2007. Disappearing Private Reputations in Long-Run Relationships. *Journal of Economic Theory* 134, pp. 287-316.

Englmaier, F., Filipi, A., Singh, R., 2010. Incentives, Reputation and the Allocation of Authority. *Journal of Economic Behavior & Organization* 76, pp. 413-427.

Fuchs, W., 2007. Contracting with Repeated Moral Hazard and Private Evaluations. *The American Economic Review* 97, pp. 1432-1448.

Fudenberg, D., Levine, D.K., 1989. Reputation and Equilibrium Selection in Games with a Patient Player. *Econometrica* 57, pp. 759-778.

Fudenberg, D., Levine, D.K., 1992. Maintaining a Reputation when Strategies Are Imperfectly Observed. *The Review of Economic Studies* 59, pp. 561-579.

Frankel, A., 2010. Aligned Delegation. NYU Working Paper: http://www. econ.as .nyu.edu/docs/IO/18572/Frankel_20110126.pdf [22.3.2013]

Goltsman, M., Mukherjee, A., 2011. Interim Performance Feedback in Multistage Tournaments: The Optimality of Partial Disclosure. *Journal of Labour Economics* 29, pp. 229-265.

Holmström, B., 1977. *On Incentives and Control in Organizations.* Ph.D. Dissertation, Graduate School of Business, Stanford University.

Holmström, B., 1984. On the Theory of Delegation. In M. Boyer and R. Kihlstrom, eds., *Bayesian Models in Economic Theory.* New York: North-Holland.

Holmström, B., 1999. Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies* 66, pp. 169-182.

Jensen, M., 1986. Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review* 76, pp. 323–329.

Jensen, M., Meckling, W., 1992. Specific and General Knowledge and Organizational Structure. In L. Werin AND H. Wijkander, eds., *Contract Economics.* Oxford: Blackwell.

Kovác, E., Krähmer, D., 2012. Optimal Sequential Delegation. University of Bonn Working Paper: http://www.wiwi.uni-bonn.de/kraehmer/2012-07-05_DK_SeqDel.pdf [22.3.2013]

Kreps, D.M., Wilson, R., 1982. Reputation and Imperfect Information. *Journal of Economic Theory* 27, pp. 253-279.

MacLeod, W.B., Malcomson, J.M., 1989. Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment. *Econometrica* 57, pp. 447-480.

Mailath, G.J., Samuelson, L., 2001. Who Wants a Good Reputation? *The Review of Economic Studies* 68, pp. 415-441.

Mailath, G.J., Samuelson, L., 2006. *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

Mas-Colell, A., Whinston, M., Green, J., 1995. *Microeconomic Theory.* Oxford University Press.

Melumad, N.D., Shibano, T., 1991. Communication in Settings with No Transfers. *The RAND Journal of Economics* 22, pp. 173-198.

Morris, S., 2001. Political Correctness. *Journal of Political Economy* 109, pp. 231-265.

Milgrom, P., Roberts, J., 1982. Predation, Reputation, and Entry Deterrence. *Journal of Economic Theory* 27, pp. 280-312.

Morris, S., 2001. Political Correctness. *Journal of Political Economy* 109, pp. 231-265.

Mukherjee, A., 2008. Career Concerns, Matching, and Optimal Disclosure Policy. *International Economic Review* 49, pp. 1211-1250.

Ottaviani, M., Sørensen, P., 2006a. Reputational Cheap Talk. *RAND Journal of Economics* 37, pp. 155-175.

Ottaviani, M., Sørensen, P., 2006b. Professional Advice. *Journal of Economic Theory* 126, pp 120-142.

Prat, A., 2005. The Wrong Kind of Transparency. *American Economic Review* 95, pp. 862-877.

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

Schmidt, K.M., 1993. Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests. *Econometrica* 61, pp. 325-351.

Sobel, J., 1985. A Theory of Credibility. *Review of Economic Studies* 52, pp. 557-573.

Spence, M., 1973. Job Market Signaling. *The Quarterly Journal of Economics* 87, pp. 355-374.

Zabojnik, J., 2011. Subjective Evaluations with Performance Feedback. Queen's Economics Department Working Paper No. 1283: http://qed.econ.queensu.ca/working_papers/papers/qed_wp_1283.pdf [22.3.2013]