TI 2014-115/VII
Tinbergen Institute Discussion Paper

Biased Supervision

Josse Delfgaauw
Michiel Souverijn

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

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

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

Biased Supervision

Josse Delfgaauw and Michiel Souverijn*

June 2016

Abstract

Organizations can use subjective performance pay when verifiable performance measures are imperfect. However, this gives supervisors the power to direct employees towards tasks that mainly benefit the supervisor rather than the organization. We cast a principal-supervisor-agent model in a multitask setting, where the supervisor has an intrinsic preference towards specific tasks and may receive soft information on the agent's efforts. We show that subjective performance pay based on evaluation by a biased supervisor has the same distorting effect on the agent's effort allocation across tasks as incentive pay based on an incongruent performance measure. Combining incongruent performance measures with biased supervision can mitigate, but does not always eliminate this distortion. We apply our results to the choice between specialist and generalist middle managers, where a trade-off between monitoring ability and bias arises.

JEL codes: J24, M12, M52.

Kewyords: subjective performance evaluation, middle managers, incentives, multitasking.

*Erasmus University Rotterdam and Tinbergen Institute. E-mail: delfgaauw@ese.eur.nl, souverijn@ese.eur.nl.
We thank the Editor, two referees, Daniel Ferreira, Jurjen Kamphorst, Dirk Sliwka, Otto Swank, Jan Tichem, Jan Zabojnik, and participants at the POEK 2014 conference in Cologne for valuable comments and suggestions. We gratefully acknowledge financial support from NWO through an Open Competition grant, no. 404-10-352.

1 Introduction

In many organizations, middle managers' assessment of employees' performance is an important determinant of bonus pay and career prospects.[1] If verifiable performance measures are imperfect, subjective performance evaluation may provide a more accurate assessment of employees' performance, thereby providing better incentives for employees. On the other hand, subjective performance evaluation can be manipulated, weakening the link between actual and reported performance. Furthermore, their role in determining pay and promotion opportunities gives managers (more) power over their subordinates. Earlier work has shown that performance pay based on middle managers' evaluations can be prone to favoritism (Prendergast and Topel 1996, Bol 2011, Dur and Tichem 2015), collusion (Tirole 1986, Vafaï 2010, Thiele 2013), extortion (Laffont 1990, Vafaï 2002, 2010), and a lack of incentives or ability to monitor (Gibbs et al. 2004, Bol 2011, Kamphorst and Swank 2015).

In this paper, we study subjective performance pay in a principal-supervisor-agent model, where the supervisor uses her discretionary power to pull the agent towards tasks that benefit the supervisor more than the organization. The agent exerts effort on multiple tasks, which is unobserved by the principal. Depending on her monitoring ability, the supervisor may receive soft information on the agent's efforts. The supervisor provides a report on the agent's performance to the principal, which can be used in determining the agent's incentive pay. Crucially, we assume that the supervisor has an intrinsic preference for particular tasks exerted by the agent. This makes that she overemphasizes these tasks when providing directions to the agent. Anticipating that not living up to the supervisor's expectations results in a bad evaluation, the agent works towards the supervisor's goals. As a consequence, akin to the standard multitasking model (Holmström and Milgrom 1991, Baker 1992, 2002), we show that the principal optimally sets weaker subjective performance pay when the supervisor's preferences are less aligned, as well as when the supervisor has lower ability.[2]

[1] For instance, Eccles and Crane (1988), Gibbs (1995), and Bol (2011) document the use of subjective performance evaluation in (financial) service firms, Breuer et al. (2013) in a large call-center, Gibbs et al. (2004) in car-dealerships, Woods (2012) in an internal audit firm, and Medoff and Abraham (1980) in manufacturing firms.

[2] Supervisors can also use their power to affect (the behaviour of) employees in ways that are not directly linked to employees' tasks at work, e.g. by engaging in bullying, extortion, (sexual) harassment, etc. Our interest

This changes when the principal has access to a verifiable, but possibly incongruent, performance measure. To structure ideas, consider a salesman of a local store owned by a retail chain. The store's manager is an active member of the local community, so that she cares a lot about her store's reputation for providing good service. The salesman contributes to long-run store performance through sales effort and service effort. The latter does not contribute directly to short-run sales, but increases the reputation of the local store, which has long-run benefits to the retail chain. The manager monitors the salesman's efforts, but the chain's headquarters only observes sales. If headquarters uses the salesman's sales figures to provide incentive pay, he will focus disproportionately on sales at the expense of service. Alternatively, headquarters could relate the salesman's pay to his performance evaluation as provided by the store manager. However, in evaluating performance, the manager will put too much emphasis on service provision, inducing the salesman to exert suboptimally low sales effort. Combining verifiable sales figures with subjective performance evaluation in the salesman's bonus plan brings several advantages. First, sales targets constrain the store manager in emphasizing service at the expense of sales. Second, the use of subjective performance evaluation allows the manager to pull the salesman away from the disproportionate focus on sales induced by sales targets. Third, the sales figures provide additional information on the salesman's efforts, allowing for better monitoring.[3]

We show that by offering bonus pay conditional on achieving both a performance target and a favorable subjective evaluation, the principal may mitigate the distortion that arises when using either objective or subjective performance pay exclusively. This relates to the literature on contracting with multiple incongruent performance measures (Feltham and Xie 1994, Datar et al. 2001, Budde 2007), where it has been shown that full congruence can be achieved if the number of verifiable measures meets or exceeds the number of tasks. Even when all measures are biased towards the same task, congruence is feasible by placing a negative weight on the most biased measure. In contrast, we show that this does not hold when some measures are subjectively determined. Placing a negative weight on the subjective evaluation is ineffective.

lies with supervisors' incentives to provide misaligned directions regarding employees' efforts at work.

[3]Manthei and Sliwka (2014) provide a subset of local managers of a retail bank, who previously allocated bonus pay based on subjective assessment, with individual sales data of their employees. This increased both employees' sales activities and profit.

If a good evaluation would have a negative effect on the agent's compensation, the supervisor could still direct the agent towards the tasks she considers important by threatening to provide a good evaluation unless the agent follows her directions. Hence, congruence is not feasible when the supervisor is more biased than the verifiable performance measure.

When the verifiable performance measure and the supervisor are biased towards different tasks, the principal implement non-distorted efforts, unless either the supervisor's monitoring ability is low or the performance measure is unreliable. We model the latter as the probability with which the agent can ex post costlessly manipulate measured performance.[4] If this probability is too high, the supervisor ignores the principal's performance target and induces her most preferred effort allocation. Similarly, if the supervisor's ability is too low, the agent ignores her instructions and meets the performance target at lowest effort cost by working purely towards measured performance. To prevent these outcomes, the principal must allow for some bias in effort allocation and optimally reduces the agent's incentive pay.

The key assumption of our model is that the supervisor has intrinsic preferences over the agent's tasks, which may differ from the principal's relative valuation of these tasks. Such preferences could be driven by private benefits, by career concerns, or by professional norms. The supervisor may overemphasize providing input into her own work, or overemphasize tasks that benefit the supervisor's unit at the expense of activities that benefit other units. Alternatively, the supervisor may intrinsically consider particular tasks more important, as in e.g. Akerlof and Kranton (2005) and Prendergast (2007). In Guth and Macmillan (1984), middle managers admit to making decisions that are not aligned with corporate strategy and goals, in order to protect their self-interest. Burgelman (1994) argues that in the 1980s, Intel had to change corporate strategy after middle managers made resource allocation decisions that went against the initial strategy. Our analysis shows that misaligned middle management may, but need not be detrimental for firm performance, depending on the available performance measures.

[4]Examples of manipulation of performance information abound. Nagin et al. (2004) study monitoring of call-center agents who can falsely report sales. Alternatively, employees may be able to influence the timing of sales around target commencement dates, as documented by Asch (1990), Oyer (1998), Courty and Marschke (2004), and Larkin (2014). In the accounting literature, manipulation of information is an important theme, ranging from earnings management to accounting fraud (see e.g. Holthausen et al. 1995, Efendi et al. 2007, and Goldman and Slezak 2006)

Supervisor's biased preferences over tasks differ from interpersonal preferences such as altruism, spite or favouritism, as in Prendergast and Topel (1996), Giebe and Gürtler (2012), and Dur and Tichem (2015). Typically, interpersonal preferences mute the incentive effect of subjective performance pay by weakening the link between employees' effort and pay, and can explain the well-documented leniency bias in performance appraisal (Jawahar and Williams 1997, Prendergast 1999, Bol 2011). When the supervisor has better information about performance than employees, she may provide overly positive evaluations attempting to increase employees' perception of their ability (Suvorov and Van de Ven 2009, Zabojnik 2014, Kamphorst and Swank 2015). In our framework, biased supervision reduces the value of subjective performance pay through misallocation of employees' efforts across tasks. Müller and Weinschenk (2015) consider persistence of a supervisor's opinion regarding employees' performance. This induces higher effort in early periods at the expense of later periods, which need not be costly if the principal responds by setting weaker (stronger) incentives in early (later) periods.

Most earlier work on combining subjective and objective performance measures considers subjective evaluation by the principal (Baker et al. 1994, Schmidt and Schnitzer 1995, Pearce and Stacchetti 1998, Budde 2007). Following Bull (1987), the emphasis lies on self-enforcing relational contracts, where the size of the subjectively determined bonus is restricted by the principal's incentive to give low evaluations despite good performance in order to save on bonus payments. Thiele (2013) shows that in this setting delegation of subjective performance evaluation to a supervisor also entails low-powered incentives, in order to prevent collusion. Following Tirole (1986), collusion is the main issue studied in static three-tier hierarchy models; for overviews see Laffont and Rochet (1997) and Mookherjee (2013). We assume away the problem of collusion in our static model by assuming that side-contracts are not enforceable.

Most related to our work are Laffont (1990) and Vafaï (2002, 2010), who study abuse of authority by the supervisor. In a setting with verifiable information on team output, Laffont (1990) shows that a supervisor can extort side-payments from her subordinates by threatening to misreport their individual performance. In response, the principal optimally distorts the agents' incentive pay by reducing the weight on individual performance relative to the weight on team output. In Vafaï (2002), the supervisor can decide to conceal hard information on the agent's performance from the principal. This allows the supervisor to demand a bribe from the agent when performance is good. Preventing this abuse of authority requires that a good performance report yields the same pay for the agent as a report without performance information, which hampers incentive provision. Vafaï (2010) extends the analysis to possible side-payments between the principal and the supervisor. In our setting, abuse of authority arises from the supervisor's soft information on the agent's efforts and materializes in a distorted effort allocation. We analyze how the principal can use imperfect verifiable performance measures to constrain this abuse of authority.

We use our results to contribute to the debate on the relative merit of specialists and generalists in managerial positions. Garicano (2000) considers task allocation between employees. He argues that more able agents should be assigned to higher hierarchical positions, so that lower-level generalists can screen for tasks that can only be properly conducted by specialists. In contrast, Ferreira and Sah (2012) consider communication and argue that generalists should be higher in the hierarchy to facilitate information transmission with lower-level units consisting of (different) specialists. Prasad (2009) argues that in a setting where tasks are complements, generalists are more likely to work on multiple tasks than specialists. He finds supporting evidence among non-academic researchers in the US, for whom the probability of getting management tasks is decreasing in past research success. We complement this work by focussing on supervisors' monitoring role.

We argue that while specialists may have better monitoring ability, they may also have more biased preferences (for instance arising from professional norms). In the absence of verifiable performance measures, this gives a trade-off between the strength of subjective performance incentives and the distortion induced in the agent's efforts. The availability of a verifiable performance measure decreases both the cost of supervisor bias and the benefit of higher monitoring ability. We find that the first effect typically outweighs the second effect, implying that better performance measures increase the attractiveness of specialist supervisors.

In the next section, we set up the model. We analyze benchmark cases in Section 3 and the full model in Section 4. In Section 5, we apply our framework to the choice between specialists and generalists for supervisory positions. Section 6 concludes.

2 The model

We consider a principal-supervisor-agent model in which all players are risk neutral. The principal (P) employs one agent (A) and one supervisor (S). The outside option utility of both the agent and the supervisor is zero, and they are both protected by limited liability such that $w_A \geq 0$ and $w_S \geq 0$, where w_A (w_S) is the total wage payment to the agent (supervisor).[5]

The agent works on two tasks $i \in \{1, 2\}$. The principal values the two tasks equally, and his utility is given by

$$U_P = e_1 + e_2 - w_A - w_S \tag{1}$$

where e_i is the agent's effort in task i. The agent's utility is given by:

$$U_A = w_A - \frac{1}{2}(e_1)^2 - \frac{1}{2}(e_2)^2 \tag{2}$$

The principal cannot observe the agent's efforts. There is a verifiable performance measure of the agent's efforts, which is imperfect in two ways akin to the dimensions of distortion and noise in e.g. Feltham and Xie (1994) and Baker (2002). First, the performance measure is biased towards one of the tasks. The level of measured performance m is given by

$$m(e_1, e_2) = \varphi e_1 + (1 - \varphi) e_2 \tag{3}$$

with $\varphi \in [0, 1]$. Hence, if $\varphi \neq \frac{1}{2}$, the relative importance of the two tasks in determining measured performance differs from the relative valuation of the tasks by the principal. Second, the performance measure is not perfectly reliable. We model this as follows. After choosing efforts, with probability $1 - q$ the agent can costlessly manipulate measured performance into showing any level of m as preferred by the agent. The principal cannot detect this type of gaming.[6] As discussed below, the use of a less reliable performance measure (lower q) implies

[5] Limited liability also implies that the principal cannot sell the firm to either the supervisor or the agent.

[6] The assumption of (un)reliability provides a tractable way of introducing a second dimension of performance measure imperfection into our model (on top of the bias). This allows us to compare our results with earlier work that studies the trade-off between distortion and noise in performance measures. Alternative ways of modeling unreliability yielding equivalent outcomes include a setting where the agent can always costlessly manipulate measured performance while the principal detects gaming with probability q, and a setting where the objective performance measure fails to provide any measured performance with probability $1 - q$.

higher rents for the agent.

The only role of the supervisor is to monitor the agent. With probability p, the supervisor observes effort levels e_1 and e_2 as chosen by the agent. We interpret parameter p as the supervisor's effectiveness in monitoring the agent. The supervisor's information is soft and cannot be made verifiable. The supervisor creates a verifiable report r regarding her assessment of the agent's performance, which can provide a basis for (subjective) performance pay. It follows that the supervisor's report is cheap talk: she can provide any report independent of the agent's actual efforts.[7] Before making her report, the supervisor observes measured performance m.[8] She does not observe directly whether the agent manipulated measured performance (but can infer this when observing effort).

Crucially, we assume that the supervisor cares about the tasks performed by the agent. As discussed in the Introduction, these preferences may stem from career concerns, professional norms, or intrinsic care for the tasks' output. The supervisor's utility is given by:

$$U_S = w_S + \eta e_1 + (1 - \eta)\, e_2 \tag{4}$$

where $\eta \in [0, 1]$. Hence, the supervisor has biased preferences relative to the principal's valuation of tasks whenever $\eta \neq \frac{1}{2}$.[9] Note that the supervisor does not incur monitoring cost.

The above implies that the agent's wage can depend on both measured performance m and the supervisor's report r. For expositional reasons we explicitly distinguish between purely objective performance pay $b(m)$ and subjective performance pay $c(m, r)$, such that the agent's wage equals $w_A(m, r) = b(m) + c(m, r)$.[10] The supervisor's wage could also depend on m. However, we show in Section 4 that the principal optimally abstains from providing incentive

[7] This differs from Tirole (1986) and Vafaï (2010), where the supervisor reports either the true (outcome of the agent's) efforts or reports that she received no information.

[8] Supervisors typically have access to the verifiable performance information when evaluating employees. For instance, Bol (2011) shows that the form supervisors had to fill out contained both verifiable and subjective items, where the items based on verifiable measures came with strict guidelines on how to translate performance into rating. The results are qualitatively similar when assuming that the supervisor does not observe m before determining r.

[9] In our analysis, only the relative weights of the supervisor's preferences over tasks matters, not the absolute level. Hence, we could multiply the last two terms of (4) with the same parameter without affecting results.

[10] On top of this, the principal may offer fixed wage a to the agent, so that $w_A = a + b(m) + c(m, r)$. However, our assumption of limited liability ($w_A \geq 0$) makes that the principal always sets $a = 0$. It also implies that the agent's participation constraint is always fulfilled.

pay to the supervisor. Hence, given limited liability, the principal optimally offers $w_S = 0$, which the supervisor accepts.[11]

Provided that the agent's wage depends on the supervisor's evaluation, the supervisor can make demands to the agent. We assume that side-contracts are not verifiable (neither between the supervisor and the agent nor between the supervisor and the principal).[12] Instead, the supervisor and the agent engage in an implicit agreement. As the supervisor does not bear the cost of the agent's bonus pay, she is ex post indifferent between providing different subjective evaluations. We assume that the supervisor adheres to the implicit agreement. Given that the supervisor's evaluation is cheap talk, her demands have most influence when she promises to provide evaluations such that her report has a maximal effect on the agent's wage. Without loss of generality, we assume that the supervisor demands effort levels $\underline{e_1}$ and $\underline{e_2}$, and promises to provide in return the report that yields the highest wage (given measured performance m). If the supervisor learns that the agent did not adhere to the demand, she provides the report that yields the lowest possible wage. We assume that if the supervisor is unsure about the agent's efforts, which happens when she does not observe effort and measured performance m is in line with demanded performance (which could be the result of manipulation), she provides the report that yields the highest subjectively determined bonus.[13] Effectively, this implies that the supervisor will provide one of two reports. Hence, without loss of generality, we can reduce the set of possible reports to $r \in \{r_G, r_B\}$, where r_G (r_B) denotes a 'good' ('bad') report. Note that the cheap-talk nature of subjective evaluation implies that a demand to disclose the exact effort

[11] Our assumption of supervisor's limited liability ($w_S \geq 0$) implies that the principal cannot acquire the rents from intrinsic utility obtained by the supervisor. We make this assumption to ensure that the optimal contract as designed by the principal is aimed at optimizing the agent's incentives. It implies that when the agent's efforts are verifiable, the principal optimally induces balanced efforts, $e_1 = e_2$. If, instead, the supervisor's participation constraint would be binding, the principal would optimally demand biased efforts to increase the supervisor's intrinsic utility, which facilitates rent extraction by the principal. Apart from this difference, all results derived below are qualitatively similar.

[12] This makes collusion non-sustainable. The bonus is paid to the agent after the supervisor has reported to the principal, implying that the agent has no incentive to transfer part of the bonus to the supervisor. Collusion is studied by e.g. Tirole (1986), Vafaï (2002, 2010), and Thiele (2013).

[13] This assumption implies that ineffective supervision (low p) yields rents to the agent. Alternatively, the supervisor might send any report that yields a lower bonus, including a report that explicitly states that she is unsure about the agent's efforts. Then, the principal could eliminate the agent's rents by offering $w_A(m) = 0$ after any such report. In equilibrium, neither the agent's efforts nor the principal's payoff would depend on p. All other results would be qualitatively similar. This assumption on the supervisor's reporting strategy allows us to study the effects of ineffective supervision.

levels is ineffective. In other words, it is possible to interpret r_G as $r(e = e^P)$ denoting a report which states that the agent has exerted effort e^P as demanded (and rewarded) by the principal. However, the cheap-talk nature of r implies that the supervisor can also report $r(e = e^P)$ when in fact $e \neq e^P$.

The timing of the game is as follows:

1. The principal offers the agent a contract, determining $w_A(m, r)$.
2. The agent accepts or rejects the contract. If the agent rejects, all players receive their outside option payoff.
3. The supervisor demands effort levels $\{\underline{e_1}, \underline{e_2}\}$ from the agent.
4. The agent chooses effort, which is observed by the supervisor with probability p.
5. With probability $1 - q$, the agent can manipulate measured performance m.
6. The supervisor observes m and sends report $r \in \{r_G, r_B\}$.
7. Payoffs are realized.

3 Benchmarks

3.1 Complete information

Suppose the principal can contract on effort directly. Neither the supervisor nor the performance measure have any use in this case. The principal demands the effort levels that maximize his utility (1) subject to the agent's participation constraint $U_A \geq 0$, where U_A is given by (2). Ignoring the supervisor's intrinsic utility, this gives first-best levels of effort $e_1 = e_2 = 1$. The participation constraint of the agent is satisfied by setting $w_A = 1$ if and only if $e_1 = e_2 = 1$ and $w_A = 0$ otherwise. This results in $U_P = 1$ and $U_A = 0$. Hence, in the absence of moral hazard problems, the principal optimally induces the agent to balance effort levels across tasks.

3.2 Pure objective performance pay

In the absence of subjective performance evaluation, the model is a standard multitasking model. Without loss of generality, we assume that the principal offers the agent a fixed bonus b if measured performance (3) is above a specified target, $m \geq \underline{m}$, and no bonus if $m < \underline{m}$. Using backward induction, if the agent gets the opportunity to manipulate measured performance, he secures his bonus by setting $m = \underline{m}$ if his true performance is below $\underline{m}$.

In choosing effort, the agent derives no benefits from outperforming the principal's target $\underline{m}$. Hence, conditional on meeting the principal's target, the agent's optimal effort levels maximize (2) subject to $m\left(e_1, e_2\right) = \underline{m}$. This gives

$$e_1 = \frac{\varphi}{\varphi^2 + (1-\varphi)^2}\underline{m} \tag{5}$$

$$e_2 = \frac{1-\varphi}{\varphi^2 + (1-\varphi)^2}\underline{m} \tag{6}$$

It follows that $e_1 = e_2$ only when $\varphi = \frac{1}{2}$. If $\varphi \neq \frac{1}{2}$, the agent provides more effort on the task that impacts measured performance m most. Given that the agent satisfies $m = \underline{m}$, he optimally chooses an effort combination such that

$$\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi} \tag{PMR}$$

which we refer to as the Performance Measure's Ratio (PMR).

If the agent decides not to meet the performance target $\underline{m}$, he only receives bonus b if he can manipulate measured performance, which happens with probability $1-q$. As this probability is independent of effort, optimal effort is zero. Hence, using (2), the agent chooses to meet the performance target if

$$b - \frac{1}{2}\frac{1}{\varphi^2 + (1-\varphi)^2}\underline{m}^2 \geq (1-q)b \tag{7}$$

which shows that the agent's rents $(1-q)b$ are decreasing in the reliability of the performance measure.

The principal chooses b and $\underline{m}$ to maximize his utility (1) subject to the agent's incentive compatibility constraint (7), which yields the following solution:

$$b = \frac{1}{2}\frac{q}{\varphi^2 + (1-\varphi)^2} \tag{8}$$

$$\underline{m} = q \tag{9}$$

The principal sets a lower bonus when the performance measure is more biased, leading to a reduction in total efforts. The optimal performance target is independent of φ. Measured performance is convex in the distortion through its effect on the agent's effort ratio (PMR). Yet, in determining the optimal performance target, this effect is exactly offset our linear-quadratic framework by the effect of the distortion on the optimal total efforts through bonus b. Equilibrium payoffs are given by

$$U_A^{ob} = \frac{1}{2}\frac{q\,(1-q)}{\varphi^2+(1-\varphi)^2} \tag{10}$$

$$U_P^{ob} = \frac{1}{2}\frac{q}{\varphi^2+(1-\varphi)^2} \tag{11}$$

The principal benefits from having a more effective and better aligned performance measure, as U_P increases in q and decreases in $|\frac{1}{2}-\varphi|$. For a given φ, the sum of U_A and U_P is maximal when $q=1$, reflecting that the principal sacrifices surplus to reduce the agent's rent when $q<1$.

3.3 Pure subjective performance pay

Without objective performance measures, incentive pay is solely based on the supervisor's subjective report r. In the implicit agreement between the supervisor and the agent, the supervisor promises to provide a positive report to the principal if the agent performs at least effort levels $\underline{e_1}$ and $\underline{e_2}$. The agent follows the supervisor's demands when a good report yields a sufficiently higher bonus than a bad report. Given limited liability, the optimal incentive scheme has wage zero after a bad report, $c(r_B)=0$. Below, we denote $c(r_G)=c$.

The least costly way for the agent to satisfy the supervisor's demands is to provide exactly the demanded effort levels. If the agent decides to exert lower effort levels, he only receives the bonus when the supervisor does not observe the agent's efforts, which happens with probability $1-p$. In this case, the best alternative is to provide no effort at all. Hence, the supervisor sets $\underline{e_1}$ and $\underline{e_2}$ to maximize her utility (4), subject to the agent's incentive compatibility constraint:

$$c-\frac{1}{2}\underline{e_1}^2-\frac{1}{2}\underline{e_2}^2 \geq (1-p)\,c \tag{12}$$

The incentive compatibility constraint binds as the supervisor values additional effort. The agent's rents $(1-p)\,c$ decrease in the supervisor's effectiveness p. The supervisor's optimal effort demands are given by:

$$\underline{e_1} = \eta\sqrt{\frac{2pc}{\eta^2+(1-\eta)^2}} \tag{13}$$

$$\underline{e_2} = (1-\eta)\sqrt{\frac{2pc}{\eta^2+(1-\eta)^2}} \tag{14}$$

It follows that $\underline{e_1}$ and $\underline{e_2}$ are increasing in bonus c and in the supervisor's effectiveness p. Both allow the supervisor to make stronger demands. Furthermore, the supervisor induces the agent to focus disproportionately on her preferred task, demanding effort levels such that:

$$\frac{\underline{e_2}}{\underline{e_1}} = \frac{1-\eta}{\eta} \tag{SR}$$

which we will refer to as the Supervisor's preferred effort Ratio (SR).

It follows from (13) and (14) that for given c and p, the principal's valuation of implemented efforts, $e_1 + e_2$, is higher when the supervisor's preferences are more aligned with the principal, i.e. when η is closer to $\frac{1}{2}$. Hence, a more aligned supervisor induces higher total effort for a given bonus, which benefits the principal.

The principal chooses c to maximize utility (1), taking into account the supervisor's effort demands (13) and (14). The optimal subjective bonus is given by:

$$c = \frac{1}{2}\frac{p}{\eta^2+(1-\eta)^2} \tag{15}$$

The bonus is increasing in p, as more effective supervisors leave fewer rents to the agent. The bonus decreases in the disalignment between the supervisor and the principal, as more disaligned supervisors use their discretionary power to induce more distorted effort levels.[14] Equilibrium

[14]For this effect, it is crucial that the supervisor's information is soft. If, instead, the supervisor would obtain hard information on the agent's effort, the principal could condition the payout of the bonus to the agent on the supervisor's report showing balanced efforts. Combined with a bonus for the supervisor for reporting the (hard) information, the principal could induce balanced efforts, independent of the supervisor's preferences over tasks. Subjective performance evaluation in firms as described in the literature (see footnote 1) seems well in line with the assumption that supervisors base their evaluation (partially) on soft information.

effort levels are given by:

$$e_1 = \frac{\eta p}{\eta^2 + (1-\eta)^2} \tag{16}$$

$$e_2 = \frac{(1-\eta)\,p}{\eta^2 + (1-\eta)^2} \tag{17}$$

Hence, if $p \neq 1$ or $\eta \neq \frac{1}{2}$, effort provision under subjective pay for performance is below first-best levels. The agent's and principal's equilibrium payoffs are given by:

$$U_A^{sub} = \frac{1}{2}\frac{p\,(1-p)}{\eta^2 + (1-\eta)^2} \tag{18}$$

$$U_P^{sub} = \frac{1}{2}\frac{p}{\eta^2 + (1-\eta)^2} \tag{19}$$

Note that the payoffs of the agent and the principal are identical to their payoffs when using pure objective pay for performance (see (10) and (11)), with $p = q$ and $\eta = \varphi$ (or $\eta = 1 - \varphi$). The principal responds to biased supervision by reducing subjective performance pay, as in case of incentive pay based on the incongruent verifiable performance measure. Hence, biased supervision is as harmful for the principal (and the agent) as an incongruent performance measure.[15]

In equilibrium, the supervisor's utility is independent of her bias: $U_S = p$. The higher utility attained by a more biased supervisor for a given bonus c is exactly offset by the reduction in the bonus set by the principal. Strictly speaking, this implies that allowing for collusion through side-contracting between the agent and the supervisor would not affect the outcomes above. Given a side-contract stipulating a payment from the agent to the supervisor following report r_G, the supervisor would ex post always provide report r_G independent of the agent's efforts. This eliminates the agent's incentive to exert effort. Hence, as the supervisor's equilibrium utility is (weakly) higher than the bonus for the agent, the supervisor would not agree ex ante to send report r_G in exchange for receiving part of c.However, none of the results above

[15] If the supervisor's bias is private information, such that the principal only knows the distribution of η but not the current supervisor's bias, a similar result obtains. From effort levels (16) and (17), it follows that given bonus c, the principal's payoff is decreasing and convex in supervisor bias $|\frac{1}{2} - \eta|$. For any distribution of bias $|\frac{1}{2} - \eta|$, the principal optimally sets a bonus below the optimal bonus in case of a supervisor with average bias. Hence, uncertainty about supervisor bias leads to (even) weaker subjective performance pay.

rely on the intensity of the supervisor's intrinsic preferences (cf footnote 10). Yet, if these intrinsic preferences are relatively weak, collusion would be mutually beneficial to the supervisor and the agent. As noted by Tirole (1986), the combination of side-contracting and subjective performance evaluation based on soft information renders subjective performance pay useless. Where Tirole (1986) studies performance evaluation with (collusive) side-contracting under hard information, we study the complementary situation of performance evaluation under soft information in the absence of side-contracting.

3.4 Subjective versus objective performance pay

Suppose that the principal must choose between implementing either objective or subjective performance pay. Comparing (11) and (19) shows that the principal is better off using subjective performance pay when:

$$\frac{p}{q} > \frac{\eta^2 + (1-\eta)^2}{\varphi^2 + (1-\varphi)^2} \tag{20}$$

Even if the supervisor is more biased than the performance measure, subjective performance pay is still preferred when the supervisor is sufficiently more effective than the performance measure, and vice versa. Figure 1 gives the preferred type of performance pay for given values of η, φ, p, and q.[16]

4 Combining subjective and objective performance pay

Now consider the case where both objective and subjective performance measures are available. We first establish several general features of the optimal contract.

Lemma 1 *The optimal contract is a forcing contract, where the agent receives compensation unless either objectively measured performance m differs from a pre-determined target $\underline{m}$ or the supervisor reports bad performance r_B.*

Proof. The proof is given in the Appendix. ■

[16] It is also possible to interpret Figure 1 as a comparison between supervisors who differ in bias and effectiveness.

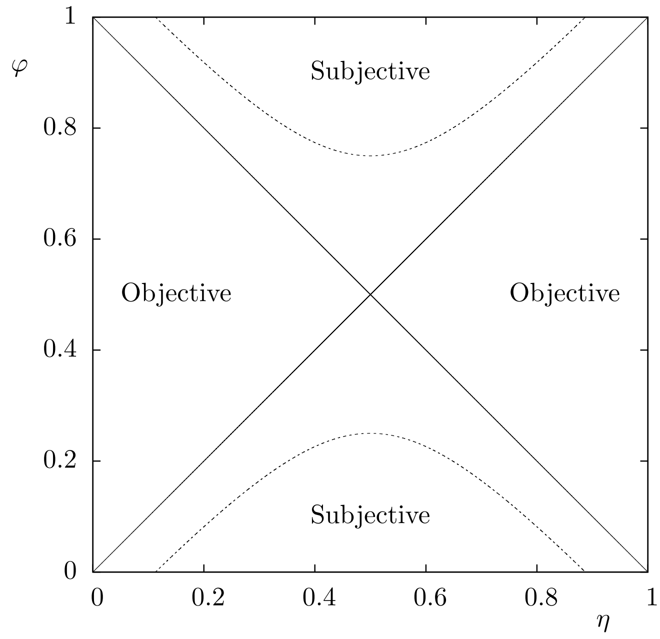

Figure 1: Preferred pure incentive mechanism if $p = q$ (solid) and $p = \frac{4}{5}q$ (dashed).

To understand Lemma 1, consider first a supervisor with the same bias as the objective performance measure, $\eta = \varphi$. Combining subjective and objective performance pay would not affect the distortion in the effort ratio. Still, conditioning the agent's pay on both measures of performance yields the highest probability of punishing a shirking agent. The forcing contract minimizes the agent's rents. This benefit of a forcing contract carries over when $\eta \neq \varphi$. Then, the supervisor uses any subjectively determined bonus to pull the agent towards her own preferred effort ratio (SR). By conditioning the payout of the subjective bonus on objectively measured performance, the principal also constrains the supervisor. Denying the agent the bonus despite a positive report if measured performance does not meet a pre-determined target $\underline{m}$ reduces the supervisor's power to demand efforts that fall short of the target.

A direct implication of Lemma 1 is that there is no clear separation between objective and subjective performance pay. Optimally, the agent's compensation depends on both objectively measured performance as well as the supervisor's subjectively determined report. Without loss of generality, we assume that the agent's contract has subjective performance pay $c(m, r_B) = 0$ and $c(m, r_G) > 0$, where the bonus is forfeit when measured performance differs from the target:

$b(m = \underline{m}) = 0$ and $b(m \neq \underline{m}) = -c(m, r)$ for all m and r.[17] For ease of notation, below we denote $c(m, r_G) = c$.

Before deriving the optimal contract, we first establish that not all possible effort allocations can be implemented through the combination of objective and subjective performance pay.

Lemma 2 *If $\eta > \varphi$ ($\eta < \varphi$), effort allocations such that $\frac{e_2}{e_1} > \frac{1-\varphi}{\varphi}$ ($\frac{e_2}{e_1} < \frac{1-\varphi}{\varphi}$) cannot be implemented.*

Proof. In this proof, we focus on the case where $\eta > \varphi$. The case $\eta < \varphi$ is the mirror image. By Lemma 1, the agent receives bonus c unless $r = r_B$ or $m \neq \underline{m}$. Several outcomes are possible. First, the agent can ignore both the principal's target and the supervisor's request. Then, optimal effort $e_1 = e_2 = 0$. Second, if the agent ignores the supervisor's request but adheres to the principal's target $m = \underline{m}$, the agent's optimal effort allocation has $\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi}$ (see (PMR)). Third, the supervisor can ignore the principal's target and induce the agent to follow her request. When ignoring the target, the supervisor's optimal effort allocation is such that $\frac{e_2}{e_1} = \frac{1-\eta}{\eta} < \frac{1-\varphi}{\varphi}$ (see (SR)). Lastly, the agent can follow the supervisor's request while the supervisor adheres to the principal's target. Then, using (2), the agent's incentive compatibility constraint is given by:

$$c - \frac{1}{2}\underline{e_1}^2 - \frac{1}{2}\underline{e_2}^2 \geq \max\left\{(1-p)(1-q)c, (1-p)c - \frac{1}{2}\frac{1}{\varphi^2 + (1-\varphi)^2}\underline{m}^2\right\} \tag{21}$$

The first term in braces is the agent's expected utility when $e_1 = e_2 = 0$. The second term in braces is the agent's expected utility when using the performance measure effort ratio (PMR) to reach $m = \underline{m}$ at lowest cost as given by effort levels (5) and (6). To induce $\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi}$, (21) requires that $c = \frac{1}{(p+(1-p)q)}\frac{\underline{m}^2}{2\left(\varphi^2+(1-\varphi)^2\right)}$. Keeping $\underline{m}$ constant, an increase in c makes different effort allocations on $m = \underline{m}$ feasible. As $\eta > \varphi$, the supervisor prefers effort allocations on $m = \underline{m}$ with more e_1 over effort allocations with less e_1. Hence, for any $c > \frac{1}{(p+(1-p)q)}\frac{\underline{m}^2}{2\left(\varphi^2+(1-\varphi)^2\right)}$, the supervisor optimally requests an effort allocation on $\underline{m}$ such that $\frac{e_2}{e_1} \leq \frac{1-\varphi}{\varphi}$. Summarizing, none of these outcomes has $\frac{e_2}{e_1} > \frac{1-\varphi}{\varphi}$. ■

[17] Alternatively, the contract could have verifiable performance pay $b(m \neq \underline{m}) = 0$ and $b(\underline{m}) > 0$, with the 'disqualifier' that the bonus is forfeit after a bad subjective evaluation: $c(m, r_B) = -b(m)$ and $c(m, r_G) = 0$.

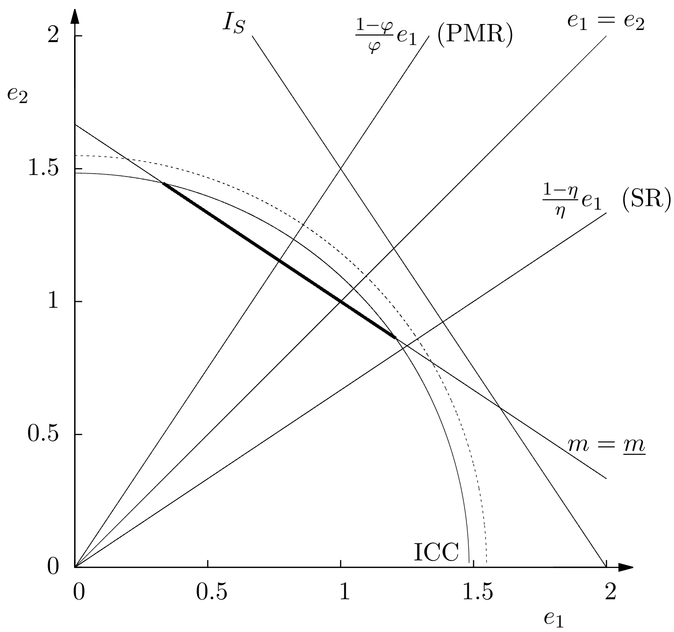

Figure 2: An example with $\eta = \frac{3}{5}$ and $\varphi = \frac{2}{5}$. The feasible set of demands for the supervisor is given by the thicker line segment, which is always centered around (PMR). Lowering m or raising the ICC can lengthen this feasible set.

Lemma 2 is illustrated by Figure 2, for the case where $\eta > \varphi$. The figure depicts the $m = \underline{m}$ performance target which runs orthogonal to the line representing the performance measure's effort ratio (PMR), as well as the set of points where the agent's incentive compatibility constraint (ICC) given by (21) is binding. The supervisor can only implement effort allocations on the $m = \underline{m}$ line below the ICC, as indicated by the thicker line segment. This line segment always includes effort ratio $\frac{1-\varphi}{\varphi}$ (PMR), as this yields $m = \underline{m}$ at minimal effort cost. The supervisor's indifference lines run orthogonal to the line representing the supervisor's preferred effort ratio (SR). Hence, given $\eta > \varphi$ and $m = \underline{m}$, the supervisor prefers $\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi}$ over any $\frac{e_2}{e_1} > \frac{1-\varphi}{\varphi}$. This implies that the principal cannot induce the supervisor to request an effort allocation beyond (PMR) from the supervisor's perspective.

Lemma 2 underlines an important difference between objective and subjective performance evaluation. If the supervisor would be a second verifiable performance measure with bias $\eta \neq \varphi$, any effort allocation would have been feasible (Feltham and Xie, 1994). This holds even when the two measures are biased in the same direction, by placing a negative weight on the most-biased measure. In case of subjective performance evaluation, however, a negative weight is ineffective.

The supervisor would react by switching good and bad evaluations, allowing her to still induce the agent to choose an effort allocation closer to (SR). As a consequence, congruence is not feasible when the supervisor is more biased than the verifiable performance measure ($\eta > \varphi > \frac{1}{2}$ or $\frac{1}{2} > \varphi > \eta$). In this situation, the least-biased effort ratio that can be implemented is given by (PMR), $\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi}$.[18]

Using Lemma's 1 and 2, we can derive the optimal contract. We focus on the case where $\eta > \varphi$, the case $\eta < \varphi$ is the mirror image. Proposition 1 derives the conditions under which the principal optimally implements non-distorted efforts.

Proposition 1 *Given $\eta > \varphi$, the optimal contract induces balanced effort $\frac{e_2}{e_1} = 1$ if and only if (i) $\varphi \leq \frac{1}{2}$ and (ii) $\frac{1}{(1-2\eta)^2}\frac{q}{(1-q)} \geq p \geq \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$.*
Under these conditions, the optimal contract has $c = \underline{m} = (p + (1-p)\,q)$, $b(m = \underline{m}) = 0$, and $b(m \neq \underline{m}) = -c$.

Proof. The proof is given in the Appendix. ■

Despite having an incongruent performance measure and a biased supervisor, under some conditions the principal optimally induces congruent efforts by combining objective and subjective performance pay. Figure 2 helps to build intuition for this result. By adjusting c and $\underline{m}$, the verifiable performance measure allows the principal to affect the set of effort allocations the supervisor can demand from the agent (the thicker line segment). This restricts the supervisor, who cannot simply implement her most-preferred effort ratio (SR). At the same time, subjective performance evaluation can be used to implement effort ratio's different from bias (PMR) inherent in the verifiable performance measure. When conditions (i) and (ii) are met, the outcome and the optimal contract are as if the principal has access to an unbiased performance measure that is as effective as the supervisor and the verifiable measure combined.

Condition (i) follows from Lemma 2: it is not feasible to induce the supervisor to implement balanced efforts when she is more biased than the verifiable performance measure. Condition (ii) indiciates that if the effectiveness of the supervisor and the performance measure are too

[18] In line with Feltham and Xie (1994), it is possible to implement effort ratio's $\frac{e_2}{e_1} < \frac{1-\eta}{\eta}$ when $\eta > \varphi$ by placing a negative weight on the verifiable measure. This is optimally implemented by setting bonus $c = 0$ for $m > \underline{m}$.

different, it is not optimal to induce non-distorted efforts. If the supervisor's effectiveness p is too low, the agent is tempted to feign satisfying the supervisor's demands by generating $m = m(\underline{e_1}, \underline{e_2})$ at lower cost along the PMR. The benefits of this deviation for the agent are larger when the performance measure is more biased. In contrast, if the performance measure is relatively unreliable, the supervisor would ignore the performance target $\underline{m}$. If the probability that the agent can manipulate measured performance is high, the principal is unlikely to detect the deviation. Deviating is more beneficial for more biased supervisors. The following Propositions describe the optimal outcomes when the conditions in Proposition 1 are not met. Proposition 2 considers the case where $\varphi > \frac{1}{2}$, while Propositions 3 and 4 consider the cases where p and q are too different.

Proposition 2 *Given $\eta > \varphi > \frac{1}{2}$, the principal optimally induces efforts along the PMR ($\frac{e_2}{e_1} = \frac{1-\varphi}{\varphi}$), if and only if $p \leq \frac{q}{(1-q)} \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}$.*
Under these conditions, the optimal contract has $c = \frac{(p+(1-p)q)}{2\left(\varphi^2+(1-\varphi)^2\right)}$, $\underline{m} = (p + (1-p)\,q)$, $b(m = \underline{m}) = 0$, and $b(m \neq \underline{m}) = -c$.

Proof. The proof is given in the Appendix. ■

Figure 3 helps in explaining this result. Given $\eta > \varphi > \frac{1}{2}$, the supervisor uses any available room to move away from the PMR to distort the agent's efforts even further away from the principal's objectives. Hence, it is optimal for the principal to constrain the supervisor to the PMR. Compared to the case where $\varphi \leq \frac{1}{2}$, the principal demands the same measured performance $\underline{m}$, but sets a smaller bonus. The resulting outcome is as if the principal has access to a biased performance measure as effective as the supervisor and the verifiable measure combined. The only constraint on the outcome is that the supervisor should not be tempted to ignore the target. As before, this happens when the verifiable measure is sufficiently unreliable relative to the effectiveness of the supervisor, and is more likely when the PMR and the SR are more apart.[19]

[19] As the contract in Proposition 2 implements effort levels that meet the principal's performance target $\underline{m}$ at minimal effort costs, the agent never benefits from choosing different effort levels while still generating measured performance $\underline{m}$.

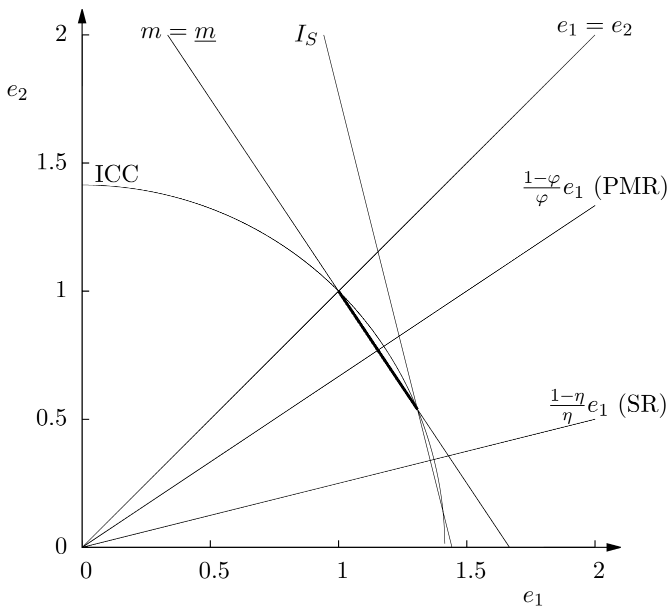

Figure 3: An example with $\eta = \frac{4}{5}$ and $\varphi = \frac{3}{5}$. The supervisor could induce balanced efforts here, but prefers to induce highly distorted efforts. Hence, the principal is better off raising m to the intersection between (PMR) and the ICC.

Combining Propositions 1 and 2, it follows that if p and q are sufficiently large, the supervisor's bias only affects whether balanced efforts can be implemented. Other than that, neither the optimal contract nor equilibrium efforts depend on the exact bias of the supervisor.

Proposition 3 *Suppose $\eta > \varphi$ and $\varphi \leq \frac{1}{2}$. If $p < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, the optimal contract induces unbalanced efforts biased towards the PMR, $1 < \frac{e_2}{e_1} \leq \frac{1-\varphi}{\varphi}$.*
Under these conditions, the optimal contract has $c = \frac{\left(\sqrt{(1-p)q}+(1-2\varphi)\sqrt{p}\right)^2}{2\left(\varphi^2+(1-\varphi)^2\right)}$, $\underline{m} = (1-p)q + (1-2\varphi)\sqrt{p(1-p)q}$, $b(m = \underline{m}) = 0$, and $b(m \neq \underline{m}) = -c$.

Proof. The proof is given in the Appendix. ■

If the supervisor is ineffective, the agent is tempted to feign compliance to the supervisor's request by meeting performance target $m = \underline{m}$ at minimal cost using effort ratio PMR. Feigning compliance is particularly attractive for the agent when the performance measure is both highly effective and highly biased while the supervisor is weak. Anticipating the agent's incentive to deviate, the supervisor is forced to shift her requested effort levels closer to the PMR. This increases the agent's rents.

The principal optimally responds to this inefficiency in two ways. First, he increases the performance target $\underline{m}$. While this forces the supervisor to request a biased effort ratio closer to the PMR, it also reduces the agent's rents when feigning compliance, which makes following the supervisor's request relatively more attractive. Second, as the implemented effort levels remain biased toward the PMR, the principal lowers the bonus. Hence, compared to the case with a more effective supervisor (Proposition 1), the bonus is smaller while the performance target is higher.

The imbalance in the effort levels is decreasing in supervisor effectiveness p and increasing in the bias of the verifiable performance measure $|\frac{1}{2} - \varphi|$. Provided that the supervisor is not more biased than the verifiable performance measure ($\eta > \varphi$ and $\varphi \leq \frac{1}{2}$), the supervisor's bias is irrelevant given that she is weak. This differs when the supervisor is relatively strong, as shown in Proposition 4.

Proposition 4 *Suppose $\eta > \varphi$. If $p > \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$ and $\varphi \leq \frac{1}{2}$, or if $p > \frac{q}{(1-q)}\frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}$ and $\varphi > \frac{1}{2}$, the optimal contract induces unbalanced efforts biased towards the SR.*
Under these conditions, the optimal contract has $c = \frac{\left(\sqrt{p(1-q)}+|1-2\eta|\sqrt{q}\right)^2}{2\left(\eta^2+(1-\eta)^2\right)}$, $b(m = \underline{m}) = 0$, and $b(m \neq \underline{m}) = -c$.
If $\eta > \frac{1}{2}$, $\underline{m} = \lambda\left((\eta\varphi + (1-\eta)(1-\varphi))\sqrt{p(1-q)} + (\eta - \varphi)\sqrt{q}\right)$ and $\frac{1-\eta}{\eta} \leq \frac{e_2}{e_1} < 1$.
If $\eta < \frac{1}{2}$, $\underline{m} = \lambda\left((\eta\varphi + (1-\eta)(1-\varphi))\sqrt{p(1-q)} - (\eta - \varphi)\sqrt{q}\right)$ and $1 < \frac{e_2}{e_1} \leq \frac{1-\eta}{\eta}$.

Proof. The proof is given in the Appendix. ■

Proposition 4 shows that an unreliable performance measure not only hampers the provision of incentives to the agent, but also makes monitoring the supervisor difficult. Strong supervisors use their information advantage to induce efforts that are biased towards their most-preferred task.[20] Figure 4 depicts this situation. The solid ICC curve gives the effort allocations the supervisor could induce provided that the penalty $b(m \neq \underline{m})$ would never be incurred, while the dashed ICC gives all implementable effort allocations if penalty $b(m \neq \underline{m})$ is incurred with probability q. From the supervisor's perspective, the best-feasible effort allocation implementing

[20]Note that if $q = 0$, the equilibrium is identical to the equilibrium derived in subsection 3.3 where only subjective performance evaluation was available. For $q = 0$, the (irrelevant) performance target $\underline{m}$ simply gives the performance as measured given the agent's efforts.

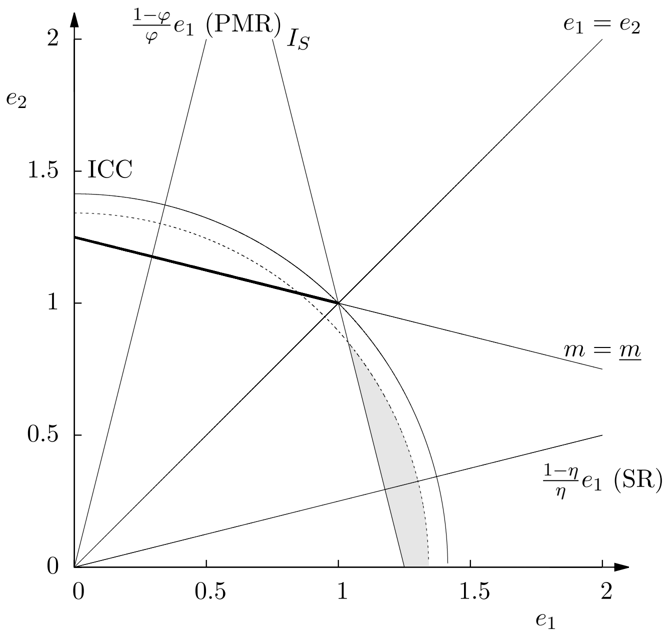

Figure 4: An example with $\eta = \frac{4}{5}$ and $\varphi = \frac{1}{5}$. For relatively low q, it is feasible and optimal for the supervisor to ignore $m = \underline{m}$, as shown by the shaded area.

$m = \underline{m}$ is the rightmost point on the thick segment of the $m = \underline{m}$ line. However, given that the supervisor is very effective compared to the verifiable measure, she can induce efforts off the $m = \underline{m}$ line that yield her higher utility, as depicted by the shaded area. Optimally, she would induce the agent to choose the effort levels determined by the intersection of the SR and the dashed ICC.

Anticipating the supervisor's incentive to deviate, the principal optimally adjusts the contract. First, given bonus c, the principal adjusts the performance target $\underline{m}$ such that the supervisor can induce the agent to meet $m = \underline{m}$ with an effort allocation closer to SR. Given that $\eta > \varphi$, this adjustment is upward (downward) when $\eta < \frac{1}{2}$ ($\eta > \frac{1}{2}$). Given q, this distortion in equilibrium efforts is increasing in p. Second, the principal optimally sets a smaller bonus in response to the distortion. However, this does not imply that the principal would prefer a less effective supervisor when verifiable performance measures are unreliable. Effective supervisors also reduce the agent's rents, which in turn makes implementing higher efforts more attractive to the principal. This positive effect on efficiency dominates the negative effect of a more biased outcome. Given supervisor effectiveness p, an increase in the supervisor's bias harms the principal.

The following proposition gives, for all cases considered above, the comparative statics on the optimal level of bonus pay and on the principal's payoff, which are identical in our linear-quadratic framework.

Proposition 5 *Given $\eta > \varphi$, comparative statics on the optimal level of bonus pay and the principal's equilibrium payoff are as follows:*
(i) $\frac{\partial U_P}{\partial p} = \frac{\partial c}{\partial p} \geq 0$, *with equality only if* $q = 1$.
(ii) $\frac{\partial U_P}{\partial q} = \frac{\partial c}{\partial q} \geq 0$, *with equality only if* $p = 1$.
(iii) $\frac{\partial U_P}{\partial \varphi} = \frac{\partial c}{\partial \varphi} \geq 0$ *if* $\varphi \leq \frac{1}{2}$, *with equality only if* $p \geq \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, *and* $\frac{\partial U_P}{\partial \varphi} = \frac{\partial c}{\partial \varphi} \leq 0$ *if* $\varphi > \frac{1}{2}$, *with equality only if* $p > \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$.
(iv) $\frac{\partial U_P}{\partial \eta} = \frac{\partial c}{\partial \eta} \geq 0$ *if* $\eta \leq \frac{1}{2}$, *with equality only if* $p \leq \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$, *and* $\frac{\partial U_P}{\partial \eta} = \frac{\partial c}{\partial \eta} \leq 0$ *if* $\eta > \frac{1}{2}$, *with equality only if* $\varphi \leq \frac{1}{2}$ *and* $p \leq \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$ *or if* $\varphi > \frac{1}{2}$ *and* $p < \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$.

Proof. The proof is given in the Appendix. ■

To summarize our findings above, we find that the principal always benefits from more effective performance measurement and supervision. Both reduce the agent's rents and allow for stronger incentive pay. Supervisor ability and reliability of the verifiable measure are substitutes. A biased supervisor need not be detrimental to the principal, provided that the principal has access to a verifiable performance measure that is sufficiently effective and the supervisor is either less biased than the verifiable performance measure or biased towards the other task. However, when the verifiable performance measure is too unreliable, a biased supervisor forces the principal to accommodate to her preferences, leading to lower optimal incentive pay. Similarly, a biased performance measure is not problematic as long as the supervisor is sufficiently effective. If not, the agent's incentive to ignore the supervisor leads to an allocation of efforts biased towards the task that receives most weight in the performance measure.

Combining the results in Proposition 5, it is easily derived that the principal is better off combining subjective and objective performance evaluation compared to using only subjective or only objective performance evaluation. Combined evaluation always reduces the agent's rents and allows for the implementation of higher efforts. Furthermore, if both subjective and objective performance measurement are sufficiently effective, combined evaluation eliminates any bias in effort allocation. If either subjective or objective performance evaluation is not sufficiently effective, combined evaluation may increase the bias. However, the cost implied by this extra bias is outweighed by the benefits of better monitoring. This is in line with Manthei and Sliwka (2014), who provide a subset of managers of local units of a bank in Germany with individual sales figures of their employees, which previously allocated bonus pay purely on the basis of a subjective assessment. This led to an increase in both employee sales activities and in profits, particularly in large branches (where, arguably, monitoring by the manager is more difficult than in small units).[21]

Corollary 5.1 *The principal weakly benefits from combining subjective and objective performance evaluation, and strictly so unless either* $p = 1$ *and* $\eta = \frac{1}{2}$ *or* $q = 1$ *and* $\varphi = \frac{1}{2}$.

So far, we have ignored the possibility of providing incentive pay to the supervisor. This would have to be based on measured performance m. The cheap-talk nature of the subjective report r implies that incentive pay based on r would induce the supervisor ex post to send the report that yields her the highest pay. Anticipating this, the agent would ignore any requests from the supervisor, which destroys the value of subjective performance evaluation altogether. Incentive pay based on m effectively changes the supervisor's relative preferences over tasks, by drawing them closer to (PMR). Proposition 5 shows that a change in supervisor's preferences affects the principal's payoff only in the case – covered in Proposition 4 – where the verifiable performance measure is relatively weak. However, given the supervisor's limited liability, providing a bonus to the supervisor is highly costly precisely when the verifiable measure is weak. Deriving the optimal non-linear incentive pay for the supervisor shows that even in the case of Proposition 4 the cost of incentive pay are larger than the benefits, yielding the following result.[22]

Corollary 5.2 *The principal optimally provides no incentive pay to the supervisor.*

[21] We have assumed that the (fixed) cost of obtaining both the objective and the subjective performance measure is zero. If these costs are positive, a trade-off naturally arises.

[22] The proof is a straightforward extension of the proof of Proposition 4, and is avaliable on request.

5 Generalists versus specialists as supervisors

We use our results to discuss the principal's trade-off when choosing between different supervisors. In particular, we consider the choice between a specialist and a generalist. We assume that a specialist is more effective in monitoring the agent than a generalist, $p_S > p_G$, where subscript S (G) denotes specialist (generalist). However, the specialist has stronger preferences regarding the execution of the tasks than the generalist. We assume that the generalist has unbiased preferences, $\eta_G = \frac{1}{2}$, while the specialist has biased preferences $\eta_S \neq \frac{1}{2}$. This combination of expertise and bias corresponds to findings by Li (2013), who looks at decisions made by reviewers of grant proposals at the US National Institute of Health. She finds that reviewers are better able to judge the quality of proposals in their own area, but that they are also biased in favour of proposals in their own area. While the relation between grant reviewers and potential grant recipients differs from the relation between managers and employees, the reviewers decide about resource allocation across different fields, much like the supervisor in our model.

To determine whether the principal prefers a generalist or a specialist in the absence of verifiable performance measures, we can directly use the principal's payoff under pure discretionary pay (19). It follows that the principal prefers the specialist manager if

$$\frac{p_S}{p_G} > 2\left((1-\eta_S)^2 + \eta_S^2\right) \tag{22}$$

Hence, the principal tolerates the bias of a specialist only if she is sufficiently more effective in monitoring.

Now suppose that the principal has access to a verifiable performance measure. Without loss of generality, we focus on the case where $\eta_S > \frac{1}{2}$ and $\eta_S > \varphi$. Better performance measurement (higher q) affects the choice between the generalist and the specialist in two ways. First, the marginal benefit to the principal of a better supervisor is smaller, as the effectiveness of subjective and objective performance measurement are substitutes. Second, the principal can use the verifiable performance measure to neutralize or at least mitigate the bias that is induced by the specialist supervisor. The latter effect dominates, unless the specialist's bias cannot be fully eliminated.

Combining Propositions 1, 3, and 5, we have that if $\varphi \leq \frac{1}{2}$ and $p_S < \frac{1}{(1-2\eta_S)^2}\frac{q}{(1-q)}$ the supervisor's bias is irrelevant while the principal benefits from more effective supervision. It follows that the principal prefers the specialist: when the verifiable performance measure is sufficiently reliable and biased towards the opposite task as compared to the specialist, the principal can neutralize the specialist's bias while obtaining the benefits from better supervision. If the objective measure is relatively unreliable, the principal is forced to accommodate towards the specialist's bias. If the bias of the specialist is sufficiently strong while she is only slightly more effective, the principal prefers the generalist (in the limit where $q = 0$, this condition is given by (22)). Interestingly, if $\varphi > \frac{1}{2}$ and the verifiable measure is sufficiently reliable, the bias induced by the specialist is equal to the bias in the verifiable performance measure. In this case, a more aligned objective performance measure (φ closer to $\frac{1}{2}$) makes the specialist more attractive, whereas a more reliable objective performance measure (higher q) increases the relative attractiveness of the generalist. These results are summarized in the following proposition.

Proposition 6 *Consider the principal's choice of supervisor between a specialist and a generalist, where $\eta_S > \eta_G = \frac{1}{2}$ and $p_S > p_G$. The principal prefers the specialist over the generalist, unless (i) $\eta_S > \varphi > \frac{1}{2}$ and $p_S < \frac{(\eta_S\varphi+(1-\eta_S)(1-\varphi))^2}{(\eta_S-\varphi)^2}\frac{q}{(1-q)}$ or (ii) $p_G > \frac{1}{(1-2\eta_S)^2}\frac{q}{(1-q)}$. In case (i), if $p_G \geq \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, the principal prefers the generalist if and only if $p_S - p_G < (1-2\varphi)^2\left(\frac{q}{(1-q)} + p_G\right)$, while if $p_G < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, the principal prefers the generalist if and only if $p_S - p_G < \frac{2\sqrt{p_G}}{(1-q)}\left((2\varphi-1)\sqrt{(1-p_G)q} - 2\varphi(1-\varphi)\sqrt{p_G}\right)$. In case (ii), the principal prefers the generalist if and only if $p_S - p_G < \frac{\left((2\eta-1)\sqrt{p_S}-\sqrt{\frac{q}{(1-q)}}\right)^2}{2\left(\eta^2+(1-\eta)^2\right)}$.*

Proof. The proof is given in the Appendix. ■

6 Concluding remarks

Supervisors can (ab)use their discretion in determining subjective performance evaluations by directing subordinates towards activities that are not valued by the organization. Biased supervision is costly and reduces the optimal strength of subjective performance pay, as in case of incongruent verifiable performance measurement. This effect can be eliminated when a (possibly incongruent) verifiable performance measure is available, by using performance targets in the agent's contract to constrain the supervisor. However, we have shown that, in contrast to incentive pay based on multiple incongruent verifiable measures, biased supervision remains costly when the supervisor is more biased than the verifiable measures.

We have derived the optimal contract assuming that the supervisor's bias and ability are observed by the principal. If the supervisor's type is unobservable and supervisors self-select into organizations, a given contract is most attractive to supervisors with high-ability and with a bias close to the bias of the verifiable performance measure. This implies that in determining performance targets and the supervisor's discretion, the principal faces a trade-off between attracting aligned but low-ability supervisors and attracting high-ability but more biased supervisors. The effects of (incentive) wages on the self-selection of workers based on intrinsic motivation and/or ability is studied by e.g. Handy and Katz (1998), Besley and Ghatak (2005), Delfgaauw and Dur (2008, 2010), Prendergast (2007), and Dal Bo et al. (2013). How self-selection of managers is affected by the degree of discretion over their subordinates' activities is an interesting question that we leave for future work.

References

[1] Akerlof, G., and R. Kranton (2005), Identity and the Economics of Organizations, *Journal of Economic Perspectives*, 19(1), pp. 9-32.

[2] Asch, B. (1990), Do Incentives Matter? The Case of Navy Recruiters, *Industrial and Labor Relations Review,* 43, pp. 89S-106S.

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

[4] Baker, G. (2002), Distortion and Risk in Optimal Incentive Contracts, *Journal of Human Resources*, 37(4), pp. 728-751.

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

[6] Besley, T., and M. Ghatak (2005), Competition and Incentives with Motivated Agents, *American Economic Review,* 95(3), pp. 616-636.

[7] Bol, J. (2011), The Determinants and Performance Effects of Managers' Performance Evaluation Biases, *The Accounting Review*, 82, pp. 1549-1575.

[8] Breuer, K., P. Nieken, and D. Sliwka (2013), Social Ties and Subjective Performance Evaluations: An Empirical Investigation, *Review of managerial Science*, 7(2), pp. 141-157.

[9] Budde, J. (2007), Performance Measure Congruity and the Balanced Scorecard, *Journal of Accounting Research*, 45(3), pp. 515-539.

[10] Bull, C. (1987), The Existence of Self-Enforcing Implicit Contracts, *Quarterly Journal of Economics*, 102, pp. 147-159.

[11] Burgelman, R.A. (1994), Fading Memories: A Process Theory of Strategic Business Exit in Dynamic Environments, *Administrative Science Quarterly,* 39(1), pp. 39-56.

[12] Courty, P., and G. Marschke (2004), An Empirical Investigation of Gaming Responses to Explicit Performance Incentives, *Journal of Labor Economics* 22(1), pp. 23-56.

[13] Dal Bo, E., F. Finan, and M. Rossi (2013), Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service, *Quarterly Journal of Economics*, 128(3), pp. 1169-1218.

[14] Datar, S., S. Kulp, and R. Lambert (2001), Balancing Performance Measures, *Journal of Accounting Research*, 39(1), pp. 75-92.

[15] Delfgaauw, J., and R. Dur (2008), Incentives and Workers' Motivation in the Public Sector, *Economic Journal,* 118(525), pp. 171-191.

[16] Delfgaauw, J., and R. Dur (2010), Managerial Talent, Motivation, and Self-selection into Public Management, *Journal of Public Economics,* 94(9-10), pp. 654-660.

[17] Dur, R., and J. Tichem (2015), Altruism and Relational Contracts in the Workplace, *Journal of Economics and Management Strategy,* 24(3), pp. 485-500.

[18] Eccles, R. and D. Crane (1988), *Doing Deals: Investment Banks at Work.* Boston: Harvard Business School Press.

[19] Efendi, J., A. Srivastava, and E. Swanson (2007), Why Do Corporate Managers Misstate Financial Statements? The Role of Option Compensation and Other Factors, *Journal of Financial Economics,* 85(3), pp. 667-708.

[20] Feltham, G., and J. Xie (1994), Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations, *The Accounting Review*, 69(3), pp. 429-453.

[21] Ferreira, D., and R.K. Sah (2012), Who Gets To the Top? Generalists versus Specialists in Organizations, *RAND Journal of Economics*, 43(4), pp. 577-601.

[22] Garicano, L. (2000), Hierarchies and the Organization of Knowledge in Production, *Journal of Political Economy*, 108(5), pp. 874-904.

[23] Gibbs, M. (1995), Incentive Compensation in a Corporate Hierarchy, *Journal of Accounting and Economics*, 19(2), pp. 247-277.

[24] Gibbs, M., K.A. Merchant, W.A. Van der Stede, and M.E. Vargus (2004), Determinants and Effects of Subjectivity in Incentives, *The Accounting Review*, 79(2), pp. 409-436.

[25] Giebe, T., and O. Gurtler (2012), Optimal Contracts for Lenient Supervisors, *Journal of Economic Behavior & Organization*, 81, pp. 403-420.

[26] Goldman, E., and S. Slezak (2006), An Equilibrium Model of Incentive Contracts in the Presence of Information Manipulation, *Journal of Financial Economics,* 80, pp. 603–626.

[27] Guth, D.G. and I.C. MacMillan (1986), Strategy Implementation versus Middle Management Self-interest, *Strategic Management Journal,* 7, pp. 313-327.

[28] Handy, F. and E. Katz (1998), The Wage Differential between Nonprofit Institutions and Corporations: Getting More by Paying Less?, *Journal of Comparative Economics,* 26, pp. 246-261.

[29] Holmström, B., and P. Milgrom (1991), Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics and Organization*, 7, pp. 24-52.

[30] Holthausen, R., D. Larcker, and R. Sloan (1995) Annual Bonus Schemes and the Manipulation of Earnings. *Journal of Accounting and Economics,* 19, pp. 29–74.

[31] Jawahar, I., and C. Williams (1997), Where all the Children are Above Average: The Performance Appraisal Purpose effect, *Personnel Psychology,* Vol. 50(4), pp. 905-925.

[32] Kamphorst, J., and O. Swank (2015), The Role of Performance Appraisals in Motivating Employees, Tinbergen Institute Discussion Paper no. 12-034/1.

[33] Laffont, J. (1990), Analysis of Hidden Gaming in a Three-Level Hierarchy, *Journal of Law, Economics & Organization*, 6(2), pp. 301-324.

[34] Laffont, J., and J. Rochet (1997), Collusion in Organizations, *Scandinavian Journal of Economics*, 99(4), pp. 485-495.

[35] Larkin, I. (2014), The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales. *Journal of Labor Economics,* 32(2), pp. 199 - 227.

[36] Li, D. (2013), Expertise vs. Bias in Evaluation: Evidence from the NIH, *American Economic Journal: Applied,* forthcoming.

[37] Manthei, K., and D. Sliwka (2014), Multitasking and the Benefits of Objective Measurement - Evidence of a Field Experiment, mimeo, Univerisity of Cologne.

[38] Mookherjee, D., (2013), Incentives in Hierarchies, in R. Gibbons and J. Roberts (eds.), *The Handbook of Organizational Economics.* Princeton, NJ: Princeton University Press.

[39] Muller, D., and P. Weinschenk (2015), Rater Bias and Incentive Provision, *Journal of Economics and Management Strategy,* vol. 24(4), pp. 833-862.

[40] Oyer, P. (1998), Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality, *Quarterly Journal of Economics,* 113(1), pp. 149-185.

[41] Pearce, D.G., and E. Stacchetti (1998), The Interaction of Implicit and Explicit Contracts in Repeated Agency, *Games and Economic Behavior*, 23, pp. 75-96.

[42] Prasad, S., (2009), Task Assignment and Incentives: Generalists versus Specialists, *Rand Journal of Economics,* 40(2), pp. 380-403.

[43] Prendergast, C. (1999), The Provision of Incentives in Firms, *Journal of Economic Literature,* vol. 37(1), pp. 7-63.

[44] Prendergast, C. (2007), The Motivation and Bias of Bureaucrats, *American Economic Review,* 97(1), pp. 180-196.

[45] Prendergast, C., and R. Topel (1996), Favoritism in Organizations, *Journal of Political Economy,* 104(5), pp. 958-978.

[46] Schmidt, K.M., and M. Schnitzer (1995), The Interaction of Explicit and Implicit Contracts, *Economic Letters*, 48(2), pp. 193-199.

[47] Suvorov, A. and J. van de Ven (2009), Discretionary Rewards as a Feedback Mechanism, *Games and Economic Behavior*, 67, pp. 665-681.

[48] Thiele, V. (2013), Subjective Performance Evaluations, Collusion, and Organizational Design, *Journal of Law, Economics, & Organization*, 29(1), pp. 35-59.

[49] Tirole, J. (1986), Hierarchies and Bureaucracies: On the Role of Collusion in Organizations, *Journal of Law, Economics, and Organization*, 2, pp. 181-214.

[50] Vafaï, K. (2002), Preventing Abuse of Authority in Hierarchies, *International Journal of Industrial Organization*, 20, pp. 1143-1166.

[51] Vafaï, K. (2010), Opportunism in Organizations, *Journal of Law, Economics, & Organization*, 26(1), pp. 158-181.

[52] Woods, A. (2012), Subjective Adjustments to Objective Performance Measures: The Influence of Prior Performance, *Accounting, Organizations and Society,* 37, pp. 403-425.

[53] Zabojnik, J. (2014), Subjective Evaluations with Performance Feedback, *RAND Journal of Economics,* 45(2), pp. 341-369.

A Appendix

Proof of Lemma 1. Suppose the principal wants to induce efforts $e^* = \{e_1^*, e_2^*\}$, which would lead to measured performance $m = \underline{m}$. First, suppose the supervisor also wants to induce e^*. Suppose that agent exerts effort e^*. Provided that $w_A(\underline{m}, r_G) \geq w_A(m \neq \underline{m}, r)$, the agent will not manipulate m ex post. Using (2), this gives

$$U_A(e^*) = w_A(\underline{m}, r_G) - \frac{1}{2}(e_1^*)^2 - \frac{1}{2}(e_2^*)^2 \tag{A1}$$

Let $U_A(e^{-*}, m \neq \underline{m})$ represent the agent's expected utility after choosing $e \neq e^*$ such that $m \neq \underline{m}$. Now, the agent will manipulate m when given the option, implying that

$$U_A(e^{-*}, m \neq \underline{m}) = qw_A(m \neq \underline{m}, r_B) + p(1-q)w_A(\underline{m}, r_B) + (1-p)(1-q)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1^{-*}\right)^2 - \frac{1}{2}\left(e_2^{-*}\right)^2 \tag{A2}$$

Note that with probability $(1-p)q$ the supervisor sends report r_B even though she did not observe the agent's efforts, because she learned from $m \neq \underline{m}$ that the agent did not adhere to her demands. Lastly, for any $e \neq e^*$ such that $m = \underline{m}$, we have

$$U_A(e^{-*}, m = \underline{m}) = pw_A(\underline{m}, r_B) + (1-p)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1^{-*}\right)^2 - \frac{1}{2}\left(e_2^{-*}\right)^2$$

The agent exerts e^* when $U_A(e^*) \geq \max_e U_A(e^{-*})$. It follows that the agent's rents $\max_e U_A(e^{-*})$ are minimal when $w_A(m \neq \underline{m}, r_B) = w_A(\underline{m}, r_B) = 0$.

Next, consider the supervisor's incentive to demand $e^S \neq e^*$. First, consider any e^S yielding

$m = \underline{m}$. As measured performance cannot distinguish between e^S and e^*, the principal cannot affect the agent's choice through the wage scheme. Hence, if effort cost at e^* are at least as large as at e^S, the principal cannot prevent the supervisor from inducing e^S. Second, consider any e^S yielding $m = m^S \neq \underline{m}$. Now, the agent might optimally manipulate measured performance into either m^S or $\underline{m}$. Suppose first that the agent manipulates m into $\underline{m}$ when the opportunity arises. If the agent follows the supervisor's request, $e = e^S$, his expected utility equals

$$U_A(e^S) = qw_A(m \neq \underline{m}, r_G) + (1-q)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1^S\right)^2 - \frac{1}{2}\left(e_2^S\right)^2 \tag{A3}$$

If the agent exerts $e \neq e^S$ such that $m \neq m^S$ and $m \neq \underline{m}$, expected utility is

$$U_A(e \neq e^S) = qw_A(m \neq \underline{m}, r_B) + p(1-q)w_A(\underline{m}, r_B) + (1-p)(1-q)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1\right)^2 - \frac{1}{2}\left(e_2\right)^2 \tag{A4}$$

Next, consider $e \neq e^S$ such that $m = \underline{m} \neq m^S$. This yields

$$U_A(e \neq e^S) = pw_A(\underline{m}, r_B) + (1-p)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1\right)^2 - \frac{1}{2}\left(e_2\right)^2$$

Lastly, the agent can exert $e \neq e^S$ such that $m = m^S$, which yields

$$\begin{aligned} U_A(e \neq e^S) &= pqw_A(m \neq \underline{m}, r_B) + p(1-q)w_A(\underline{m}, r_B) + (1-p)qw_A(m \neq \underline{m}, r_G) + \\ &+ (1-p)(1-q)w_A(\underline{m}, r_G) - \frac{1}{2}\left(e_1\right)^2 - \frac{1}{2}\left(e_2\right)^2 \end{aligned}$$

The set of e^S the supervisor can demand increases in the difference between $U_A(e^S)$ and $\max_e U_A(e \neq e^S)$. Hence, the principal wants to minimize this difference. It is not possible to set $w_A(m \neq \underline{m}, r_B) > w_A(m \neq \underline{m}, r_G)$, as the cheap talk nature of the reports implies that the supervisor would increase the set e^S by switching the report labels. Hence, it is optimal for the principal to set $w_A(m \neq \underline{m}, r_G) = w_A(m \neq \underline{m}, r_B) = 0$. This also implies that manipulation of m into $m^S \neq \underline{m}$ is never beneficial (for neither the agent nor the supervisor).

Lastly, we show that $w_A(\underline{m}, r_B) > 0$ does not affect the supervisor's decision to induce e^S rather than e^*, implying that optimally $w_A(\underline{m}, r_B) = 0$ to minimize the agent's rents as derived above. We focus on the case where the agent's best alternative to following the supervisor's demand is an effort allocation such that $m \neq \underline{m}$. The case where the best alternative yields $m = \underline{m}$ is analog and therefore omitted. If the agent decides not to follow the supervisor's demands (either e^* or e^S), it follows from substituting for $w_A(m \neq \underline{m}, r_G) = w_A(m \neq \underline{m}, r_B) = 0$ into (A2) and (A4) that exerting $e_1 = e_2 = 0$ is optimal. Using (A1), it follows that the agent exerts $e = e^* = e^S$ when

$$(p(1-q)+q)\, w_A(\underline{m}, r_G) - \frac{1}{2}(e_1^*)^2 - \frac{1}{2}(e_2^*)^2 \geq p(1-q) w_A(\underline{m}, r_B) \tag{A5}$$

Both the principal and the supervisor are best off when these equations hold with equality. Suppose that e^* is such that $\frac{e_2^*}{e_1^*} = \frac{1-\kappa}{\kappa}$, for any $\kappa \in [0,1]$. By (A5), this implies that $e_1^* = \kappa\sqrt{2\,(p(1-q)+q)\, w_A(\underline{m}, r_G) - 2p(1-q) w_A(\underline{m}, r_B)}$ and $e_2^* = \frac{1-\kappa}{\kappa} e_1^*$. Using (4), this yields supervisor utility

$$U_S(e^*) = (\eta\kappa + (1-\eta)(1-\kappa))\sqrt{2\,(p(1-q)+q)\, w_A(\underline{m}, r_G) - 2p(1-q) w_A(\underline{m}, r_B)}$$

Using (A3) and (SR), the optimal $e^S \neq e^*$ for the supervisor is given by $e_1^S = \eta\sqrt{2p(1-q)\,(w_A(\underline{m}, r_G) - w_A(\underline{m}, r_B))}$ and $e_2^S = \frac{1-\eta}{\eta} e_1^S$. Using (4), this yields

$$U_S(e^S) = \left(\eta^2 + (1-\eta)^2\right)\sqrt{2p(1-q)\,(w_A(\underline{m}, r_G) - w_A(\underline{m}, r_B))}$$

The supervisor prefers to induce e^S rather than e^* if $U_S(e^*) < U_S(e^S)$. Now consider an increase in $w_A(\underline{m}, r_B)$. This gives $\frac{\partial U_S(e^*)}{\partial w_A(\underline{m}, r_B)} = -(\eta\kappa + (1-\eta)(1-\kappa))\frac{p(1-q)}{\sqrt{2(p(1-q)+q) w_A(\underline{m}, r_G) - 2p(1-q) w_A(\underline{m}, r_B)}}$ and $\frac{\partial U_S(e^S)}{\partial w_A(\underline{m}, r_B)} = -\left(\eta^2 + (1-\eta)^2\right)\frac{p(1-q)}{\sqrt{2p(1-q)(w_A(\underline{m}, r_G) - w_A(\underline{m}, r_B))}}$. It follows that if $U_S(e^*) < U_S(e^S)$, we also have $\frac{\partial U_S(e^*)}{\partial w_A(\underline{m}, r_B)} < \frac{\partial U_S(e^S)}{\partial w_A(\underline{m}, r_B)}$. When the supervisor prefers some feasible e^S over e^*, an increase in $w_A(\underline{m}, r_B)$ does not induce the supervisor to demand e^*. Hence, given that the agent's rents increases in $w_A(\underline{m}, r_B)$ as shown above, it is optimal for the principal to set $w_A(\underline{m}, r_B) = 0$. ■

Proof of Proposition 1. Condition (i) follows from Lemma 2. By Lemma 1, $b(m = \underline{m}) = 0$ and $b(m \neq \underline{m}) = -c$. Let $\varphi \leq \frac{1}{2}$. Given c, $\underline{m}$, $b(m = \underline{m}) = 0$, and $b(m \neq \underline{m}) = -c$, the supervisor maximizes utility (4) with respect to $\underline{e_1}$ and $\underline{e_2}$, subject to the agent's incentive compatibility constraint. If the supervisor requests effort levels that yield $m = \underline{m}$, this constraint is given by (21). If the first term between braces in (21) is larger than the second term, i.e. if $c \leq \frac{1}{2(1-p)q} \frac{1}{\varphi^2+(1-\varphi)^2} \underline{m}^2$, the supervisor optimally requests

$$\underline{e_1} = \frac{\varphi}{\varphi^2 + (1-\varphi)^2} \underline{m} + \frac{(1-\varphi)}{\varphi^2 + (1-\varphi)^2} \sqrt{2\left(\varphi^2 + (1-\varphi)^2\right)(p + (1-p)\,q)\,c - \underline{m}^2} \quad \text{(A6)}$$

$$\underline{e_2} = \frac{(1-\varphi)}{\varphi^2 + (1-\varphi)^2} \underline{m} - \frac{\varphi}{\varphi^2 + (1-\varphi)^2} \sqrt{2\left(\varphi^2 + (1-\varphi)^2\right)(p + (1-p)\,q)\,c - \underline{m}^2} \quad \text{(A7)}$$

Anticipating this, the principal maximizes (1) with respect to c and $\underline{m}$. The first-order conditions for c and $\underline{m}$ are, respectively, given by

$$\frac{(1-2\varphi)(p + (1-p)\,q)}{\sqrt{2\left(\varphi^2 + (1-\varphi)^2\right)(p + (1-p)\,q)\,c - \underline{m}^2}} - 1 = 0$$

$$\frac{1}{\varphi^2 + (1-\varphi)^2}\left(1 - \frac{(1-2\varphi)\,\underline{m}}{\sqrt{2\left(\varphi^2 + (1-\varphi)^2\right)(p + (1-p)\,q)\,c - \underline{m}^2}}\right) = 0$$

This gives $c = \underline{m} = (p + (1-p)\,q)$. Substituting for c and $\underline{m}$ in (A6) and (A7) yields $\underline{e_1} = \underline{e_2} = (p + (1-p)\,q)$.

If, instead, $c > \frac{1}{2(1-p)q} \frac{1}{\varphi^2+(1-\varphi)^2} \underline{m}^2$, the agent deviates by generating $m = \underline{m}$ along (PMR). Substituting for c and $\underline{m}$ yields condition $p \geq \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$. The second condition follows from the supervisor's incentive to deviate. When ignoring $\underline{m}$, the supervisor optimally requests

$$\underline{e_1} = \eta \sqrt{\frac{2p\,(1-q)\,c}{\eta^2 + (1-\eta)^2}} \quad \text{(A8)}$$

$$\underline{e_2} = (1-\eta) \sqrt{\frac{2p\,(1-q)\,c}{\eta^2 + (1-\eta)^2}} \quad \text{(A9)}$$

where we have used that the agent's incentive compatibility constraint is now given by $(1-q)\,c - \frac{1}{2}\underline{e_1}^2 - \frac{1}{2}\underline{e_2}^2 \geq (1-q)(1-p)\,c$. Comparing supervisor utility levels from adhering to and ignoring $\underline{m}$ implies that the supervisor adheres to target $\underline{m}$ when

$$p+(1-p)\,q \geq \sqrt{2\left(\eta^2+(1-\eta)^2\right)p\,(1-q)\,(p+(1-p)\,q)}$$

which can be rewritten to $p \leq \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$. ■

Proof of Proposition 2. By Lemma 2, it is not possible to induce $\frac{e_2}{e_1} > \frac{1-\varphi}{\varphi}$. By Lemma 1, $b(m=\underline{m})=0$ and $b(m \neq \underline{m}) = -c$. For given c and $\underline{m}$, the supervisor maximizes (4) with respect to $\underline{e_1}$ and $\underline{e_2}$, subject to (21). Assuming that $(1-p)qc \leq \frac{1}{2}\frac{1}{\varphi^2+(1-\varphi)^2}\underline{m}^2$, optimally requested efforts are (A6) and (A7). Substituting for $\underline{e_1}$ and $\underline{e_2}$ into (1) gives

$$U_P = \frac{1}{\varphi^2+(1-\varphi)^2}\underline{m} + \frac{(1-2\varphi)}{\varphi^2+(1-\varphi)^2}\sqrt{2\left(\varphi^2+(1-\varphi)^2\right)(p+(1-p)\,q)\,c-\underline{m}^2} - a - c - w_s \tag{A10}$$

By $\varphi > \frac{1}{2}$, (A10) is decreasing in c. This implies that (21) is binding. Using this condition to substitute for c in (A10) and maximizing with respect to $\underline{m}$ gives first-order condition

$$\frac{1}{\varphi^2+(1-\varphi)^2} - \frac{\underline{m}}{\left(\varphi^2+(1-\varphi)^2\right)(p+(1-p)\,q)} = 0$$

which yields $\underline{m} = p+(1-p)\,q$. Substituting for $\underline{m}$ in (21) gives $c = \frac{p+(1-p)q}{2\left(\varphi^2+(1-\varphi)^2\right)}$, and efforts follow from substituting for $\underline{m}$ and c in (A6) and (A7). Lastly, at this solution, we have $(1-p)qc < \frac{1}{2}\frac{1}{\varphi^2+(1-\varphi)^2}\underline{m}^2$ as assumed.

The supervisor prefers adhering to the principal's demand $m=\underline{m}$ rather than deviating to the best-feasible effort allocation along (SR) when

$$(\eta\varphi+(1-\eta)(1-\varphi))\sqrt{(p+(1-p)\,q)} \geq \sqrt{\left(\eta^2+(1-\eta)^2\right)\left(\varphi^2+(1-\varphi)^2\right)p\,(1-q)}$$

where the left-hand (right-hand) side follows from substituting for efforts (A6) and (A7) ((A8) and (A9)) into (4). Rewriting yields $p \leq \frac{q}{(1-q)}\frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}$. ■

Proof of Proposition 3. By Lemma 1, $b(m=\underline{m})=0$ and $b(m \neq \underline{m}) = -c$. If $p < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, under the contract derived in Proposition 1, the agent optimally deviates to the effort levels on the (PMR) that generate $m = \underline{m}$. Hence, the principal faces an additional constraint, given by equating the two terms between braces in (21), which yields $c = \frac{1}{2(1-p)q} \frac{1}{\varphi^2 + (1-\varphi)^2} \underline{m}^2$. The optimal contract follows from substituting for efforts (A6) and (A7) and for c into (1) and maximizing with respect to $\underline{m}$. The first-order condition for $\underline{m}$ is

$$1 + (1 - 2\varphi) \sqrt{\frac{p}{(1-p)q}} - \frac{1}{(1-p)q} \underline{m} = 0$$

This can be rewritten to $\underline{m} = (1-p)q + (1-2\varphi) \sqrt{p(1-p)q}$, yielding $c = \frac{\left(\sqrt{(1-p)q} + (1-2\varphi)\sqrt{p}\right)^2}{2\left(\varphi^2 + (1-\varphi)^2\right)}$. The effort levels follow from substituting for c and $\underline{m}$ into (A6) and (A7).

Lastly, we must show that the supervisor optimally adheres to $m = \underline{m}$. The best deviation is to request effort levels along (SR). Then, the agent's incentive constraint equals $(1-q)c - \frac{1}{2}\underline{e_1}^2 - \frac{1}{2}\underline{e_2}^2 \geq (1-q)(1-p)c$. Given c, the deviating supervisor optimally requests

$$\begin{aligned} \underline{e_1} &= \eta \sqrt{\frac{2(1-q)pc}{\eta^2 + (1-\eta)^2}} \\ \underline{e_2} &= (1-\eta) \sqrt{\frac{2(1-q)pc}{\eta^2 + (1-\eta)^2}} \end{aligned} \tag{A12}$$

Given $\varphi < \frac{1}{2}$, the incentive to deviate is strongest for a supervisor with $\eta = 1$. Given $\eta = 1$, supervisor utility (4) reduces to $U^S = e_1$. Substituting for c, $\underline{m}$, and $\eta = 1$ into (A6) and (A12) implies that the supervisor adheres to $m = \underline{m}$ when

$$\frac{1}{\varphi^2 + (1-\varphi)^2} \left(\varphi(1-p)q + \left(1 - 2\varphi^2\right) \sqrt{p(1-p)q} + (1-\varphi)(1-2\varphi)p\right) >$$
$$\left(\sqrt{(1-p)q} + (1-2\varphi)\sqrt{p}\right) \sqrt{\frac{(1-q)p}{\left(\varphi^2 + (1-\varphi)^2\right)}}$$

This expression increases in q. Rewriting condition $p < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, the lowest value of q considered in Proposition 3 is given by $q = \frac{p}{(1-p)(1-2\varphi)^2}$. Substituting for this level of q yields $\frac{2p\left(\varphi^2+(1-\varphi)^2\right)}{(1-2\varphi)^2} \left(\sqrt{\varphi^2 + (1-\varphi)^2} - \sqrt{\left((1-2\varphi)^2 - \frac{p}{1-p}\right)}\right) \geq 0$, which holds for any φ and p, as

$\varphi^2 + (1-\varphi)^2 \geq (1-2\varphi)^2$ given that $0 \leq \varphi \leq 1$. ■

Proof of Proposition 4. If $\varphi \leq \frac{1}{2}$ and $p > \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$, the optimal contract derived in Proposition 1 is not attainable. Similarly, if $\varphi > \frac{1}{2}$ and $p > \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$, the optimal contract derived in Proposition 2 is not feasible. In both cases, the supervisor has an incentive to deviate from inducing efforts that would satisfy $m = \underline{m}$ (i.e. (A6) and (A7)) to efforts along her most-preferred effort ratio (SR), as given by (A8) and (A9). Anticipating this, the principal must design a contract that meets the supervisor's incentive compatibility constraint:

$$\frac{\eta\varphi + (1-\eta)(1-\varphi)}{\varphi^2 + (1-\varphi)^2}\underline{m} + \frac{\eta - \varphi}{\varphi^2 + (1-\varphi)^2}\sqrt{2\left(\varphi^2 + (1-\varphi)^2\right)(p + (1-p)\,q)\,c - \underline{m}^2} \geq \sqrt{2\left(\eta^2 + (1-\eta)^2\right)(1-q)\,pc} \tag{A13}$$

where the left-hand side gives the supervisor's utility when meeting the principal's target and the right-hand side gives his utility when ignoring this target, both following from substituting the effort levels into supervisor utility (4).

By Lemma 1, $b(m = \underline{m}) = 0$ and $b(m \neq \underline{m}) = -c$. Substituting for efforts (A6) and (A7) into (1) and maximizing with respect to c and $\underline{m}$, taking into account limited liability constraints and the binding constraint (A13) gives lengthy first-order conditions, which, after straightforward (but tedious) rewriting, yield the expressions for c and $\underline{m}$ given in the proposition. Substituting for c and $\underline{m}$ into (A6) and (A7) yields the expressions for e_1 and e_2. ■

Proof of Proposition 5. Substituting for $a = w_s = 0$, and the optimal bonus c and efforts as given in Propositions 1-4 into (1) yields the following expressions for the principal's payoff U_P^j, where superscript $j \in [1, 4]$ indicates the corresponding proposition (with some abuse of notation):

$$\begin{aligned}
U_P^1 &= p + (1-p)\,q \\
U_P^2 &= \frac{p + (1-p)\,q}{2\left(\varphi^2 + (1-\varphi)^2\right)} \\
U_P^3 &= \frac{\left(\sqrt{(1-p)q} + (1-2\varphi)\sqrt{p}\right)^2}{2\left(\varphi^2 + (1-\varphi)^2\right)} \\
U_P^4 &= \frac{\left(\sqrt{p\,(1-q)} + |1-2\eta|\sqrt{q}\right)^2}{2\left(\eta^2 + (1-\eta)^2\right)}
\end{aligned}$$

Note that $U_P^j = c^j$ for all j. Hence, all comparative statics are identical too.

First, we compare the principal's payoffs at the exact parameter thresholds that determine which proposition is relevant. Substituting for $\varphi = \frac{1}{2}$ into U_P^2, we have $U_P^1 = U_P^2$. Similarly, when $p = \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, we have $U_P^1 = U_P^3$, and when $p = \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$ we have that $U_P^1 = U_P^4$. Lastly, when $p = \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$ and $\eta > \varphi > \frac{1}{2}$ we have that $U_P^2 = U_P^4$. Hence, marginal changes in parameter values do not lead to jumps in the principal's payoff, allowing us to focus on comparative statics within each proposition.

(i) The (weakly) positive effect of p follows directly for Propositions 1, 2, and 4. $\frac{\partial U_P^3}{\partial p} = -\frac{\sqrt{p}q-(1-2\varphi)\sqrt{(1-p)q}}{2\sqrt{p}\sqrt{(1-p)q}} > 0$, where the sign follows from the conditions $\varphi \leq \frac{1}{2}$ and $p < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$ under which Proposition 3 is relevant.

(ii) The (weakly) positive effect of q follows directly for Propositions 1, 2, and 3. $\frac{\partial U_P^4}{\partial q} = -\frac{\sqrt{q}p-(1-2\eta)\sqrt{p(1-q)}}{2\sqrt{q}\sqrt{p(1-q)}} > 0$, where the sign follows from the conditions $p > \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$ or $p > \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$ and $\eta > \varphi > \frac{1}{2}$ under which Proposition 4 is relevant.

(iii) $\frac{\partial U_P^1}{\partial \varphi} = \frac{\partial U_P^4}{\partial \varphi} = 0$. U_p^2 decreases in bias $|\varphi - \frac{1}{2}|$. $\frac{\partial U_P^3}{\partial \varphi} = -\frac{(1-2\varphi)(p-q+pq)+4\varphi(1-\varphi)\sqrt{p(1-p)q}}{\left(\varphi^2+(1-\varphi)^2\right)^2}$ and, hence, also decreases in bias $|\varphi - \frac{1}{2}|$ given the conditions $\varphi \leq \frac{1}{2}$ and $p < \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$ under which Proposition 3 is relevant.

(iv) $\frac{\partial U_P^1}{\partial \eta} = \frac{\partial U_P^2}{\partial \eta} = \frac{\partial U_P^3}{\partial \eta} = 0$. $\frac{\partial U_P^4}{\partial \eta} = -\frac{(1-2\eta)(q-p+pq)+4\eta(1-\eta)\sqrt{p(1-q)q}}{\left(\eta^2+(1-\eta)^2\right)^2}$ and, hence, decreases in supervisor bias $|\eta - \frac{1}{2}|$ given the conditions $p > \frac{1}{(1-2\eta)^2}\frac{q}{(1-q)}$ or $p > \frac{(\eta\varphi+(1-\eta)(1-\varphi))^2}{(\eta-\varphi)^2}\frac{q}{(1-q)}$ and $\eta > \varphi > \frac{1}{2}$ under which Proposition 4 is relevant. ■

Proof of Proposition 6. If $p_S < \frac{(\eta_S\varphi+(1-\eta_S)(1-\varphi))^2}{(\eta_S-\varphi)^2}\frac{q}{(1-q)}$ and $\varphi \leq \frac{1}{2}$ or $\varphi > \eta_S > \frac{1}{2}$, the principal's payoff with both supervisors is either U_P^1 or U_P^3, which are both increasing in p and independent of η. If $p_S > \frac{1}{(1-2\eta_S)^2}\frac{q}{(1-q)} > p_G$, the specialist yields payoff U_P^4, while the generalist yields U_P^1. By result (i) in Proposition 5, we have that $U_P^1 > U_P^4$ for any η. This proves the first part. When $\eta_S > \varphi > \frac{1}{2}$ and $p_S < \frac{(\eta_S\varphi+(1-\eta_S)(1-\varphi))^2}{(\eta_S-\varphi)^2}\frac{q}{(1-q)}$, employing the specialist yields U_P^2. Employing the generalist yields U_P^1 if $p_G \geq \frac{q(1-2\varphi)^2}{1+q(1-2\varphi)^2}$, otherwise it yields $U_P = \frac{\left(\sqrt{(1-p)q}+(2\varphi-1)\sqrt{p}\right)^2}{2\left(\varphi^2+(1-\varphi)^2\right)}$. Comparing these payoffs gives the two conditions in case (i), respectively. Lastly, when $p_G > \frac{1}{(1-2\eta_S)^2}\frac{q}{(1-q)}$, the specialist yields U_P^4, while the generalist yields U_P^1. Comparing these payoffs gives the condition in case (ii). ■