

# Biased Performance Evaluation in A Model of Career Concerns: Incentives versus Ex-Post Efficiency

Sergey Stepanov\*

March 2, 2015

## Abstract

I study a career concerns model in which the principal obtains information about the agent's performance from an intermediary (evaluator), whose interests diverge from those of the principal. I show that, while the evaluator's bias leads to ex-post suboptimal decisions regarding the agent (e.g., inefficient promotion or dismissal), it incentivizes the agent to exert more effort. As a result, generally, a non-zero bias is optimal. The optimal bias is "anti-agent" ("pro-agent") when the agent is of high value (low value) for the principal from the ex-ante perspective. The magnitude of the optimal bias is increasing in the strength of the agent's career concerns and decreasing in the degree of uncertainty about the agent's ability. I also obtain that delegating decision rights to the evaluator may be preferred to communication when a sufficiently large bias is required to create incentives. I discuss applications of my results to promotion policies in organizations, evaluation of government programs and evaluation of CEOs by boards of directors.

JEL classification: D82, D83, M51

Keywords: career concerns, performance evaluation, information transmission, biased intermediary, strategic delegation

## 1 Introduction

In the traditional career concerns framework, the performance of the agent is directly observed by the principal or the market.<sup>1</sup> The agent benefits from the higher performance through a higher evaluation of his ability, which leads to a greater wage (Holmström, 1999; Dewatripont, Jewitt, and Tirole, 1999a,b), greater chances of promotion or retention (Gibbons and Murphy,

---

\*Higher School of Economics, Moscow. Email: sstepanov@hse.ru. I am very grateful to Mikhail Drugov, Guido Friebel, Sergei Guriev, Sergei Izmalkov, Tymofiy Mylovanov, Anders Olofsgård, Sergei Severinov, Giancarlo Spagnolo, seminar participants at Higher School of Economics (Moscow), Goethe University (Frankfurt am Main), Paris School of Economics, Stockholm Institute of Transition Economics, participants of the 2014 North American Summer Meeting of the Econometric Society, and of the 29th European Economic Association Annual Congress for comments.

<sup>1</sup>More generally, by those who take a decision regarding the agent, be they the agent's current boss, potential employers, or voters, depending on the specific setting.

1992; Chevalier and Ellison, 1999) or higher reelection probability (Persson and Tabellini, 2000, ch. 4.5), depending on the specific setup. In this way, career concerns create incentives for the agent to exert effort.

Yet, in many real-life circumstances, information about the agent's performance is transmitted to the principal through intermediaries who do not necessarily have the same objectives as those of the principal. For example, when deciding on a promotion of an employee, the CEO of a company would rely, at least in part, on the evaluation of the employee's ability by the employee's peers or immediate boss. Because such people may be natural competitors of the employee inside the firm, their objectives may differ from the one of the CEO who wants to implement the optimal allocation of people within the firm. Another example is evaluation of a government program by an ad hoc committee (in this case, the agent is the bureaucrat implementing the program). Depending on the political stances of the committee members, its aggregate preferences may be more or less biased with respect to the principal's preferences, be it the government or the society at large. Yet another example is evaluation of a politician by a media outlet, the objectives of which may differ from the ones of the society.

This paper introduces a biased intermediary (evaluator) in the traditional career concerns framework. Biased evaluation is similar (although not equivalent, as I show) to delegation of decision-making to a biased intermediary. In line with many works on strategic delegation, my paper finds that delegating evaluation to a party whose ex-post preferences deviate from those of the principal can benefit the latter from the ex-ante perspective: a biased evaluator serves as a commitment device to follow a certain ex-post policy regarding the agent in order to provide incentives to him. Thus, my paper is closely related to both career concerns and strategic delegation literatures. Yet, to my knowledge, no paper has studied strategic delegation of evaluation in a career concerns setting.

I derive several results absent in the previous literature and provide implications for a number of real-life settings. In particular I show that:

- the optimal evaluator can be tougher as well as softer on the agent relative to the principal's ex-post preferences; the direction of the bias depends on how valuable the agent is for the principal from the ex-ante perspective,
- the optimal evaluator's bias is increasing in the strength of career concerns for the agent,
- the optimal evaluator's bias is decreasing in the ex-ante uncertainty about the agent's ability.

I discuss implications of these results for evaluation and promotion policies in organizations, evaluation of government programs, and evaluation of CEOs by boards of directors. I also examine when delegating decision-making to the evaluator can dominate acting on information communicated by the evaluator.

In my model, the principal wants to maximize the agent's output over the two periods (with a discount factor applied to the second period). The first period output depends both on the

agent’s unobservable effort and his ability, unknown to anyone. The output is non-contractible. After the first period, the evaluator (but not the principal!) observes the output and makes a report to the principal. The principal forms a posterior belief about the agent’s ability and takes a binary decision, regarding the agent: either “favorable” or “unfavorable”. Examples of a favorable decision could be “retention”, “promotion”, “continuation of the agent’s project”, “allocation of a greater budget”. Correspondingly, “dismissal”, “no promotion”, “termination of the agent’s project”, “no increase in the budget” are examples of an unfavorable decision.

The agent receives a private benefit from the favorable decision – this is the source of his incentives to exert effort.

The second period output depends on the agent’s ability if the decision is favorable, but does not depend on it in the case of the unfavorable decision<sup>2</sup>. Hence, the principal’s decision will be described by a threshold, such that the decision is favorable (unfavorable) if the principal’s posterior belief about the agent’s ability is above (below) the threshold.

The principal’s posterior is determined both by the prior and the message he receives from the evaluator. The evaluator’s ex-post preferences can be biased with respect to the principal’s ones. I say that the bias is “anti-agent” when the evaluator prefers to implement the unfavorable decision more often (for a larger set of beliefs about the agent’s ability) than is optimal from the principal’s perspective. The opposite situation corresponds to a “pro-agent” bias. Thus, in my model, the bias can be described by the difference between the evaluator’s threshold on the (perceived) agent’s ability, above which she would prefer the favorable decision, and the principal’s one.

I assume that evaluators’ messages are cheap talk. In my simple two-actions setting, equilibrium communication will be equivalent to a binary advice: “favorable”/“unfavorable”. A bias in the evaluator’s objectives creates incentives for her to distort information about the agent’s performance. However, unless the bias is too large, the evaluator’s message will be decision-relevant for the principal in equilibrium, despite the fact that following the evaluator’s advice leads to systematic errors in the direction of the bias.

Suboptimal ex-post decisions arising due to these errors are an obvious cost of a biased evaluator. From the ex-post perspective, thus, the bias should be zero. At the same time, the *bias can improve the agent’s incentives* to exert effort in the first period. The evaluator’s preferences effectively set the bar on performance that the agent needs to surpass in order to be treated favorably. The agent is discouraged from working both when it is too easy and when it is too hard to clear the bar. Hence, his effort is non-monotonic in this bar and, consequently, in the evaluator’s threshold on the agent’s ability (because the performance bar is monotonic in the evaluator’s preferences in equilibrium).

The crucial thing is that the effort-maximizing evaluator’s threshold is generally different

---

<sup>2</sup>This assumption fits well the example of “retention versus dismissal”. In other cases it may seem too extreme, but a weaker and realistic assumption that the organization’s output depends less on the agent’s performance in the case of the unfavorable decision would suffice.

from the principals' ex-post optimal threshold. These two thresholds are completely independent: the ex-post optimal threshold is not affected by the agent's effort, while the effort-maximizing threshold does not depend on the principal's ex-post preferences. This means that a non-zero evaluator's bias is generally needed if the principal wants to maximize the agent's incentives. This, in turn, implies that the *ex-ante optimal bias is also non-zero*, albeit smaller in magnitude than the effort-maximizing one, because it trades-off ex-post efficiency against effort provision.

Furthermore, *the optimal bias can be anti-agent as well as pro-agent*. In papers on strategic delegation, the optimal bias is usually in one direction: e.g., the central bank has to be more inflation-averse than the society (Rogoff, 1985; Persson and Tabellini, 1993), the manager needs to be more sales-oriented than the firms' owners in a Cournot oligopoly setting (Vickers, 1985; Fershtman, 1985; Fershtman and Judd, 1987; Sklivas, 1987), a regulatory agency has to be less protectionist than the government (Ludema and Olofsgård, 2008).

In my model, the direction of the optimal bias is determined by the value, *for the principal*, of the unfavorable decision relative to the *ex-ante* value of the favorable one, or, to put it another way, by the *ex-ante* value of the agent for the principal. It turns out that *the more valuable the agent is ex-ante, the more likely it is that the optimal bias is "anti-agent" rather than "pro-agent"*. A low value of the unfavorable decision means that the agent is very valuable for the principal. Then, the principal is too "lenient" to the agent ex-post, meaning that he finds it ex-post optimal to take the unfavorable decision only if his posterior about the agent's ability is very low. For example, if the decision is about firing versus retaining the agent, then, if there is no good alternative to the agent, the principal will fire him only if the agent is believed to be really bad. Such "leniency" results in too weak incentives for the agent. Therefore, the optimal evaluator's bias needs to be anti-agent in order to induce more effort. In contrast, a too high value of the unfavorable decision makes the principal too demanding, which also discourages effort. In this case, the optimal bias has to be pro-agent in order to restore incentives.

Next, I analyze the effect of the strength of career concerns, by which I mean the value of the agent's discount factor (importance of the future) and/or the magnitude of the benefits he obtains from the favorable decision. An increase in either of these parameters leads to a greater marginal effect of the bias on the agent's effort. Since the optimal bias is always smaller than the effort-maximizing one, there is always under-provision of effort in the optimum. Consequently, an increase in the marginal effect of the bias on effort calls for an increase in the bias in order to induce more effort. The implication of this result is that *more career-concerned agents should be evaluated by more biased evaluators*. Notice that this statement holds regardless of the direction of the optimal bias.

Finally, I examine the effect of the ex-ante uncertainty about the agent's ability, modeled as the variance of the prior ability distribution. Under the assumptions of the normal distribution and quadratic cost of effort, the conclusion is that *higher ex-ante uncertainty about the agent's*

*ability is likely to reduce the optimal bias.* The key thing is that a higher variance of the ability generally reduces the sensitivity of the agent's equilibrium effort to the bias, while increasing or just mildly decreasing the effect of the bias on the expected loss from ex-post inefficient decisions. That is, the effectiveness of the bias in providing incentives diminishes, whereas the ex-post efficiency remains sufficiently sensitive to the bias. Therefore, the bias needs to be lowered.

I also consider delegating decision-making to the evaluator as an alternative to communication. When the bias is not too large, the two modes of decision-making are equivalent in my binary model, because, in the case of communication, the principal simply follows the evaluator's preferred policy. However, *when the bias is so large* that the principal ignores the evaluator's message, *delegation can be better than communication*, despite its large ex-post inefficiency. The reason is that decision-irrelevant communication does not induce any effort, while delegation still generates some incentives even when the bias is high. This result contrasts with Dessein (2002) who showed, in a traditional Crawford-Sobel type of framework<sup>3</sup>, that delegation is preferred to communication only when the divergence of preferences between the principal and the expert is small enough. The difference in the results is because, in my model, the mode of decision making has an effect on the third party's incentives (the agent), which affect the principal's welfare.

Moreover, I show that *delegation may dominate communication even when the bias is chosen optimally*. This happens when inducing incentives is important but the maximum bias that still ensures decision-relevant communication results in too low effort.

A caveat is in order here: any conclusions about superiority of delegation in my model rely on the assumption that the principal can actually commit to delegation even when ex-post he would want to deprive the delegatee of the decision-making power.

The rest of the paper is organized as follows. Section 2 discusses related literature. In section 3 I set up the model. Section 4 presents the solution. Section 5 is devoted to the comparative statics. In section 6 I compare communication with delegation. Section 7 discusses real-life applications. Section 8 concludes the paper.

## 2 Related literature

### 2.1 Strategic delegation and commitment to ex-post inefficient actions

Optimality of biased evaluation in my model stems from the principal's commitment (or time-consistency) problem. Academic literature has offered various ways of mitigating time inconsistency; my paper is closely related to two of them: strategic delegation of decision-making to a party with different ex-post preferences and distorting or limiting access to information about the agent's performance. In the industrial organization literature, it has been shown that

---

<sup>3</sup>Crawford and Sobel (1982)

owners of a firm may benefit from creating an incentive scheme which would make the manager more sales-oriented than the owners themselves in a Cournot oligopoly setting (Vickers, 1985; Fershtman, 1985; Fershtman and Judd, 1987; Sklivas, 1987) and less sales-oriented in Bertrand oligopoly setting (Fershtman and Judd, 1987; Sklivas, 1987). Such mechanisms allow the firm to commit to ex-ante profit-maximizing strategies in the presence of strategic interactions with rivals.<sup>4</sup>

In the political economy literature, Besley and Coate (2001) show that it may be optimal, from the median voter perspective, to elect a representative whose preferences will differ from those of the median voter, in order to neutralize the effect of lobbies on policy choices ex-post. In Persson and Tabellini (1994), in order to commit to avoid excessive taxation of capital, the majority elects a policymaker with a lower preference for taxing capital (higher capital income) than that of the median voter.<sup>5</sup>

In the monetary policy literature, a well known solution to the time-consistency problem is to delegate monetary policy to a conservative (inflation-averse) central banker (Rogoff, 1985; Persson and Tabellini, 1993).

Similarly, in Ludema and Olofsgård (2008), a government finds it optimal to delegate either protection of a firm or collecting information about the cost of protection to an agency who is less protectionist than government. Such delegation improves the investment policy of the firm. Similarly to the section 6 of my paper, section Ludema and Olofsgård compare delegation of decision-making to simply communication of decision-relevant information by the agency (this is actually the focus of their work); I am discussing the relation of my work to their analysis in section 6.

Many papers in the principal-agent literature, e.g., on financial contracting, share the feature that the principal's commitment to ex-post inefficient actions can induce better agent's incentives (e.g., Dewatripont and Maskin, 1995; Dewatripont and Tirole, 1994, 1996; Berglöf and Von Thadden, 1994; Bolton and Scharfstein, 1996; Hart and Moore, 1998; Crémer, 1995). While my paper also makes use of this idea, its application to the evaluation of career-motivated agents is novel. Among the mentioned papers, Crémer (1995) is probably closest to mine. In a setup with explicit incentives, Crémer (1995) argues that it can be optimal for the principal to commit to stay uninformed about the causes of poor agent's performance, because it creates commitment not to renegotiate with the agent. This, in turn, raises the agent's effort. Essentially, in Crémer's paper, garbling information about the agent's performance helps to create incentives. This is a feature of my model too (biased evaluation generates information garbling).

Dewatripont et al. (1999a) also notice that garbling of a performance measure (in a career concerns setup) may result in a higher agent's effort. However, in contrast to my paper, they

---

<sup>4</sup>See also Zhang and Zhang (1997), who add an R&D dimension into the setting. Fershtman et al. (1991) derive benefits of delegation in a more general game-theoretic setup. Sengul et al. (2012) provide a review of the strategic delegation literature in IO.

<sup>5</sup>See Persson and Tabellini (2002) for a few other examples of strategic delegation in political economics and corresponding references.

focus on the negative effect of information garbling and apply it to show deficiency of creating “fuzzy missions” for an agent.

Gehlbach and Simpser (2015), in a political economy setting, show that garbling information about the ruler’s popularity through electoral manipulation helps the ruler to encourage bureaucrats to exert more effort, because bureaucrats lose incentives when they believe that the ruler’s hold on power is weak.

## 2.2 Other related literature

Ivanov (2010) and Ambrus et al. (2013) derive optimality of having a biased intermediary in information transmission in a different context. They consider traditional sender-receiver cheap talk games, in which the sender’s preferences diverge from those of the receiver. There is no issue of effort provision; the receiver is just concerned with obtaining as precise information as possible, because this information is decision-relevant for him. Both papers show that intermediated communication can dominate direct one if the intermediary’s preferences are biased from the receiver’s ones in the direction opposite to the sender’s bias. In those papers the intermediary’s bias is good because it improves information transmission, whereas in my model it is good precisely because it distorts information transmission.

Finally, in the context of applying my analysis to the internal labor market, I compare my work with Friebel and Raith (2004) and Friebel and Raith (2013), see subsection 7.3 for details. A talented subordinate (employee) may be a threat to a division manager, who is afraid of being replaced. Friebel and Raith (2004) argue that restricting direct communication between the CEO (principal) and the employee (i.e., bypassing the division manager) can be efficiency improving because then the division manager is less afraid to recruit or develop talented subordinates. In my model, evaluation of an employee by his division manager, rather than the CEO or an unbiased colleague, may benefit the organization through improving the *employee’s* incentives.

Friebel and Raith (2013) compare two promotion systems: “silo”, in which junior employees can be promoted only within their initial divisions, and “lattice”, in which promotions across divisions are allowed. In that paper, a well trained employee benefits the division manager through a higher productivity of the division. Then, a silo has a positive effect on the division manager’s incentives to train employees, because in a silo employees cannot be moved to another division. In my model, a silo in which the employee is evaluated by the immediate boss can be good thanks to a positive effect on the *employee’s* incentives.

## 3 Model

I will first present a formal model and then discuss its assumptions and applicability to various real life settings.

### 3.1 Players and timing

There are three players in the model: the principal ( $P$ ), the agent ( $A$ ) and the evaluator ( $E$ ).  $A$  has ability  $\theta \sim F(\cdot)$ , with the density function  $f(\cdot)$  having full support and differentiable everywhere. A priori  $\theta$  is unknown to anyone, while  $F(\cdot)$  is common knowledge.

There are two periods. The sequence of actions in the first period is as follows:

- $A$  exerts effort  $e$  at cost  $c(e)$ ,  $c'(0) = 0$ ,  $c'(\cdot) > 0$  for any  $e > 0$ ,  $c''(\cdot) > 0$  for any  $e$
- $A$ 's first period performance is realized:  $y(\theta, e) = \theta + e$ .
- $E$  (but not  $P$ ) observes  $y$  and makes a report  $r \in \mathbb{R}$  to  $P$ .
- $P$  takes a binary decision regarding  $A$ : "favorable" or "unfavorable".
- The organization's first period output is realized:  $\Pi_1 = y$ .

In the second period no actions are taken, only the second period output  $\Pi_2$  is realized:

$$\Pi_2 = \begin{cases} \theta & \text{if } P\text{'s decision was favorable} \\ z & \text{if } P\text{'s decision was unfavorable} \end{cases}$$

### 3.2 Payoffs

$P$ 's payoff is  $\Pi_1$  in period 1 and  $\Pi_2$  in period 2. Hence, the principal's welfare is identical to the organization's payoff.  $P$  has a discount factor  $\delta$  between the periods. Thus, the ex-ante  $P$ 's welfare is

$$W \equiv \mathbf{E}\Pi_1 + \delta\mathbf{E}\Pi_2$$

I assume that the agent cares only about  $P$ 's decision being favorable and the cost of effort. Specifically,  $A$ 's payoff is  $-c(e)$  in period 1, and in period 2 he gets

$$\begin{cases} B & \text{if } P\text{'s decision was favorable} \\ 0 & \text{if } P\text{'s decision was unfavorable} \end{cases}$$

$B$  is the exogenous private benefit the agent derives from the favorable decision. Depending on the context, it can stem from prestige, reputation (which may translate into higher wealth in the future), possibility to extract more rents at a higher position or from a larger project, etc.

The agent's discount factor is  $\delta_A$ , so his ex-ante welfare is

$$-c(e) + \delta_A B \cdot I(\text{favorable decision}),$$

where  $I(\cdot)$  is the indicator function.

Finally,  $E$ 's second payoff is as follows. Her first period payoff is irrelevant because she takes no decisions then, so I just ignore it. Her second period payoff is:

$$\begin{cases} \theta & \text{if } P\text{'s decision was favorable} \\ z + b & \text{if } P\text{'s decision was unfavorable} \end{cases}$$

Thus,  $E$ 's ex-post preferences are similar to those of  $P$  except being biased; the bias is parametrized by  $b$ . If  $b > 0$ , we will say that the bias is "anti-agent", because the unfavorable decision is relatively more attractive for  $E$  than for  $P$ . Correspondingly, if  $b < 0$ , the bias is "pro-agent".

Let us introduce the following notation:

$$\begin{aligned} \theta_P &\equiv z \\ \theta_E &\equiv z + b = \theta_P + b \end{aligned}$$

That is,  $\theta_P$  is the value of  $\theta$  for which  $P$  is indifferent between the favorable and unfavorable decision, and  $\theta_E$  is the similar threshold for  $E$ . Thus, the bias can be measured simply as the difference between  $\theta_E$  and  $\theta_P$ .

### 3.3 Discussion of the setup

#### 3.3.1 Basic features

This model fits a variety of real life settings. The crucial ingredients are:

- i. The principal cares about both the agent's effort and the agent's ability.
- ii. The principal makes a decision regarding the agent based on the past evaluation of the agent's performance.
- iii. The principal's future welfare is more sensitive to the agent's ability if the principal takes a favorable decision regarding the agent.
- iv. The agent receives a private benefit from the favorable decision.
- v. The principal does not evaluate the agent's performance directly. Instead, the agent is evaluated indirectly by another agent (intermediary) whose preferences are biased with respect to those of the principal.
- vi. Possibilities to write contracts on measures of performance (including evaluator's reports) are limited.

**Features (i) and (ii)** obviously fit many real life setups. Even if a principal's decisions regarding an agent are not always binary, they are often discrete. Depending on a specific

setting, examples of a favorable decision could be “retention”, “promotion”, “continuation of the agent’s project”, “allocation of a greater budget”. Correspondingly, “dismissal”, “no promotion”, “termination of the agent’s project”, “no increase in the budget” would be examples of an unfavorable decision.

**Feature (iii)** is realistic too. In the model, feature (iii) takes an extreme form: if the decision is unfavorable, the second period output does not depend on the agent’s ability at all. This fits well a situation in which the principle decides whether to fire or to retain the agent. However, for the results of my model, it would be enough to assume that the second period output just depends *less* on the agent’s ability in the case of the unfavorable decision. This would be true in many settings: think, for example, of keeping the agent at a low position instead of promotion (higher rank employees are more influential), cutting the budget for the agent’s project, or allocating a less important project to him. In that case, the principal’s and evaluator’s ex-post optimal policies would still have the threshold structure of the same type: “take the favorable decision if and only if the estimated ability is above a certain threshold”.

Consider now **feature (iv)**. While agents sometimes receive monetary rewards from favorable decisions (e.g., promotions normally imply salary increases), they are also likely to derive various pecuniary and non-pecuniary private benefits, which are non-contractible. Prestige, status, high reputation raise an individual’s utility both per se and because they help to increase the individual’s future wealth. Moreover a favorable decision may directly imply greater control over resources, which allows extracting higher rents.

**Feature (v)**, intermediated evaluation, is a feature of many organizations. Principals are busy with many tasks, thus they often lack time to collect information on all their agents’ performance. They may also lack skills to collect and analyze this information. Finally, principals sometimes consist of many individuals rather than a single person (think of such principals as “society” or “shareholders”), in which case there exists a coordination problem. Therefore, it is natural for a principal to delegate evaluation of an agent to parties who have better access to or ability and expertise to collect and interpret information about the agent’s performance. Peers evaluating an employee in an organization, boards of directors evaluating CEOs on behalf of shareholders, special committees evaluating the progress of a government program are examples of such intermediaries.

In many instances, these intermediaries would not have the same preferences as the principal does. Sometimes the divergence of preferences would arise naturally: e.g., peers of an employee may be his/her competitors for a promotion, in which case their preferred decision regarding the employee would naturally differ from the principal’s one. In other cases the bias could be purposefully created by selecting evaluators with certain preferences. For example, a CEO can be evaluated by boards with greater or lower proportion of insiders, a bureaucrat in charge of a government reform can be evaluated by a committee with greater or lower proportion of reform members, a university’s tenure committee may contain members which different views

on tenure standards. The bias can also be affected by reputational concerns or gender or race prejudices.

The important implicit assumption of my model is that the principal does not have access to sources of information other than the evaluator or that such sources are very costly. In particular, once the evaluator has been chosen and the agent has exerted effort, the principal cannot use another evaluator. Without this assumption, the principal would fall into a commitment problem. However, I believe that this assumption is realistic: in many instances, it is difficult or costly to find an alternative evaluator.

Let me also remark that introducing a *costly* alternative source of information would not be a problem for my model. Provided that the evaluator’s bias is not too large, the principal would abstain from spending resources on the alternative information channel. Thus, such a modification would just reduce the range of “credible biases” in equilibrium, but would not change the qualitative results.

The final basic feature of the setup is lack of contractibility (**feature (vi)**). I assume that performance  $y$  and both periods outputs  $\Pi_1$  and  $\Pi_2$  are *non-verifiable*. I also assume that  $E$ ’s report  $m$  is non-contractible. Finally, I assume that  $P$ ’s decision (equivalently,  $A$ ’s private benefit) is non-contractible either. The complete absence of contractibility is not necessary for my results and is driven by the desire to simplify the model by keeping only one channel of incentives provision: biased evaluation. However, some sufficient degree of contractual imperfection is needed in the model. For example, I could allow for some contractible signal about  $y$ , but this signal would have to be noisy enough. In the Appendix I consider an extension in which I allow for contracts stipulating payments from  $P$  to  $A$  contingent on  $P$ ’s decision (favorable/unfavorable). Although such payments help to provide incentives, they are costly to  $P$ . As a result, it remains optimal to have a biased evaluator.

### 3.3.2 Other assumptions

It is crucial that  $P$  needs to take his decision before  $\Pi_1$  is observed, otherwise he could infer  $A$ ’s ability in equilibrium before taking a decision, and  $E$  would play no role then. I also implicitly assume that  $A$ ’s participation constraint is satisfied (e.g., his reservation utility is zero), and  $A$  has no money to pay to  $P$ .

The setup was silent about the agent being able or unable to send his own message to the principal. In fact, in my setup,  $A$ ’s messages would have no effect, because he always strictly prefers the favorable decision, and, hence, any cheap talk between him and  $P$  would be uninformative. So, ignoring this possibility is innocuous.

I also make a few technical assumptions.

**Assumption 1** Distribution  $F(\cdot)$  is unimodal

This assumption is made for simplicity.

**Assumption 2** Functions  $\delta_A Bf(a - x)$  and  $c'(x)$  cross only once on  $x \in [0, \infty)$ , for any  $a$ .

This assumption ensures that  $A$ 's effort is *uniquely* determined by the first-order condition, as will be clear in subsection 4.2.

As can be easily predicted, if  $E$ 's bias is too large,  $P$  will always ignore  $E$ 's messages. Such a bias cannot be optimal for it will neither result in decision-relevant communication nor create any incentives for  $A$ . Hence, for given other parameters, the set of biases which can potentially be optimal will be limited by some  $b_{\min}$  and  $b_{\max}$ . For simplicity, I make an assumption that ensures the interior solution for the optimal bias:

**Assumption 3**  $\mathbf{E}(\theta \mid \theta < Mo(\theta)) < \theta_P \leq \mathbf{E}(\theta \mid \theta > Mo(\theta))$ , where  $Mo(\theta)$  is the mode of  $F(\cdot)$ .

The meaning of this assumption is that  $\theta_P$  should not be far from the mode in order to guarantee that the optimal bias will stay moderate and, hence, will not hit the bounds. Why it is important to compare  $\theta_P$  with the mode will be clear in subsections 4.2 and 4.3.

## 4 Solution

### 4.1 Principal's decision and communication

Let us assume that in the case of indifference, both  $P$  and  $E$  prefer the favorable decision.<sup>6</sup>

The principal's optimal policy is to take the favorable decision iff

$$\mathbf{E}(\theta \mid r) \geq z \equiv \theta_P \tag{1}$$

From  $E$ 's standpoint, the favorable decision should be taken iff

$$\theta \geq z + b \equiv \theta_E = \theta_P + b \tag{2}$$

For given  $E$ 's belief  $\hat{e}$  about  $A$ 's choice of effort, the observation of performance  $y$  allows  $E$  to infer  $\theta$ :

$$\hat{\theta} = y - \hat{e}$$

Suppose  $E$  simply announces whether  $\hat{\theta} < \theta_E$  or  $\hat{\theta} \geq \theta_E$  (below I will show that this is indeed the case in any equilibrium with decision-relevant communication). Such communication will be decision-relevant (that is, will affect  $P$ 's decision) if and only if  $P$ 's updated belief about the agent's talent crosses  $\theta_P$  as  $E$ 's announcement changes:

$$\mathbf{E}(\theta \mid \theta < \theta_E) < \theta_P \leq \mathbf{E}(\theta \mid \theta \geq \theta_E) \tag{3}$$

---

<sup>6</sup>This is not crucial, it merely simplifies the exposition.

Let us introduce the following notation:

$$\begin{aligned} b_{\max} &\text{ solves } \mathbf{E}(\theta \mid \theta < \theta_E(b_{\max})) = \theta_P \\ b_{\min} &\text{ solves } \mathbf{E}(\theta \mid \theta \geq \theta_E(b_{\min})) = \theta_P, \end{aligned}$$

where  $\theta_E(b) \equiv \theta_P + b$ . It is straightforward that  $b_{\max} > 0$  and  $b_{\min} < 0$ .

Then (3) can be rewritten as

$$b_{\min} \leq b < b_{\max} \tag{4}$$

**Lemma 1**

1. When (4) holds, there always exists an equilibrium with decision-relevant communication. All such equilibria are equivalent:  $E$  simply reports whether  $\hat{\theta} < \theta_E$  or  $\hat{\theta} \geq \theta_E$ , and  $P$  "follows  $E$ 's advice", i.e., takes the unfavorable decision after the former announcement and the favorable one – after the latter announcement.
2. When (4) does not hold, there exists no equilibrium with decision-relevant communication

**Proof.** See the Appendix. ■

**4.2 Agent's choice of effort, equilibrium, and effort-maximizing bias**

If (4) is not satisfied, the agent will not exert effort because he cannot affect  $P$ 's perception of his ability.

Suppose now (4) is satisfied and denote  $y_E(\hat{e}) \equiv \theta_E + \hat{e}$ . The agent's incentive is to beat the performance threshold  $y_E(\hat{e})$ . Given the effort  $e$  and  $E$ 's belief  $\hat{e}$ , the probability of that is

$$\Pr[\theta + e \geq y_E(\hat{e})] = \Pr[\theta \geq y_E(\hat{e}) - e] = 1 - F(y_E(\hat{e}) - e)$$

Hence,  $A$  solves

$$\max_e \delta_A [1 - F(y_E(\hat{e}) - e)] B - c(e)$$

Given our assumptions on  $c(\cdot)$  and Assumption 2, the solution is uniquely determined by the first-order condition

$$\delta_A B f(y_E(\hat{e}) - e) = c'(e) \tag{5}$$

Assumption 2 guarantees a unique solution to the above equation ( $c'(0) = 0$  ensures that a solution exists).

In equilibrium it must be that  $\hat{e} = e^*$ , where  $e^*$  is the equilibrium effort. Hence, the equilibrium effort is determined by

$$\delta_A B f(\theta_E) = c'(e^*) \tag{6}$$

Let us denote the effort-maximizing (incentive-maximizing)  $\theta_E$  by  $\theta_{inc}$ . Since  $c'(\cdot)$  is a strictly increasing function,  $\theta_{inc}$  is the mode of the distribution of  $\theta$ . Notice that  $\theta_{inc}$  does not generally coincide with  $\theta_P$ . In particular, if  $\theta_P < \theta_{inc}$ , then some positive  $b$ , i.e., an *anti-agent* bias, maximizes  $E$ 's effort. If  $\theta_P > \theta_{inc}$ , then some negative  $b$ , i.e., a *pro-agent* bias, maximizes  $E$ 's effort. Notice that thanks to Assumption 3, (3) (and thus (4)) is satisfied at the mode.

The intuition behind the result that the effort maximizing  $\theta_E$  is at the mode can be grasped from the following informal reasoning. If  $\theta_E$  is too low, the agent thinks in the following way: “The performance threshold will be low, so given that my true  $\theta$  is most likely well above  $\theta_E$ , I can ensure passing the performance threshold with a high probability even with little effort; any extra effort will not add much to this probability, thus, there is no point in working harder.” The agent’s reasoning for very high  $\theta_E$  is similar, except that his chances of clearing the performance threshold are too low instead of too high. Only when  $\theta_E$  is close to the mode, working hard really pays off, because the likelihood that the agent’s  $\theta$  is close to  $\theta_E$  is very high and, thus, effort strongly affects the likelihood of passing the threshold.

### 4.3 Optimal bias

The principal’s preferred value of  $\theta_E$  maximizes

$$\begin{aligned} W &= \mathbf{E}\Pi_1 + \delta\mathbf{E}\Pi_2 = \\ &= \mathbf{E}y(\theta, e^*) + \delta [F(\theta_E)z + (1 - F(\theta_E))\mathbf{E}(\theta \mid \theta \geq \theta_E)] \equiv \\ &\equiv t + e^*(\theta_E) + \delta A(\theta_E) \end{aligned} \tag{7}$$

where  $A(\theta_E)$  is the expression in the square brackets, and  $t$  is the unconditional expectation  $\mathbf{E}(\theta)$ . The first order condition with respect to  $\theta_E$  yields

$$\begin{aligned} \frac{dW}{d\theta_E} &\equiv \frac{de^*}{d\theta_E} + \delta \frac{dA}{d\theta_E} \equiv \\ &\equiv \frac{de^*}{d\theta_E} + \delta f(\theta_E)(z - \theta_E) = 0 \end{aligned} \tag{8}$$

or, given that  $\theta_E - z = b$ ,

$$\frac{dW}{d\theta_E} \equiv \frac{de^*}{d\theta_E} - \delta f(\theta_E)b = 0 \tag{9}$$

The principal’s trade-off is simple. If he only cared about the effort maximization, he would choose  $\theta_E$  so that  $de^*/d\theta_E = 0$ , i.e.,  $\theta_E = \theta_{inc}$ . Alternatively, if  $P$  cared solely about the ex-post efficiency, he would maximize  $A(\theta_E)$ , which implies setting  $\theta_E = \theta_P$ , or, equivalently,  $b = 0$ . Since  $P$  cares about both, optimal  $\theta_E$  will be between  $\theta_{inc}$  and  $\theta_P$ . Thus, the optimal bias trades off provision of incentives against optimality of ex-post decisions. Let us denote optimal  $\theta_E$  by  $\theta_W$  and the optimal bias by  $b_W$ . Notice that thanks to Assumption 3, (3) (and thus (4)) is satisfied at  $\theta_W$ .

**Proposition 1** *The ex-ante optimal bias is generally non-zero. It has the same direction as the effort-maximizing bias, but is smaller in magnitude.*

Figure 1 illustrates the proposition. This proposition is the first important result of this work. It establishes the general optimality of biased evaluation in a career concerns framework and illustrates the trade-off that determines the optimal bias.

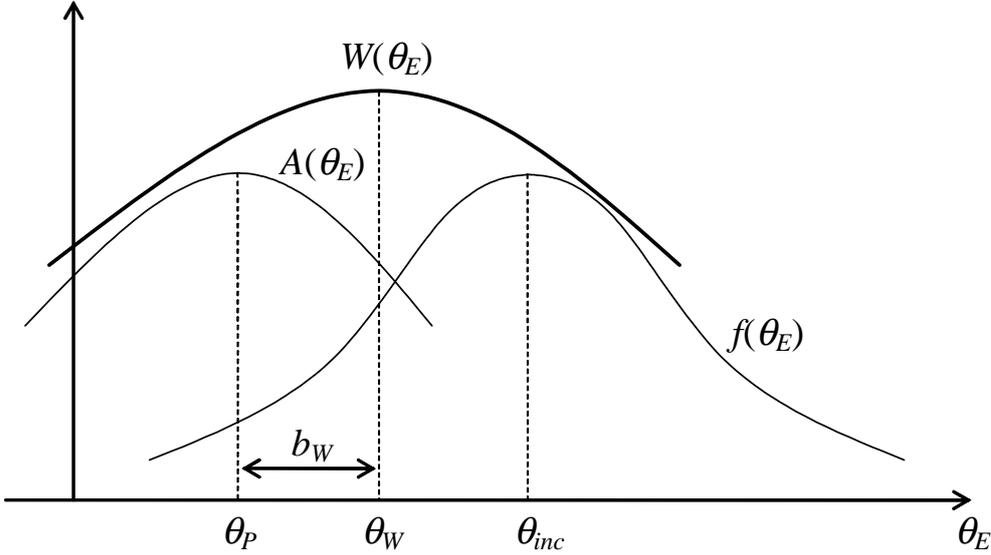


Figure 1. The case when  $\theta_P < Mo(\theta)$ . The optimal bias is  $b_W$ .

Optimality of biased evaluation stands in stark contrast with the vast majority of the principal-agent literature, in which distortions or other imperfections in performance measures are unambiguously bad for the principal (e.g., Holmström, 1979; Feltham and Xie, 1994; Baker, 2002; Prendergast and Topel, 1996). The reason is lack of contractible variables *together with*  $P$ 's commitment problem. For example, if  $y$  were observable and contractible, *and* the principal could commit not to renegotiate, the same result as here could be obtained without any biased evaluation by explicitly including the ex-ante optimal performance threshold in the contract. If  $y$  were not contractible but  $E$ 's report were contractible (like in Prendergast and Topel, 1996), we could again achieve the same result with an unbiased evaluator by conditioning the decision on reported  $\theta$ ; yet, again,  $P$ 's commitment not to renegotiate would be necessary.

The evaluator's bias serves as a commitment device to take ex-post suboptimal decisions in order to induce effort provision. The optimally biased evaluator garbles information about the agent's performance in a way that makes the principal take biased ex-post decisions optimal from the ex-ante perspective. This result parallels the argument by Crémer (1995). In a setup with explicit incentives and contractible performance, Crémer (1995) shows that it can be optimal for the principal to commit to stay uninformed about the causes of poor agent's performance in order to avoid renegotiation. This is a mechanism to create greater incentives for the agent in his model. Thus, the role of biased evaluation in my model is similar to the role of commitment to stay uninformed in Crémer (1995).

In the Appendix I consider an extension of the model in which I allow for contracts stipulating payments from  $P$  to  $A$  contingent on  $P$ 's decision. Although such payments help to provide incentives, they are costly to  $P$ . As a result, it remains optimal to have a biased evaluator.

## 5 Comparative statics

Let us now examine how various parameters affect the optimal bias. I will focus on four things:

- i. The value (to  $P$ ) of the unfavorable decision relative to the *ex-ante* (expected) value of the favorable decision. That is, I will be looking at  $z$  for a fixed distribution of the agent's ability. This is the same as analyzing shifts in the ability distribution relative to  $z$ . Thus, the question can restated as "what is the effect of the *ex-ante* value of the agent?"
- ii. The importance of the future for  $P$
- iii. The strength of career concerns for  $A$  (his discount factor and the benefits he receives from the favorable decision)
- iv. Ex-ante uncertainty about  $A$ 's ability (i.e., the variance of the ability distribution)

### 5.1 The value of the unfavorable decision (inverse of the ex-ante "value" of the agent)

The unfavorable decision is valuable for  $P$  relative to the favorable one from the ex-ante perspective whenever  $z$  is high.

First, for small enough  $z$ ,  $\theta_P \equiv z$  will be below the mode, i.e., below  $\theta_{inc}$ . Since  $\theta_W$  lies between  $\theta_P$  and  $\theta_{inc}$ , for sufficiently small  $z$  the optimal bias must be positive (anti-agent). Analogously, for high enough  $z$ , the optimal bias must be negative (pro-agent).

What is the marginal effect of  $z$ ? Consider (9). Suppose  $z$  goes up. Then  $\theta_E$  increases too, because  $\theta_E \equiv z + b$ . Since  $\theta_W$  is the point where  $W$  is maximized, function  $dW/d\theta_E$  is sloping downwards around  $\theta_W$ , implying that  $dW/d\theta_E$  becomes negative. This means that  $b$  must be decreased in order to restore  $dW/d\theta_E = 0$  (a decline in  $b$  both decreases  $\theta_E$  directly and shifts  $dW/d\theta_E$  upwards).

Thus, a rise in the value of the unfavorable decision leads to a decrease in the *absolute* value of the optimal bias if the bias was positive (i.e., the bias becomes less anti-agent) and to an increase in the *absolute* value of the optimal bias if it was negative (i.e., the bias becomes less pro-agent).

Thus, we can formulate the following proposition.

**Proposition 2** *When the agent is of high value for the principal from the ex-ante perspective (equivalently, the value of the unfavorable decision is small), the optimal bias is anti-agent. As the ex-ante value of the agent falls (equivalently, the value of the unfavorable decision grows),*

the optimal bias monotonically decreases until it becomes pro-agent. A further decrease in the ex-ante value of the agent (an increase in the value of the unfavorable decision) leads to an increase in the pro-agent bias.

The intuition behind this result is as follows. A low value of the unfavorable decision means that  $\theta_P$  is below the mode. This results in too weak incentives for the agent if  $E$  is unbiased. At the same time, the marginal ex-post efficiency loss from making the bias slightly positive is zero at  $\theta_E = \theta_P$ . So, the optimal bias has to be positive. An increase in the value of the unfavorable decision moves  $\theta_P$  closer to the mode, which increases  $A$ 's incentives and reduces the marginal effect of the bias on  $A$ 's incentives for any given bias. This results in a smaller optimal bias. At some point  $\theta_P = \theta_E = \theta_{inc}$ , and further growth of  $z$  pushes  $\theta_P$  to the right of the mode. The incentives fall at  $\theta_P$ , calling for a pro-agent bias in the optimum ( $b < 0$ ). What happens to the right of the mode is just a “mirror image” of the effects to the left of the mode. Figure 2 illustrates the proposition.

In short, this analysis suggests that the evaluator has to be “anti-agent” when the agent is ex-ante very valuable to the principal or there is no good alternative to the agent, and “pro-agent” when the ex-ante value of the agent is rather low or there exists a good alternative to the agent.

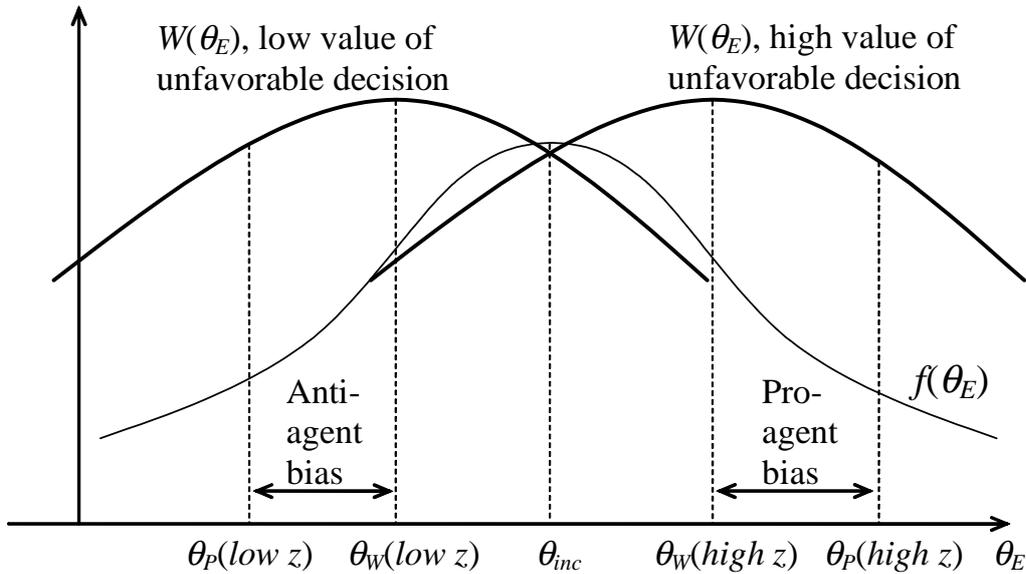


Figure 2. Effect of the value of the unfavorable decision.

## 5.2 Importance of the future and the strength of career concerns

The principal’s discount factor has a straightforward effect: an increase in  $\delta$  results in a lower optimal bias. This follows almost immediately from (9) and is natural: a higher weight of the future makes ex-post efficiency more important than incentives.

**Proposition 3** *The importance of the future for the principal reduces the optimal bias.*

The effect of an increase in  $A$ 's discount factor is not that obvious. As it follows from (6), an increase in  $\delta_A$  induces greater effort. Yet we are interested in the marginal effects of the bias on the effort and ex-post efficiency. Looking at (6) one can notice that  $\delta_A$  amplifies the effect of  $\theta_E$  (hence, the bias) on  $e^*$ . Since  $c'(e^*)$  is an increasing function,  $f'(\theta_E) > 0$  implies  $de^*/d\theta_E > 0$ , and then, clearly,  $de^*/d\theta_E$  grows with  $\delta_A$ . Similarly, if  $f'(\theta_E) < 0$ , then  $de^*/d\theta_E < 0$ , and  $de^*/d\theta_E$  falls with  $\delta_A$ .

Hence, if the optimal bias was anti-agent ( $\theta_P < \theta_E < Mo(\theta) \Leftrightarrow f'(\theta_E) > 0$ ), then an increase in  $\delta_A$  shifts  $dW/d\theta_E \equiv \frac{de^*}{d\theta_E} + \delta \frac{dA}{d\theta_E}$  up near the optimum, implying that  $\theta_W$  and, hence, the bias increase ( $dW/d\theta_E$  is sloping downwards near  $\theta_W$ ). If the optimal bias was pro-agent ( $Mo(\theta) < \theta_E < \theta_P \Leftrightarrow f'(\theta_E) < 0$ ), then an increase in  $\delta_A$  shifts  $dW/d\theta_E$  down near the optimum, implying that  $\theta_W$  decreases. The optimal bias in this case is negative and decreases, which implies that the magnitude of the bias grows.

Since the private benefits,  $B$ , enter (6) in the same way as  $\delta_A$ , their effect is exactly the same. Both  $B$  and  $\delta_A$  reflect the strength of career concerns for the agent.

**Proposition 4** *The strength of career concerns for the agent increases the absolute value of the optimal bias.*

The intuition behind this proposition is rather simple. The role of the bias in this model is to generate incentives. Since the absolute value of the optimal bias is always smaller than the effort-maximizing one (unless  $\theta_P$  is at the mode), there is always under-provision of effort in the optimum. Hence, when the marginal effect of the bias on incentives grows, it is optimal to increase the magnitude of the bias.

The implication of this result is that more career-concerned agents should be evaluated by more biased evaluators, other things being equal. It might be tempting to conclude that, since younger agents are presumably more concerned about their future careers, the evaluator should be particularly biased (either unfriendly or friendly) to young agents. Yet, at the same time, there is usually more uncertainty about the talent of younger agents. Therefore, we cannot make such a statement without examining the effect of the uncertainty about the agent's ability, to which we turn now.

### 5.3 Uncertainty about the agent's ability

The uncertainty can be naturally modeled through an increase in the variance of the prior ability distribution. Let us look at (9). What does an increase in the variance imply for the marginal effects of the bias on the effort and ex-post efficiency? Consider first the effect on the marginal ex-post loss,  $\delta f(\theta_E)b$ . The variance of  $F(\cdot)$  does not have a uniform effect on  $f(\cdot)$ . If we take a symmetric unimodal distribution, e.g., the Gaussian one, the variance lowers  $f(\cdot)$  around the mode and raises  $f(\cdot)$  at the tails (see Figure 3). So, the marginal effect of  $b$  on the ex-post efficiency in the optimum depends on where  $\theta_W$  lies.

Consider now  $de^*/d\theta_E$ . The effort is determined by (6), using which we can write

$$\frac{de^*}{d\theta_E} = \frac{\delta_{AB} f'(\theta_E)}{c''(e^*)}$$

Here also the effect of the variance is ambiguous. Assume, for simplicity, that  $c''(\cdot)$  is a constant ( $c(\cdot)$  is quadratic). The magnitude of  $f'(\cdot)$  does not generally change uniformly with the variance. Take again the Gaussian distribution and consider the case when  $\theta_P$  is smaller than the mode.  $f'(\cdot)$  decreases with the variance close to the mode but increases at the tails.

Thus, for an arbitrary distribution of the agent's ability, the effect of the uncertainty is ambiguous. Yet, in the case of the normal distribution and the quadratic cost of effort, the conclusion turns out to be unambiguous. In such a case, a higher variance of the ability reduces the sensitivity of  $A$ 's effort to the bias, while increasing or just relatively mildly decreasing the effect of the bias on the expected loss from ex-post inefficient decisions.

**Proposition 5** *When the agent's ability is normally distributed, and the cost-of-effort function is quadratic, higher ex-ante uncertainty about the agent's ability reduces the magnitude of the optimal bias.*

*Proof.* See the Appendix. ■

Thus, if there is high uncertainty about the agent's talent, (too) biased evaluation is bad: it has a small effect on incentives, while increasing the ex-post loss. Yet, it needs to be examined to what extent this result is robust to the assumptions about the distribution of the ability and the cost of effort function.

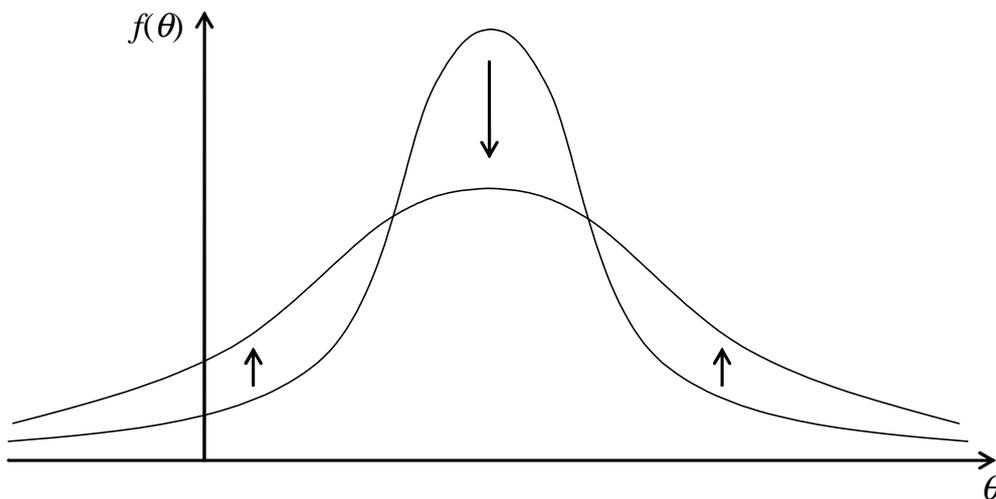


Figure 3. Increase in the variance of the ability distribution.

## 6 Communication versus delegation

In this section I discuss the possibility of delegating decision-making to the evaluator. For the start, let us abstract from the optimal bias and simply look whether delegation can improve over communication for a given bias. Clearly, for  $b_{\min} \leq b < b_{\max}$ , delegation is equivalent to communication. This is because under communication  $P$  always takes  $E$ 's preferred decision.

What if  $b$  is outside this range? We know from Lemma 1 that no decision-relevant communication occurs in such a case. Consequently, the agent cannot influence  $P$ 's decision by working harder and, therefore, exerts zero effort.  $P$  takes his decision based on the prior.

What if we delegate decision-making to  $E$ ? This will hurt ex-post efficiency: the mere fact that  $E$ 's messages would be ignored if she announced whether  $\theta$  lied below or above  $\theta_E$  under communication implies that  $P$  is worse off letting  $E$  choose her preferred actions rather than relying on the prior. Yet, delegation improves effort provision, because now  $A$  has to surpass a certain performance threshold in order to deserve the favorable decision (he actually solves the same problem as in subsection 4.2).

It is easy to invent a situation in which the ex-ante  $P$ 's welfare is improved: one just need to set  $P$ 's discount factor low enough so that ex-post efficiency is relatively unimportant. More generally, on the basis of the above discussion, we can formulate the following proposition.

**Proposition 6** *When the evaluator's bias is not too large, delegation does not improve over communication. When the evaluator's bias is very large, delegation is better than communication whenever provision of incentives is sufficiently more important than ex-post efficiency.*

This result contrasts with Dessein (2002) who showed, in a traditional Crawford-Sobel type of framework, that under large divergence of preferences between the principal and the expert, delegation is inferior to communication. This difference from Dessein's result stems from the fact that, in my model, the mode of decision making has an effect on the third party's incentives (the agent), which affect the principal's welfare. If there were no effort in the model, Dessein's result would follow, since  $P$  would care only about ex-post efficiency.

Can delegation dominate communication in the optimum if the principal can also choose the bias? If Assumption 3 holds, the answer is "no", because the optimal bias, given that  $E$ 's preferred decision is always implemented, lies within  $[b_{\min}, b_{\max})$ . However, if Assumption 3 does not hold, the answer is "yes, it can".

If we drop Assumption 3, the optimal bias under communication will hit the bounds  $b_{\min}$  or  $b_{\max}$  whenever  $P$ 's welfare maximization requires a sufficiently large bias. As we know, setting the bias beyond  $[b_{\min}, b_{\max})$  under communication renders communication decision-irrelevant. Switching to delegation in such cases allows to ignore the bounds and set the bias efficiently, because delegation ensures commitment to  $E$ 's preferred decisions. Thus, the factors that increased the optimal bias in the basic model would also make delegation more likely to be preferred to communication.

**Proposition 7** *If the bias can be freely chosen by the principal, delegation is more likely to preferred to communication when:*

- *the value of the unfavorable decision is more extreme (in either direction),*
- *the strength of the agent's career concerns is higher,*
- *the future is less important for the principal relative to the present,*
- *the lower is ex-ante uncertainty about the agent's ability*

A caveat is in order here: any conclusions about superiority of delegation here rely on the assumption that the principal can actually commit to delegation even when ex-post he would want to deprive the delegee of the decision-making power.

Proposition 6 parallels the results obtained by Ludema and Olofsgård (2008). They consider the problem of protecting a firm (e.g., through a subsidy) that chooses an investment policy. If the government is too protectionist, the firm overinvests. The government is faced with the time-inconsistency problem: ex-post it would like to provide protection more often than is optimal from the ex-ante perspective. The authors compare two solutions to the problem: the government can either delegate the choice of protection to an agency which is less protectionist than the government or just collect information about the social cost of protection from the agency but keep the decision rights. In the latter case, like in my model, the agency can misrepresent information. Among other things, the authors show that delegation is preferred to communication whenever the degree of time inconsistency, i.e., the difference between the ex-ante and ex-post preferred protection policy, becomes too large. The underlying reason for their result is the same as in my model: large difference between the ex-ante and ex-post preferred protection policy implies that the agency has to be very biased with respect to the government's ex-post preferences; but then communication loses credibility in the eyes of the government, and only delegation can cope with the time-inconsistency problem.

## 7 Applications

In reality incentives of agents are affected not just by career concerns, but also by other factors, notably by contracts relating rewards to performance. So, one needs to be careful when applying my model to real-life settings: it is more safely applied to setups in which explicit contracts are not used or are rather imperfect due to lack or noisiness of contractible performance measures.

The first two applications are discussed informally, whereas for the third one I apply some formal analysis.

## 7.1 Evaluation of governmental reforms or programs

A government official who failed an important project (reform or program) is likely to suffer a reputation loss and face meagre career prospects. In contrast, success of the project implies a reputation gain and good chances for career advancement. Thus, the bureaucrat in charge of the project clearly gets a benefit from a positive evaluation of the project's outcome (or the project's progress in the case of interim evaluation). Success of the project would naturally depend on the bureaucrat's ability. The principal here is the government or the society at large, which benefits from promoting or allocating important tasks to talented bureaucrats. Thus, given that bureaucrats rarely have explicit incentives, this setup fits my model well.

The government can delegate evaluation of the project's success to a special committee. Provided that the project is not politically neutral, the committee's bias will be determined by the political preferences of its members. Arguably, most reforms or programs are not politically-neutral in the sense that different political parties have different attitude to these projects.<sup>7</sup> Thus, the evaluator's bias can be regulated by varying the committee's composition.

The favorable decision would be the decision to continue/not to reverse/not to change substantially the project, perhaps together with some decisions favoring the bureaucrat in charge of the project personally, such as a promotion or allocating a new important task. Correspondingly, the unfavorable decision would be the decision to terminate/reverse/change substantially the project, perhaps together with some personal unfavorable decisions regarding the bureaucrat, such as dismissal.

Thus the expected value for the government of the unfavorable decision relative to the favorable one will be greater, the lower are the ex-ante chances of the project's success and the lower are the costs of terminating or reversing the project. Consequently, as Proposition 2 implies, reforms that have ex-ante low likelihood of success and/or are easy to reverse or terminate should be evaluated by committees whose members are on average pro-reform politically. In contrast, reforms with ex-ante high chances for success and substantial costs of reversal or termination should be evaluated by sufficiently anti-reform committees.

## 7.2 Evaluation of CEOs by corporate boards

There has been a long-standing debate on the optimal proportion of independent directors in a corporate board (see, e.g., Adams et al., 2010). My model provides additional considerations on this issue. Altering the board composition one could make the board either more or less tough on the CEO, which would affect the board's decision whether to fire the CEO after poor performance. A reduction in the fraction of independent directors and/or an increase in the proportion of insiders on board raises the friendliness of the board to the CEO. Appointing directors with reputation of being intolerant to failing CEOs makes the board less CEO-friendly

---

<sup>7</sup>Even if a reform benefits everybody in the society, the mere fact that it was proposed by representatives of a certain political party implies that its assessment is likely to depend on the political preferences of the evaluator.

and, possibly, even tougher than the firm’s shareholders would be (corresponding to an “anti-agent bias” in my model).

Of course, in reality CEO incentives are shaped in a much more complex way than just via a threat of dismissal. In particular, there are contracts relating CEO remuneration to various measures of performance.<sup>8</sup> Yet, boards do take dismissal decisions, and, arguably, a CEO suffers a substantial loss of private benefits from being fired. Thus, though I cannot derive any definitive conclusions without a more complete model of CEO incentives, my model provides some arguments on the optimal board composition. To the extent that the threat of dismissal plays a role in CEOs’ incentives, these arguments, I believe, have a value.

First, firms whose CEOs have proven to be extremely talented (superstar CEOs) should have CEO-unfriendly boards.<sup>9</sup> This conclusion follows from Proposition 2: there is a relatively low expected value from replacing a talented CEO. Whether the board’s bias has to be large or not is ambiguous a priori. On the one hand, there is low uncertainty once a CEO has proven his skills. Then, Proposition 5 tells us that the bias should be large. On the other hand, a superstar CEO may have weak career concerns, which implies that the bias should be small, by Proposition 4.

In contrast, other things being equal (strength of career concerns, in particular), a CEO with a short track record of performance should be evaluated by an unbiased (“balanced”) board, i.e., not too tough, not too friendly, because of the high uncertainty about the manager’s ability (Proposition 5).

Finally, turnaround CEOs in highly troubled companies may need to be evaluated by friendly boards due to a relatively high outside option of liquidating the company (value of  $z$ ).

### 7.3 Promotions in organizations

When the head of an organization decides on promotion of an employee, he often relies on information from the employee’s immediate superior or colleagues. While such people are naturally most informed about the potential promotee’s ability, they may also be natural competitors/rivals of him. Then, their evaluations are likely to be biased. Should the principal try to establish less biased channels of evaluation/design a promotion scheme less dependent on opinions of employees with vested interests?

I will now present a formal application of my model to analyzing such questions. Although the below analysis is far from comprehensive, it provides a couple of insights and demonstrates how my framework can be applied for analyzing the internal labor market.

---

<sup>8</sup>Furthermore, subjective board decisions are not just about firing or retaining a CEO: the board may reduce or increase a bonus in a given year or change the remuneration scheme altogether.

<sup>9</sup>One may say that this conclusion contradicts the reality: boards seem friendly to superstar CEOs as they rarely challenge decisions of or dismiss such CEOs. However, first of all, arrangements that are optimal from the shareholders’ perspective should not necessarily arise in reality. Second, according to my model, an optimally biased (anti-CEO) board would fire a superstar CEO just more often than is ex-post optimal, but such events would still occur rarely because a very talented agent would most likely clear the performance bar even when it is set by the (ex-ante optimal) anti-agent evaluator.

### 7.3.1 Adjusted model

Let us consider a hierarchy with the head of the organization (Principal,  $P$ ), a middle manager (Manager,  $M$ ) and an Agent ( $A$ ).  $M$  observes  $A$ 's performance and reports it to  $P$ . Hence, the manager is the evaluator in this example.  $P$  then takes a replacement/promotion decision: whether to promote  $A$  to the  $M$ 's position, in which case the current manager is assumed to be demoted.<sup>10</sup>

The timing is the same as in the basic model, but I adjust the output and payoff functions to capture the specifics of this particular setting. In particular, the 1-st period output now is

$$\Pi_1 = \mu m + (1 - \mu)y,$$

where  $y = \theta + e$  as before, and  $m$  is  $M$ 's ability, which is common knowledge for simplicity.<sup>11</sup>

The 2-nd period output is

$$\Pi_2 = \begin{cases} \mu\theta + (1 - \mu)m & \text{if } A \text{ is promoted} \\ \mu m + (1 - \mu)\theta & \text{if } A \text{ is not promoted} \end{cases}, \quad \mu > 1/2$$

Thus, naturally, both  $A$  and  $M$  contribute to the 2-nd period output, with  $\mu$  being the weight of the contribution of a person holding the managerial position (if  $A$  and  $M$  are switched,  $A$  becomes the manager, and  $M$  becomes the agent). The assumption  $\mu > 1/2$  reflects the (realistic) idea that higher level employees have a greater impact on the organization's welfare.

$M$ 's period 2 payoff is

$$x\Pi_2 + I(\text{retain position}) \cdot B,$$

where  $B$  is the private benefit from holding a managerial position, and  $x$  is  $M$ 's sensitivity to the organization's output. Both parameters are exogenous. Hence, the evaluator's (manager's) bias here stems from here desire to keep the private benefit. Notice that because  $B$  is positive, the bias is always anti-agent. The size of the bias will depend on the relative magnitudes of  $B$  and  $x$ .

$A$ 's payoff in period 2 is

$$I(\text{promoted}) \cdot B$$

Thus, I assume, for simplicity, that  $A$  cares only about being promoted, but not about the organization's welfare.<sup>12</sup>

Despite the modifications, the analysis is essentially the same as in the basic model. The favorable decision yields  $\mu\theta + (1 - \mu)m$  to  $P$ , while the unfavorable one results in  $\mu m + (1 - \mu)\theta$ .

<sup>10</sup>Alternatively, I could assume that  $M$  is fired with a certain probability.

<sup>11</sup>It is natural to assume that more is known about the manager's ability, who presumably has worked longer in the organization, compared to the agent. Also, because the focus is on the agent's incentives, I am assuming here, for simplicity, that the manager's effort plays no role.

<sup>12</sup>The conclusions would not change if I assumed that  $A$  cared about the organization's payoff too.

Since  $\mu > 1/2$ ,  $P$  prefers to promote  $A$  whenever his belief about  $\theta$  exceeds  $m$ , i.e.:

$$\theta_P = m$$

$M$ 's threshold solves  $x[\mu\theta_E + (1 - \mu)m] = x[\mu m + (1 - \mu)\theta_E] + B$ . Hence,

$$\theta_E = m + \frac{B}{x(2\mu - 1)}$$

Thus,  $\frac{B}{x(2\mu - 1)}$  is the measure of the bias, which is always positive in this example ( $\mu > 1/2$ ).

For given  $\theta_E$ ,  $A$ 's effort choice and the equilibrium are determined exactly as in subsection 4.2, i.e.,

$$\delta_A B f(\theta_E) = c'(e^*) \quad (10)$$

$P$ 's ex-ante welfare is

$$\begin{aligned} W &= \mathbf{E}\Pi_1 + \delta \mathbf{E}\Pi_2 = \\ &= \mu m + (1 - \mu) \mathbf{E}y(\theta, e^*) + \delta \left\{ \begin{array}{l} F(\theta_E) [\mu m + (1 - \mu) \mathbf{E}(\theta \mid \theta < \theta_E)] + \\ + (1 - F(\theta_E)) [\mu \mathbf{E}(\theta \mid \theta \geq \theta_E) + (1 - \mu)m] \end{array} \right\} \equiv \quad (11) \\ &\equiv \mu m + (1 - \mu)(t + e^*(\theta_E)) + \delta A(\theta_E), \end{aligned}$$

where  $A$  is the expression in braces.

This expression is essentially the same as expression (7). So, *if* we could freely alter  $\theta_E$  independently of  $B$ , we would obtain the first order condition analogous to (8) or (9):

$$\begin{aligned} \frac{dW}{d\theta_E} &\equiv (1 - \mu) \frac{de^*}{d\theta_E} + \delta \frac{dA}{d\theta_E} \equiv \\ &\equiv (1 - \mu) \frac{de^*}{d\theta_E} - \delta(2\mu - 1) f(\theta_E) (\theta_E - m) = 0 \end{aligned}$$

In this example  $\theta_E$  can be varied by changing, e.g.,  $B$  or  $x$ . In the former case, however,  $B$  also affects the effort directly, not just through  $\theta_E$  (see (10)); thus, differentiating  $P$ 's welfare with respect to  $B$  is not equivalent to differentiating it with respect to  $\theta_E$ . In either case, since  $x$  and  $B$  are restricted to be positive, an interior solution would not be guaranteed.

Whereas considering effects of  $B$  and  $x$  may make sense, I would like to discuss here the choice of the bias (or of  $\theta_E$ ) through the choice of the evaluation and promotion scheme. I will call the scheme I have just described (evaluation by the immediate boss and promotion to the boss' place) the "biased scheme". The bias in these scheme is *always* anti-agent. It can be compared to two alternative schemes, characterized by zero bias. I called them "unbiased schemes":

- i.  $A$  is evaluated by people who have no conflict of interest regarding  $A$ 's promotion (e.g., by colleagues from a parallel division who do not compete with  $A$  for promotion or by  $P$  himself),
- ii.  $A$  is still evaluated by his immediate boss, but is promoted to a position in *another* department.

The choice of the bias is thus discrete, so we will not be able to use differentiation, but the general intuition from the basic model still applies (if you look at Figure 2, we are simply comparing the point where  $\theta_E = \theta_P$  with some point to the right of  $\theta_P$ ).<sup>13</sup> Generally, factors that called for a greater *anti-agent* bias in the basic model will make the biased scheme more likely to be preferred here. In contrast, factors that either called for a low bias or a pro-agent bias will make the unbiased scheme preferred (clearly if the optimal bias is pro-agent, then zero bias is preferred to an anti-agent bias).

### 7.3.2 Effect of middle managers' skills

The value of the “unfavorable” decision is increasing with  $M$ 's talent, so we can draw a relationship between the quality of managers and optimality of various evaluation schemes. Building on the result of Proposition 2, the biased scheme is more likely to be optimal whenever keeping the status quo (“unfavorable decision” in the terminology of the basic model) is relatively unattractive, i.e., when  $m$  is sufficiently smaller than the mode of the distribution of  $\theta$ .

When  $m$  is above the mode, a pro-agent bias would be optimal, but we do not have such a scheme here. Yet, in this case, zero bias is better than an anti-agent bias, so an unbiased scheme is preferred.

One can argue that normally  $m$  is rather high compared to the average  $\theta$  (and, most likely, to the mode), because in order to become a manager one needs to prove that his talent is above the average. Thus, normally, an unbiased scheme should be optimal.

Yet, with time a manager's skills can deteriorate (e.g., due to age) or become obsolete in a changing environment. This is especially likely to happen when an organization goes through a period of transformation that makes skills or business approaches of incumbent managers obsolete and, hence, needs new blood at the managerial level. Then we are in the case of low  $m$ , and, according to my results, the biased scheme should be optimal. That is, in such circumstances, *junior employees should be evaluated by “dead wood” managers from the same division and should be restricted from promotions to a different division.* This is a quite surprising conclusion.

---

<sup>13</sup>Of course, one can also think about “fine-tuning” the bias by mixing schemes, i.e., using them with some probabilities. I am abstracting from this here.

### 7.3.3 Effect of private benefits from a managerial position

From Proposition 4 we know that greater strength of career concerns calls for a higher optimal bias. At the same time, high  $B$  in this example means not just stronger career concerns but also a greater bias of the biased scheme. Thus, if the biased scheme initially dominates the unbiased schemes sufficiently strongly, a change in  $B$  will not alter the dominance. Similarly, if the unbiased schemes clearly dominates the biased one, changing  $B$  will have no effect on the choice of the scheme. Yet, if the preference for one of the schemes is initially slight, the effect of  $B$  is ambiguous. Suppose, for example, that the biased scheme was only slightly preferred. This means that actually the optimal bias is lower than the one induced by the biased scheme (if we could somehow reduce the bias without changing the agent's private benefit, we would do that). Let us raise  $B$ . Stronger career concerns would call for a higher optimal bias if we could set any bias we wanted. But the bias of the biased scheme was already too high, and a further increase in the bias due to an increase in  $B$  may actually tilt our preference towards an unbiased scheme. Further investigation of this question is needed.

### 7.3.4 Comparison to Friebel and Raith (2004) and Friebel and Raith (2013)

My analysis can be compared to Friebel and Raith (2004) and Friebel and Raith (2013). Similarly to my model, in Friebel and Raith (2004), a talented agent is a threat to the Manager. The authors argue that restricting direct communication between the Principal and the Agent (i.e., bypassing the Manager) can be efficiency improving because then the Manager is less afraid to recruit or develop talented subordinates. In my model, shutting down direct evaluation by the Principal can also be optimal. When it is, the reason is also related to the fact that the Manager is afraid of being replaced. However, the mechanism is totally different. In contrast to Friebel and Raith (2004), the Manager's fear plays a positive role, because it induces the Agent to work harder.

Friebel and Raith (2013) compare two promotion systems: "silo", in which junior employees can be promoted only within their initial divisions, and "lattice", in which promotions across divisions are allowed. In their paper, a well trained employee (Agent) benefits the Manager through a higher productivity of the division (there is no threat of replacement in that model, managerial departures are exogenous). Then, a silo has a positive effect on the Manager's incentives to train employees, because in a silo employees cannot be moved to another division. In my model, a silo in which the employee is evaluated by the immediate boss can be good thanks to a positive effect on the employee's incentives to work hard.

## 8 Conclusion

This paper has shown that evaluation of an agent by a biased intermediary can be optimal in a career concerns setting and proposed a framework, based on the trade-off between ex-post

efficiency and incentives, which one can apply to study the optimal bias. The optimal bias (both the direction and the magnitude) is determined by the ex-ante value of the agent to the principal, the weight of the future in the principal’s objective function, the strength of the agent’s career concerns, and the uncertainty about the agent’s ability.

I have also compared simply evaluation by the intermediary (communication) to delegation of the decision about the agent to the intermediary. Whenever the parameters of the environment are such that a large bias is needed in order to create incentives, delegation is more likely to dominate communication. This results contrasts with Dessein (2002), who found that delegation is preferred to communication only when the divergence of preferences between the principal (receiver) and the expert (sender) is small enough. The difference in the results is due to the fact that, in my model, the mode of decision making has an effect on the third party’s incentives (the agent), which affect the principal’s welfare.

I have discussed how my model can be applied to several real-life settings. Yet, in order to study any of this applications properly, one would need to set up a formal model which would take into account specific realistic features of a given setting (such as, for example, possibilities to have contracts based on some performance measures in the case of CEO evaluation). One potential application that I have not mentioned is the design of tenure review. Should the tenure committee be biased (e.g., too tough) with respect to the preferences of the university? How does the answer depend on the university’s objective function and the parameters of the academic job market?

Thus, analyzing various applications in detail could be one direction for future work. Another direction could be incorporating dynamics in the model. If one thinks of an “internal labor market” in an organization or government, favorable or unfavorable decisions affect the distribution of talent across different levels of the organization, which in turn affects the incentives of both evaluatees and evaluators (when the latter are also the organization’s employees, like in subsection 7.3). Thus, one could potentially study the evolution of the talent distribution in an organization and, perhaps, the evolution of the general “quality” of the organization.

## Appendix

### **Proof of Lemma 1.**

In any equilibrium in which  $P$  takes into account  $E$ ’s messages (decision-relevant communication), the following must hold:

First,  $E$ ’s strategy must be equivalent to sending binary messages (“yes”/“no”). For any  $\hat{\theta}$ ,  $E$  prefers either the favorable or unfavorable decision (for simplicity, we assumed that in the case of indifference, which is a zero-measure event,  $E$  prefers the favorable decision). So, if there were three or more messages, all having different effects on  $P$ ’s decision,  $E$  would only use the two that would lead to the highest and the lowest probability of taking the favorable decision.

Second, for the same reason,  $E$  will never mix between the messages.

Third, the communication strategy must have a threshold structure with a single threshold  $\theta_E$  such that  $E$  says “no” for  $\hat{\theta} < \theta_E$  and “yes” otherwise. A single threshold follows from the fact that if  $E$  prefers the favorable decision for some  $\theta'$ , she will prefer it for any  $\theta > \theta'$ ; the same is true for the unfavorable decision. The threshold must obviously be  $\theta_E$ , for if it were  $\tilde{\theta} \neq \theta_E$ , that would mean that  $E$  plays suboptimally for  $\theta$ s between  $\tilde{\theta} \neq \theta_E$ .

Now, in order for  $P$  to find it optimal to follow  $E$ 's messages it must be that the messages move his belief from the left of  $\theta_P$  to the right and the vice versa. This condition is exactly (4). So, we proved the first part of the lemma.

The proof of the second part is straightforward. Suppose (4) does not hold, but there is an equilibrium with decision-relevant communication. Then, from the proof of the first part, we know that  $E$ 's strategy is binary with threshold  $\theta_E$ . But then, since (4) does not hold,  $P$  will ignore  $E$ 's messages, which is a contradiction.

**Proof of Proposition 5.**

Assume the cost of effort is  $e^2/2$ . Then (6) becomes

$$\delta_{AB}f(\theta_E) = e^*$$

Thus

$$\frac{de^*}{d\theta_E} = \delta_{AB}f'(\theta_E),$$

and (9) becomes

$$\delta_{AB}f'(\theta_E) - \delta f(\theta_E)b = 0 \tag{12}$$

In the case of the normal distribution

$$f'(\theta_E) = \frac{t - \theta_E}{\sigma^2} f(\theta_E),$$

where  $\sigma$  is the standard deviation.

So, (12) becomes

$$\delta_{AB} \frac{t - \theta_E}{\sigma^2} - \delta b \equiv \delta_{AB} \frac{t - \theta_E}{\sigma^2} - \delta(z - \theta_E) = 0,$$

from which we obtain

$$\theta_W = z \cdot \frac{\frac{\delta_{AB}t}{\sigma^2} + \delta\sigma^2}{\delta_{AB} + \delta\sigma^2}$$

If  $\frac{t}{z} > 1$ ,  $\theta_W$  is decreasing in  $\sigma$ . But  $\frac{t}{z} > 1 \Leftrightarrow \theta_P \equiv z < t \equiv Mo(\theta)$  for the normal distribution. So, for  $\theta_P < Mo(\theta)$ ,  $\theta_W$  is decreasing in  $\sigma$ , which implies that the bias is decreasing in  $\sigma$ .

If  $\frac{t}{z} < 1$  (equivalent to  $\theta_P > Mo(\theta)$ ),  $\theta_W$  is increasing in  $\sigma$ . Thus, for  $\theta_P > Mo(\theta)$ ,  $\theta_W$  is increasing in  $\sigma$ , which again implies that the absolute value of the bias is decreasing in  $\sigma$ .

**Extension: optimal bias when monetary rewards for the favorable decision are allowed.**

Suppose  $P$  can credibly promise to reward  $A$  in the case of the favorable decision. Let the reward be  $x \geq 0$ . Let us assume for simplicity that the reward cannot be renegotiated. Because  $x$  is a loss for  $P$ , his ex-post optimal threshold for the favorable decision becomes

$$\theta_p = z + x$$

$E$ 's threshold is biased by  $b$  as before

$$\theta_E = z + x + b$$

Given decision-relevant communication,  $A$  chooses effort solving

$$\max_e \delta_A [1 - F(y_E(\hat{e}) - e)] (B + x) - c(e),$$

i.e., the same problem as before except that his benefit from the favorable decision is  $B + x$  instead of just  $B$ .

Thus, the equilibrium effort will be determined by

$$\delta_A(B + x)f(\theta_E) = c'(e^*),$$

which is analogous to (6). Since now we will be optimizing with respect to both  $b$  and  $x$ , let us rewrite this condition as

$$\delta_A(B + x)f(z + x + b) = c'(e^*) \tag{13}$$

$P$  chooses  $x$  and  $b$  so as to maximize

$$W = t + e^*(b, B, x) + \delta [F(z + x + b)z + (1 - F(z + x + b))\mathbf{E}((\theta - x) \mid \theta \geq z + x + b)],$$

where  $e^*(b, B, x)$  is the solution to (13). Notice that when the favorable decision is taken  $P$  obtains  $\theta - x$  rather than  $\theta$  as before. Notice also that  $x$  increases the probability of the unfavorable decision.

If we assume an interior solution for both  $x$  and  $b$ , the first order conditions yield:

$$\frac{dW}{dx} = e_x^*(b, B, x) + \delta [-f(z + x + b)b - (1 - F(z + x + b))] = 0 \tag{14}$$

$$\frac{dW}{db} = e_b^*(b, B, x) + \delta [-f(z + b + x)b] = 0 \tag{15}$$

Notice, however, that an interior solution for  $x$  will not always exist. Consider the case when  $\delta$  is very large so that  $P$  is concerned almost exclusively with ex-post efficiency, and  $e_x^*(b, B, x)$

and  $e_b^*(b, B, x)$  can almost be ignored. Then, from (15)  $b$  is very close to zero, which is natural in such a case. But then, the term in the square brackets in (14) becomes almost  $-(1 - F(z + x))$ , and the whole derivative  $dW/dx$  is negative regardless of  $x$ . That means that the solution for  $x$  is not interior,  $x = 0$  in such a case.

This reasoning already shows that for large enough  $\delta$ , monetary rewards will not be used, and the optimal bias will be determined precisely as in the basic model, i.e., it will generally be non-zero, though rather small. Let us show now that even when optimal  $x$  is positive, optimal  $b$  is still generally different from zero.

So, suppose the solution is interior for both  $b$  and  $x$ . Assume that  $F(\cdot)$  is Gaussian with mean  $t$  and standard deviation  $\sigma$ , and  $c(e) = e^2/2$ . Then, using the fact that for the normal distribution  $f'(\xi) = \frac{t-\xi}{\sigma^2}f(\xi)$ , we can obtain that (14) and (15) become respectively

$$\delta_A f(z+x+b) \left[ (B+x) \frac{t-(z+x+b)}{\sigma^2} + 1 \right] + \delta [-f(z+x+b)b - (1 - F(z+x+b))] = 0 \quad (16)$$

and

$$\delta_A (B+x) \frac{t-(z+x+b)}{\sigma^2} - \delta b = 0 \quad (17)$$

Now, suppose  $b = 0$ . Then, from (17) we have  $z + x + b = t$ . Then, given that  $F(t) = 1/2$ , (16) becomes

$$\delta_A f(t) - \frac{\delta}{2} = 0$$

Clearly, this condition is satisfied only in a very special case: namely, if the relationship between the parameters is given by this condition itself. Since it is generally not true,  $b$  is generally different from zero in the optimum.

Analyzing the properties of optimal  $x$  and  $b$  (how they depend on various parameters) is not an easy task. One thing, however, can be shown easily:  $b$  is decreasing in  $z$  as in the basic model. To see this, multiply (17) by  $f(z+x+b)$  and subtract it from (16). The result is

$$\delta_A f(z+x+b) - \delta(1 - F(z+x+b)) = 0$$

If  $z$  goes up,  $x + b$  must go down so that  $z + x + b$  remains constant. Then, from (17) it is clear that it cannot be that  $x$  goes down while  $b$  goes up (or vice versa): both variables have to decrease.

## References

- [1] Adams, R.B., B.E. Hermalin, and M.S. Weisbach, 2010, "The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey," *Journal of Economic Literature*, 48(1): 58-107.

- [2] Ambrus, A., E.M. Azevedo, and Y. Kamada, 2013, "Hierarchical cheap talk," *Theoretical Economics*, 8(1): 233-261.
- [3] Baker, G., 2002, "Distortion and Risk in Optimal Incentive Contracts," *The Journal of Human Resources*, Vol. 37, No. 4, pp. 728-751.
- [4] Berglöf, E., and E.L. Von Thadden, 1994, "Short-term versus long-term interests: Capital structure with multiple investors," *The Quarterly Journal of Economics*, 109(4): 1055-1084.
- [5] Besley, T., and S. Coate, 2001, "Lobbying and welfare in a representative democracy," *Review of Economic Studies*, 86: 67-82.
- [6] Bolton, P., and D.S. Scharfstein, 1996, "Optimal debt structure and the number of creditors," *Journal of Political Economy*, 104(1): 1-25.
- [7] Chevalier, J., and G. Ellison, 1999, "Career Concerns of Mutual Fund Managers," *The Quarterly Journal of Economics*, 114 (2): 389-432.
- [8] Crémer, J., 1995, "Arm's Length Relationships," *The Quarterly Journal of Economics*, Vol. 110, No. 2, pp. 275-295.
- [9] Crawford, V., and J. Sobel, 1982, "Strategic Information Transmission," *Econometrica*, 50(6), 1431-1451.
- [10] Dessein, W., 2002, "Authority and Communication in Organizations," *Review of Economic Studies*, 69 (4): 811-838.
- [11] Dewatripont, M., and E. Maskin, 1995, "Credit and efficiency in centralized and decentralized economies," *The Review of Economic Studies*, 62(4): 541-555.
- [12] Dewatripont, M., and J. Tirole, 1994, "A theory of debt and equity: Diversity of securities and manager-shareholder congruence," *The Quarterly Journal of Economics*, 109(4): 1027-1054.
- [13] Dewatripont, M., and J. Tirole, 1996, "Biased principals as a discipline device," *Japan and the World Economy*, 8(2): 195-206.
- [14] Dewatripont, M., I. Jewitt, and J. Tirole, 1999a, "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66 (1): 183-198.
- [15] Dewatripont, M., I. Jewitt, and J. Tirole, 1999b, "The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies," *Review of Economic Studies*, 66 (1): 199-217.
- [16] Feltham, G.A., and J. Xie, 1994, "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations," *The Accounting Review*, 69, No. 3, pp. 429-453.

- [17] Fershtman, C., 1985, "Managerial incentives as a strategic variable in duopolistic environment," *International Journal of Industrial Organization*, 3(2): 245-253.
- [18] Fershtman, C., and K.L. Judd, 1987, "Equilibrium incentives in oligopoly," *The American Economic Review*, 77(5): 927-940.
- [19] Fershtman, C., K.L. Judd, and E. Kalai, 1991, "Observable contracts: Strategic delegation and cooperation," *International Economic Review*, 32: 551-559.
- [20] Friebel, G., and M. Raith, 2004, "Abuse of Authority and Hierarchical Communication," *RAND Journal of Economics*, 35 (2): 224-244.
- [21] Friebel, G., and M. Raith, 2013, "Managers, Mentoring, and Internal Labor Markets," [http://works.bepress.com/michael\\_raith/16/](http://works.bepress.com/michael_raith/16/)
- [22] Gehlbach, S., and A. Simpsen, 2015, "Electoral manipulation as bureaucratic control," *American Journal of Political Science*, 59(1): 212-224.
- [23] Gibbons, R., and K.J. Murphy, 1992, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 1992, 100: 468-505.
- [24] Hart, O., and J. Moore, 1998, "Default and Renegotiation: A Dynamic Model of Debt," *The Quarterly Journal of Economics*, 113(1): 1-41.
- [25] Holmström, B., 1979, "Moral Hazard and Observability," *Bell Journal of Economics*, 10: 74-91.
- [26] Holmström, B., 1999, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economics Studies*, 66: 169-182.
- [27] Ivanov, M., 2010, "Communication via a strategic mediator," *Journal of Economic Theory*, 145(2): 869-884.
- [28] Ludema, R.D., and A. Olofsgård, 2008, "Delegation versus Communication in the Organization of Government," *Journal of Public Economics*, 92: 213-235.
- [29] Persson, T., and G.E. Tabellini, 2000, *Political Economics: Explaining Economic Policy*, MIT Press.
- [30] Persson, T., and G.E. Tabellini, 1993, "Designing institutions for monetary stability," In *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 53-84, North-Holland.
- [31] Persson, T., and G.E. Tabellini, 1994, "Representative democracy and capital taxation," *Journal of Public Economics*, 55(1): 53-70.

- [32] Persson, T., and G.E. Tabellini, 2002, "Political Economics and Public Finance," Chapter 24 in Auerbach, Alan and Martin Feldstein (eds), *Handbook of Public Economics*, Volume 3, Amsterdam: Elsevier.
- [33] Prendergast, C., and R.H. Topel, 1996, "Favoritism in Organizations," *The Journal of Political Economy*, Vol. 104, No. 5, pp. 958-978.
- [34] Rogoff, K., 1985, "The optimal degree of commitment to an intermediate monetary target," *Quarterly Journal of Economics*, 100: 1169-1189.
- [35] Sengul, M., J. Gimeno, and J. Dial, 2012, "Strategic Delegation: A Review, Theoretical Integration, and Research Agenda," *Journal of Management*, 38(1): 375-414.
- [36] Sklivas, S.D., 1987, "The strategic choice of managerial incentives," *The Rand Journal of Economics*, 18(3): 452-458.
- [37] Vickers, J., 1985, "Delegation and the theory of the firm," *Economic Journal*, 95: 138-147.