Performance measurement in multi-task agencies

Veikko Thiele

Sauder School of Business, University of British Columbia, Canada

Received 5 December 2006; accepted 25 June 2007

Abstract

This paper analyzes a multi-task agency relationship with a risk-neutral and financially constraint agent. The agent’s performance evaluation is incongruent, i.e. it does not reflect his contribution to firm value, and thus motivates an inefficient effort allocation across tasks. This paper investigates the improvement of the agent’s performance evaluation by contrasting two alternatives for the principal: (i) to invest in assets which can be utilized to generate additional measures about the agent’s performance; and (ii), to delegate this task to a supervisor. This paper demonstrates that delegation is superior whenever the costless available performance evaluation is sufficiently incongruent.

Keywords: Multi-task agencies; Performance measurement; Distortion; Congruity; Limited liability; Incentives

1. Introduction

Many employees are charged with performing multiple tasks which may contribute differently to firms’ objectives. Employees can therefore not only decide on their effort intensity, but also on how to allocate their efforts across all relevant tasks. Whenever their efforts are non-contractible, firms face a two-dimensional incentive problem: they need not only to induce a sufficient effort intensity, but also to motivate an efficient effort allocation across tasks. The latter objective could be achieved if firms are able to identify each employee’s contribution to firm value, and incorporate this information into incentive contracts. However, firm operations are often too complex to allow attribution of the contributions to each employee. Consequently, firms are compelled to employ other viable incentive mechanisms. One alternative is the application of objective performance measures in incentive contracts as advocated in the agency literature. However, when these performance measures are incongruent, i.e. they do not accurately reflect an employee’s individual contribution to firm value, their application in incentive contracts motivates employees to implement an inefficient effort allocation across relevant tasks (Feltham and Xie, 1994). For example, faculties at

Footnotes:

* Corresponding address: 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2. Tel.: +1 604 822 8491; fax: +1 604 822 8477.
E-mail address: veikko.thiele@sauder.ubc.ca.
1 For a review of agency literature see e.g.s. in Prendergast (1999), Lambert (2001), Gibbons (2005), and Christensen and Feltham (2005).
universities are primarily responsible for two tasks: teaching and research. Since teaching is harder to quantify than research output, promotion decisions are typically made on the basis of research accomplishments. This motivates particularly younger faculties to concentrate on research at the expense of teaching.\(^2\) Schools however, can modify their incentive schemes to improve the quality of teaching. Brickley and Zimmerman (2001) reported on the incentive scheme adopted by the William E. Simon Graduate School of Business Administration at the University of Rochester. While adjusting the performance evaluation and reward system during the early 90’s led to a significant improvement in teaching quality, research output declined steadily.

If firms do not have access to performance evaluations that are suitable to induce an efficient effort allocation, alternative mechanisms must be applied to mitigate potential effort distortion. Feltham and Xie (1994), Banker and Thevaranjan (2000), Datar et al. (2001), and Thiele (2006) demonstrated that the efficiency of the agent’s effort allocation can be enhanced by configuring multiple performance measures in an optimal fashion. Since they restricted their analysis to the aggregation of costlessly available performance measures, it is imperative to investigate investment decisions with the objective of improving the agent’s performance evaluation. This paper therefore elaborates on costly performance measurement in multi-task agency relationships in order to glean new insights into the improvement of an agent’s performance evaluation with the aim of mitigating effort distortion.

To exemplify the underlying idea of costly performance measurement, consider a worker who is employed for manufacturing goods. Suppose that these goods are priced according to their quality. Hence, it is intuitive to infer that the worker’s incentive contract should incorporate quality measures to ensure that the end product is of a desired quality. In contrast to the produced quantity however, verifying the achieved quality is costly. Inevitably, if the firm wants to maintain a certain quality standard, it must invest in a quality verification mechanism. For instance, the firm can invest in a machine that (e.g. randomly) verifies whether the quality of produced items is within a specified tolerance level, or alternatively, recruit a supervisor who is tasked with this verification process. Both performance measures – produced quantity and achievement of a desired quality level – can then be appropriately combined to improve the worker’s effort allocation. That is, the firm can commit to pay a bonus only if the produced quantity of items satisfying a specified quality exceeds a predetermined number. This incentive scheme guarantees that the worker is motivated to place more emphasis on the achievement of a certain quality standard besides producing a sufficient quantity.

Specifically, I analyze a multi-task agency relationship with risk-neutral parties, where the agent faces a liability limit. The agent’s performance evaluation is incongruent, i.e. it does not reflect his contribution to firm value. The application of this incongruent performance evaluation on an incentive contract would motivate the agent to implement an inefficient effort allocation across tasks. I investigate two alternatives for the principal to improve the agent’s performance evaluation and hence, to motivate a more efficient effort allocation: (i) to invest in assets which can be utilized to generate additional measures about the agent’s performance; and (ii), to delegate the information acquisition to a risk-neutral and financially constraint supervisor. As in the standard multi-task agency literature (see e.g. Feltham and Xie (1994), Banker and Thevaranjan (2000), Datar et al. (2001), and Baker (2002)), the relevant tasks are assumed to be equally costly for the agent to perform, and no synergies exist across these tasks. For this particular case, Schnedler (2006b) has shown that more congruent performance measures generate a higher benefit for the principal.

The consideration of equally costly tasks is not entirely unproblematic. As Schnedler (2006b) demonstrated, more congruent performance measures do not necessarily lead to a higher benefit for the principal whenever relevant tasks impose distinct costs on a risk-averse agent.\(^3\) The specific value of individual performance measures is therefore not exclusively dictated by their congruity and precision, but rather by their congruity relative to the respective task difficulties. Thiele (2006) also analyzed a multi-task agency relationship, where the relevant tasks impose different costs on a risk-averse agent. He elaborated on the optimal aggregation of costlessly available performance measures with the aim of mitigating agency costs. Besides the congruity and precision of available performance measures, the optimal aggregation is also found to reflect the agent’s respective costs for conducting the relevant tasks (Thiele, 2006). Assuming equally costly tasks therefore allows an in-depth examination of the efficient improvement of an agent’s performance evaluation, which is independent of agent-specific characteristics.

\(^2\) See Kerr (1975) for a discussion of this and for further examples. Additional illustrative examples have been summarized by Gibbons (1998), Prendergast (1999), and Baker (2000).

\(^3\) See Schnedler (2006a) for an insightful exposition of agency costs imposed by the application of incongruent performance measures in incentive contracts.
This paper demonstrates how costly performance measurement in multi-task agency relationships can be efficiently organized. In the event that the principal is forced to decide between delegation and an investment in suitable assets, the analysis in this paper brings to light two important implications. First, the principal generally prefers to delegate the performance measurement to the supervisor whenever the costless evaluation of the agent’s performance is sufficiently incongruent. In this case, delegation is more profitable even for a relatively small efficiency advantage of the supervisor for generating the additionally required performance measures. Second, an investment in appropriate assets is more likely for a relatively congruent evaluation of the agent’s performance. This is because delegation would impose high requirements on the supervisor’s relative measurement efficiency, which is less likely to be achieved by a potential supervisor. Finally, whenever the principal can employ a supervisor and invest in suitable assets contemporaneously, combining both considered measurement devices is optimal to alleviate the associated (convex increasing) costs. The supervisor’s relative efficiency for generating additional performance measures is shown to be crucial for the allocation of the information acquisition between the principal and the supervisor.

This paper proceeds as follows. I introduce the basic model in Section 2, and derive the agent’s optimal contract for a given performance evaluation in Section 3. Subsequently, I analyze in Section 4 the principal’s investment decision for generating additional measures about the agent’s performance. Section 5 then elaborates on the contractual arrangements when the principal employs a supervisor for generating the required performance measures. Both considered alternatives for improving the agent’s performance evaluation are compared in Sections 6 and 7. Section 8 summarizes the main results and concludes.

2. The model

Consider a single-period employment relationship between a principal and an agent. Both parties are risk-neutral and the agent faces a liability limit constraint, i.e. payments from the agent to the principal are not feasible. The agent is assigned to perform \( n > 2 \) tasks which cannot be split among different agents. Therefore, the agent needs to implement a vector of effort \( e = (e_1, \ldots, e_n)^T \), where \( e_t \in \mathbb{R}^+ \) is the agent’s non-verifiable effort allocated to task \( t \). The agent’s disutility of effort \( C(e) \) is quadratic and separable among the different activities: \( C(e) = e^T e / 2 \). By implementing effort \( e \), the agent can affect the firm value \( V \in \{0, 1\} \), where

\[
\text{Prob}(V = 1|e) = \min\{\mu^T e, 1\}, \tag{1}
\]

is the probability for realizing the high firm value conditional on \( e \). Vector \( \mu = (\mu_1, \ldots, \mu_n)^T, \mu \in \mathbb{R}^{n^+} \), reflects the sensitivity of the expected firm value in the agent’s effort. Accordingly, the agent can not only influence the expected firm value by his effort intensity, but also by his effort allocation across relevant tasks.

The realized firm value \( V \) is non-contractible, and therefore, cannot be used to provide the agent with incentives. The principal however, receives a binary and verifiable signal \( S \in \{0, 1\} \), where \( S = 1 \) denotes the favorable signal (Milgrom, 1981). The probability of realizing the favorable signal is conditional on the agent’s effort and takes the form

\[
\text{Prob}(S = 1|e) = \min\{\tilde{\omega}^T e, 1\}, \tag{2}
\]

where \( \tilde{\omega} = (\tilde{\omega}_1, \ldots, \tilde{\omega}_n)^T, \tilde{\omega} \in \mathbb{R}^{n^+} \), represents the sensitivity of the expected signal relative to the agent’s effort. The binary statistic \( S \) potentially represents several measures about the agent’s performance which are aggregated in the most efficient manner. Henceforth, I refer to \( S \) as the costless information system. To ensure interior solutions, I assume that \( \tilde{\omega} \) is characterized such that \( \text{Prob}(S = 1|e^*) = \tilde{\omega}^T e^* < 1 \) for the optimal effort vector \( e^* \).

The expected performance evaluation \( E[S|e] \) is deemed to be incongruent if it does not capture the contribution of the agent’s effort allocation to the expected firm value \( E[V|e] \). Formally, \( E[S|e] \) is incongruent if there exists no constant \( \lambda > 0 \) satisfying \( \mu = \lambda \tilde{\omega} \). For this case, Baker (2002) demonstrated that performance measure congruity can be characterized by the angle between the vector of the expected firm value sensitivities \( \mu \), and the vector of the expected performance measure sensitivities \( \tilde{\omega} \). To adopt Baker’s (2002) congruity measure for the subsequent analysis,

\[4\] All vectors are column vectors where ‘\( ^T \)’ denotes the transpose.

\[5\] One can show by using the subsequently derived solution that this requires \( \mu^T \tilde{\omega} / 2 < 1 \), which can be achieved by re-scaling \( \mu \) or \( \tilde{\omega} \).
let $\bar{\phi} \in [0, \pi/2]$ denote the angle between vector $\mu$ and vector $\bar{\omega}$.\(^6\) The relation between these two vectors is sufficient to characterize the inefficiency provoked by the application of an incongruent performance evaluation (Baker, 2002).\(^7\)

For subsequent analysis, keep in mind that $\bar{\phi}$ is negatively related to performance measure congruity, i.e. a smaller angle $\bar{\phi}$ characterizes a more congruent performance evaluation.

The principal can exploit the verifiable information system $\bar{S}$ in a bonus contract to provide the agent with appropriate incentives to implement effort. Due to the binary information structure, the incentive contract depicts a bonus payment similar in type to the one applied by Park (1995), Kim (1997), Pitchford (1998), Demougin and Fluet (2001). The contract therefore consists of a fixed transfer $\alpha^A$, and a bonus $\beta^A$ paid if the favorable signal $\bar{S} = 1$ is realized. Hence, the agent’s wage $w^A$ takes the form

$$w^A = \begin{cases} \alpha^A + \beta^A, & \text{if } \bar{S} = 1, \\ \alpha^A, & \text{if } \bar{S} = 0. \end{cases}$$

As a consequence of the agent’s liability limit, all transfers have to be non-negative for any realization of $\bar{S}$. If the agent accepts this bonus contract, it provides him with the expected utility $U^A(e) = E[w^A|e] - C(e)$. For parsimony, the agent’s reservation utility is normalized to zero.

3. Providing incentives

If the principal cannot directly contract over $e$, she faces an incentive problem. Therefore, she needs to find a bonus contract $(\alpha^{A*}, \beta^{A*})$ which aims at maximizing her expected profit $\Pi = E[V - w^A|e]$ while ensuring the agent’s participation. Formally, the optimal bonus contract solves

$$\max_{\alpha^A, \beta^A, e} \Pi = \mu^T e - \alpha^A - \beta^A \bar{\omega}^T e$$

s.t.

$$\alpha^A + \beta^A \bar{\omega}^T e - \frac{1}{2} e^T e \geq 0 \quad (5)$$

$$e = \arg \max_{\bar{e}} \alpha^A + \beta^A \bar{\omega}^T \bar{e} - \frac{1}{2} \bar{e}^T \bar{e} \quad (6)$$

$$\alpha^A \geq 0 \quad (7)$$

$$\alpha^A \geq 0 \quad (8)$$

Condition (5) is the agent’s participation constraint and ensures that it is in his interest to enter into this relationship. Further, (6) is the agent’s incentive condition, whereas (7) and (8) guarantee that the optimal bonus contract is compatible with his liability limit.

We can directly infer from the agent’s incentive constraint that he implements

$$e^* = \beta^A \bar{\omega}. \quad (9)$$

The effort vector $e^*$ consists of two components: the scalar $\beta^A$ and the vector $\bar{\omega}$. As stated in the extent multi-task agency literature, $\beta^A$ determines the overall effort intensity, whereas the relative effort allocation across tasks is exogenously characterized by $\bar{\omega}$. An incongruent performance evaluation therefore induces the agent to implement an inefficient effort allocation from the principal’s perspective. That is, the agent is motivated to place more emphasis on tasks with a higher contribution to his performance evaluation, which has potentially only a little effect on firm value, and vice versa.

---

\(^6\) Throughout this paper, angles are represented in radian measures.

\(^7\) The scaling of $E[\bar{S}|e]$, characterized by the length of vector $\bar{\omega}$, $\|\bar{\omega}\|$, does not affect the efficiency of the bonus contract since the agent is risk-neutral. With risk-averse agents however, the scaling is crucial since it affects the precision of performance measures and hence, the agent’s risk-premium. Refer to Gibbons (2005) or Thiele (2006) for further discussions.
Proposition 1. The optimal bonus contract is characterized by $\alpha^{\ast} = 0$ and
\[
\beta^{\ast}(\hat{\varphi}) = \frac{\|\mu\|}{2\|\omega\|} \cos \hat{\varphi}.
\]
For $\hat{\varphi} \in [0, \pi/2)$, the agent extracts a rent which is maximized for $\hat{\varphi} = 0$, and decreasing in $\hat{\varphi}$. The optimal bonus contract provides the principal with the expected profit
\[
\Pi^{\ast}(\hat{\varphi}) = \frac{1}{4} \|\mu\|^2 \cos^2 \hat{\varphi}.
\]

Proof. See Appendix.

The optimal bonus $\beta^{\ast}(\hat{\varphi})$ – and as a consequence the principal expected profit $\Pi^{\ast}(\hat{\varphi})$ – reflects the congruity of the costless information system $\hat{S}$ as characterized by $\hat{\varphi}$. Observe that $\beta^{\ast}(\hat{\varphi})$ and $\Pi^{\ast}(\hat{\varphi})$ are decreasing in $\hat{\varphi}$. The rationale behind this observation is that a less congruent performance evaluation (higher $\hat{\varphi}$) motivates the agent to implement a less efficient effort allocation. This entails two inefficiencies: (i) the expected firm value $E[V|e^{\ast}]$ is lower than it would have been under the implementation of less distorted effort; and (ii), the principal has to compensate the agent even for the inefficient effort allocation in order to ensure his participation. Accordingly, it is optimal from the principal’s perspective to respond by diminishing the bonus $\beta^{\ast}(\hat{\varphi})$, which in turn leads to a lower expected profit $\Pi^{\ast}(\hat{\varphi})$.

As a result of his liability limit, the agent extracts an economic rent. More interestingly however, is the observation that a more congruent evaluation of the agent’s performance leads to a higher rent extraction. This is because providing the agent with more congruent incentives (lower $\hat{\varphi}$) leads to the implementation of less distorted effort. Then, it is beneficial from the principal’s perspective to enhance the bonus in order to motivate a higher effort intensity. Despite causing the extraction of a higher rent, a more congruent performance evaluation yields a higher expected profit $\Pi^{\ast}(\hat{\varphi})$. This can be observed because the additionally generated surplus is shared between the principal and the agent.

4. Costly performance measurement

As observed in the preceding section, the contractual efficiency in multi-task agencies is exclusively determined by the congruity of an available information system whenever all involved parties are risk-neutral. Accordingly, the principal is better off if she has access to an information system that better reflects the agent’s contribution to firm value, and can therefore be utilized to induce a more efficient effort allocation.

Suppose that the principal invests in assets which are suitable to generate additional measures about the agent’s performance, henceforth referred to as central investment or central performance measurement. Let $m \in \mathbb{R}^+$ denote the principal’s measurement intensity, where a higher intensity generates more measures and thus, leads to a more congruent information system. Implementing $m$ imposes strictly convex increasing costs $C(m)$, with $C''(m) \geq 0$, $C(0) = C'(0) = 0$, and $\lim_{m \to \infty} C'(m) = \infty$.8

Assumption 1. The implementation of $m$ generates a new verifiable and binary statistic $S(m) \in \{0, 1\}$ characterized by
\begin{enumerate}
\item $\text{Prob}(S = 1|m, \hat{\varphi}, e) = \min \{\omega(\varphi(m, \hat{\varphi}))^T e, 1\}$.
\item $\varphi(m, \hat{\varphi}) < \hat{\varphi} \ \forall m > 0$.
\end{enumerate}

Implementing an arbitrary measurement intensity $m$ generates a new information system which is represented by the binary statistic $S(m) \in \{0, 1\}$. According to condition (i), the corresponding expected signal is further characterized by the new sensitivity vector $\omega(\varphi(m, \hat{\varphi}))$.9 Its relation to the sensitivity vector $\mu$ of the expected firm value is determined by the angle $\varphi(\cdot)$, which measures the congruity of $S(m)$. The angle $\varphi(m, \hat{\varphi})$ is a function

\footnote{A non-negative third derivative ensures that the second-order approach for delegating the information acquisition to a supervisor is sufficient for all values of the supervisor’s measurement intensity.}

\footnote{We can infer from the preceding analysis that the length $\|\omega(\cdot)\|$ is arbitrary and hence, does not affect the subsequent results.}
of the implemented measurement intensity \( m \), and the congruity of the costless available information system \( \bar{S} \) as measured by \( \bar{\phi} \). Finally, condition (ii) ensures that implementing a strictly positive intensity \( m \) leads eventually to a more congruent information system as characterized by a smaller angle \( \varphi(\cdot) \).\(^{10}\)

**Assumption 2.** The information system \( S(m) \) with \( \varphi(m, \bar{\phi}) \) as its congruity measure is characterized by

\[
\begin{align*}
\text{(i)} & \quad \varphi_m < 0 \text{ and } \varphi_{mm} > 0 \quad \forall m \geq 0, \\
\text{(ii)} & \quad \varphi_{\bar{\phi}} > 0 \text{ and } \varphi_{m\bar{\phi}} < 0 \quad \forall m \geq 0, \\
\text{(iii)} & \quad \varphi(0, \bar{\phi}) = \bar{\phi} \text{ and } \varphi(m, 0) = 0, \\
\text{(iv)} & \quad \bar{\phi} \in [0, \pi/4],
\end{align*}
\]

where \( \varphi_i \) denotes the first, and \( \varphi_{ii} \) the second derivative of \( \varphi(\cdot) \) with respect to \( i, i = m, \bar{\phi} \).

Condition (i) implies that a higher measurement intensity \( m \) enhances the congruity of the information system, whereas the marginal effect of reducing \( \varphi(\cdot) \) is decreasing in \( m \). Condition (ii) states that the less congruent the costless information system \( \bar{S} \) is, the less congruent is the new information system \( S(m) \) for a given measurement intensity \( m \). Moreover, the marginal effect of reducing \( \varphi(\cdot) \) by implementing an arbitrary intensity \( m \) is increasing in \( \bar{\phi} \). Condition (iii) emphasizes that without costly performance measurement \( (m = 0) \), only the costless information system \( \bar{S} \) is available. Additionally, a perfectly congruent information system \( (\bar{\phi} = 0) \) cannot be improved. To understand the last condition, observe from (11) that the principal’s expected benefit becomes

\[
\bar{V}(m) = \frac{1}{4} \| \mathbf{\mu} \|^2 \cos^2 \varphi(m, \bar{\phi}).
\]

It can be verified that \( \bar{V}(m) \) is strictly concave increasing in \( m \) for \( \varphi(\cdot) \in (0, \pi/4) \), whereas for \( \varphi(\cdot) \in [\pi/4, \pi/2) \), the shape of \( \bar{V}(m) \) depends on the particular characteristics of \( \varphi(\cdot) \). Consequently, condition (iv) guarantees that the first-order approach for identifying the optimal measurement intensity is sufficient.

As mentioned above, the principal commits to a particular measurement intensity \( m \) by investing \( C(m) \) prior to negotiating with the agent about his bonus contracts. As a consequence, the agent’s bonus contract is now characterized by the congruity measure \( \varphi(m, \bar{\phi}) \). Hence, we can directly turn to the principal’s problem for identifying the optimal measurement intensity \( m^* \):

\[
\max_m \Pi^C(m, \bar{\phi}) = \frac{1}{4} \| \mathbf{\mu} \|^2 \cos^2 \varphi(m, \bar{\phi}) - C(m).
\]

Since \( 2 \sin \varphi(\cdot) \cos \varphi(\cdot) = \sin(2\varphi(\cdot)) \), the optimal measurement intensity \( m^* \) implicitly solves

\[
\frac{1}{4} \| \mathbf{\mu} \|^2 \sin(2\varphi(m, \bar{\phi}))(\varphi_m) = C'(m).
\]

The principal chooses the measurement intensity \( m \) such that the marginal expected benefit equals marginal costs. Since the expected benefit is concave and investment costs are convex, the first-order approach is also sufficient. Observe that the optimal measurement intensity \( m^* \) depends implicitly on the congruity of the costless information system \( \bar{S} \) as measured by \( \bar{\phi} \). The effect of \( \bar{\phi} \) on \( m^* \) and the principal’s expected profit \( \Pi^C(\cdot) \) is clarified in the subsequent proposition.

**Proposition 2.** The optimal measurement intensity \( m^*(\bar{\phi}) \) is increasing in \( \bar{\phi} \). Overall, the principal’s expected profit \( \Pi^C(m^*(\bar{\phi}), \bar{\phi}) \) is decreasing in \( \bar{\phi} \).

**Proof.** See Appendix.

The less congruent the costless available information system \( \bar{S} \) (higher \( \bar{\phi} \)), the more distorted would be the agent’s induced effort allocation. Since this diminishes the principal’s expected profit, it is optimal to invest more in the improvement of the information system with the aim of mitigating effort distortion. However, improving the information system imposes convex increasing costs. Consequently, \( \Pi^C(\cdot) \) is decreasing in \( \bar{\phi} \).

---

\(^{10}\) Technically, new generated performance measures are not sufficient statistics of other available measures.
5. Delegating performance measurement

Instead of investing in the improvement of the agent’s performance evaluation, the principal can delegate the information acquisition to a supervisor. This could be the superior strategy whenever the supervisor is able to measure the agent’s performance more efficiently. However, employing a supervisor imposes a second moral hazard problem.

Like the agent, the supervisor is risk-neutral and faces a liability limit. For parsimony, his reservation utility is normalized to zero. The supervisor is charged with objectively measuring the agent’s performance with the intensity \( m^S \in \mathbb{R}_+ \) which satisfies Assumptions 1 and 2. To ease the comparability, the supervisor’s disutility of implementing \( m^S \) is \( C^S(m^S) = \eta C(m^S) \), where \( C(m^S) \) is identical to the principal’s investment costs for all \( m^S = m \). Potential differences in measurement efficiencies are characterized by \( \eta \in (0, 1] \). The ratio \( 1/\eta \) therefore measures the supervisor’s comparative advantage in generating the same additional performance measures relative to the principal.

The principal observes the realization of the binary statistic \( S(m^S) \in \{0, 1\} \), but cannot directly contract over \( m^S \). Suppose that the unfavorable signal \( S(m^S) = 0 \) is realized. This can occur because either the agent indeed failed to meet his performance objective, or the supervisor did not generate the required information. Consequently, the principal needs to provide the supervisor with incentives to motivate a desired measurement intensity \( m^S > 0 \). To do so, the principal can utilize the available information about the conducted measurement intensity \( m^S \), which are summarized by the verifiable and binary statistic \( M \in \{0, 1\} \).

**Assumption 3.** The probability for realizing the favorable signal about \( m^S \) is \( \Pr[M = 1|m^S] = \rho(m^S) \), which satisfies

1. \( \rho(0) = 0 \),
2. \( \rho(m^S) \) is twice-continuously differentiable with \( \rho'(m^S) > 0 \) and \( \rho''(m^S) \leq 0 \forall m^S \geq 0 \),
3. \( \lim_{m^S \to \infty} \rho(m^S) = 1 \) and \( \lim_{m^S \to 0} \rho'(m^S) = \infty \).

Condition (i) emphasizes that without improving the information system \( (m^S = 0) \), the binary statistic \( M \) can never be favorable. The second and third conditions are standard and guarantee an interior and unique solution for the supervisor’s optimal measurement intensity.

To induce a desired measurement intensity, the principal provides the supervisor with a bonus contract \( w^S \) based upon the verifiable statistic \( M \). In particular,

\[
w^S = \begin{cases} \alpha^S + \beta^S, & \text{if } M = 1, \\ \alpha^S, & \text{if } M = 0, \end{cases}
\]

(15)

where \( \alpha^S \) denotes the supervisor’s fixed payment, and \( \beta^S \) his bonus paid if \( M = 1 \). These contract elements are required to be non-negative as a result of the supervisor’s liability limit. Apparently, the supervisor can increase the probability \( \rho(m^S) \) to realize the favorable signal \( M = 1 \), and consequently, the likelihood to obtain the contracted bonus \( \beta^S \), by enhancing his measurement intensity \( m^S \). His bonus contract therefore provides the supervisor with the expected utility \( U^S(m^S) = E[w^S|m^S] - \eta C(m^S) \).

Before deriving the optimal bonus contract for the supervisor, it is useful to clarify the timing of this problem. First, at \( t = 0 \), the principal offers the supervisor a bonus contract \( w^S \) based upon the verifiable signal \( M \). At \( t = 1 \), the principal provides the agent with a bonus contract \( w^A \) based upon the anticipated performance evaluation \( S(m^S) \). At \( t = 2 \), the agent implements effort \( e^*(m^S) \), and at \( t = 3 \), the supervisor generates the additional performance measure(s) by exerting \( m^S \). Afterwards, at \( t = 4 \), the agent’s and the supervisor’s performance evaluations \( S(m^S) \) and \( M \) are realized and become public knowledge. Finally, at \( t = 5 \), all contracted payments take place.

The supervisor’s and the agent’s optimal contracts need to be derived by backward induction. At \( t = 1 \), the principal determines the agent’s bonus contract \( (\alpha^{A*}, \beta^{A*}) \). Note that the evaluation of the agent’s performance \( S(m^S) \) is now a function of the supervisor’s measurement intensity \( m^S \). Hence, Proposition 1 implies that \( \alpha^{A*} = 0 \) and

\[
\beta^{A*}(m^S, \varphi) = \frac{||\mu||}{2||\omega||} \cos \varphi(m^S, \varphi).
\]

(16)

---

11 I discuss in Section 8 how the subsequent results change when the supervisor is not financially constraint.

12 See Demougin and Fluet (2001) for a formal derivation that verifiable information about one-dimensional effort can be summarized by a binary statistic if parties are risk-neutral.
At $t = 0$, the principal establishes the supervisor’s contract $(\alpha^S, \beta^S)$ with the aim of maximizing her expected profit $\Pi^D = \mathbb{E}[V - w^S - w^S]\left[\alpha^S(m^S), m^S\right]$ while guaranteeing the supervisor’s participation. Given the agent’s optimal contract $(\alpha^A, \beta^A(m^S, \tilde{\varphi}))$, the supervisor’s optimal bonus contract solves

$$\max_{\alpha^S, \beta^S, m^S} \Pi^D = \frac{1}{4}\|\mu\|^2 \cos^2 \varphi(m^S, \tilde{\varphi}) - \alpha^S - \rho(m^S)\beta^S$$

subject to

$$\alpha^S + \rho(m^S)\beta^S - \eta C(m^S) \geq 0 \quad (18)$$

$$m^S = \arg \max_{\tilde{m}^S} \alpha^S + \rho(\tilde{m}^S)\beta^S - \eta C(\tilde{m}^S) \quad (19)$$

$$\alpha^S + \beta^S \geq 0 \quad (20)$$

$$\alpha^S \geq 0. \quad (21)$$

Condition (18) is the supervisor’s participation constraint and guarantees that it is in his interest to enter into this relationship. Additionally, (19) is the supervisor’s incentive constraint. Finally, conditions (20) and (21) ensure that the supervisor’s contract is compatible with his liability limit.

Tirole (1986), among others, cautioned against the employment of several agents (including supervisors) as it may lead to side-contracting – what is commonly referred to as collusion. Since collusion is deemed illegal and thus not court-enforceable, its accomplishment requires that none of the involved parties has an incentive to deviate ex post from the stipulated behavior. In this single-period framework, collusion between the agent and supervisor cannot be a dominant strategy due to one crucial reason: The supervisor generates the required performance measures before all payments take place. As a result of his liability limit, the agent can only pay the supervisor a bribe after he has implemented $m^S$. However, because side-contracting is not court-enforceable, the supervisor anticipates that the agent has ex post no incentives to honor collusive agreements, and thus implements $m^S$. The optimal bonus contract for the supervisor as presented in Proposition 3 is therefore collusion-proof.

**Proposition 3.** The supervisor’s optimal contract is characterized by the fixed transfer $\alpha^{S*} = 0$ and the expected bonus

$$B(m^S*, \eta) \equiv \beta^{S*}(m^S*, \eta)\rho(m^S*) = \frac{\eta \rho'(m^S*)\rho(m^S*)}{\rho'(m^S*)}. \quad (22)$$

The supervisor’s optimal measurement intensity $m^{S*}$ solves

$$\frac{1}{4}\|\mu\|^2 \sin(2\varphi(m^S, \tilde{\varphi}))(\tilde{\varphi} - \tilde{\varphi}) = \frac{\partial B(m^S, \eta)}{\partial m^S}. \quad (23)$$

Moreover, the supervisor extracts a rent for $m^{S*} > 0$, which is increasing in $m^{S*}$. The optimal bonus contracts for the agent and the supervisor provide the principal with the expected profit

$$\Pi^D(m^S*, \tilde{\varphi}, \eta) = \frac{1}{4}\|\mu\|^2 \cos^2 \varphi(m^S*, \tilde{\varphi}) - B(m^S*, \eta). \quad (24)$$

**Proof.** See Appendix.

Consider first the supervisor’s expected bonus $B(m^S*, \eta)$ which comprises the optimal alignment to induce $m^S*$. It is characterized by the likelihood ratio $\rho'(m^S*)/\rho(m^S*)$ which – according to Holmström (1979) – measures the precision of the evaluation of the supervisor’s performance. The expected bonus $B(m^S*, \eta)$ further consists of the supervisor’s relative measurement efficiency as parameterized by $\eta$. The less efficient the supervisor’s information acquisition (higher $\eta$), the higher must be the expected bonus to motivate an arbitrary measurement intensity $m^S$. Finally, the supervisor extracts a rent as a result of his liability limit.

---

13 See Tirole (1992) for a thorough discussion. However, some authors assume that side-contracting is per se enforceable, see e.g. Kofman and Lawarree (1993), Villadsen (1995), and Faure-Grimaud et al. (2001).
Next, consider the optimality condition for $m^{S*}$ as emphasized in Proposition 3. The optimal measurement intensity $m^{S*}$ implicitly depends on two parameters: (i) the congruity measure $\bar{\phi}$ of the costless information system $\bar{S}$; and (ii), the supervisor’s relative measurement efficiency $\eta$. Since both parameters determine $m^{S*}$, they eventually affect the congruity of the agent’s performance evaluation, and consequently, the efficiency of his effort allocation across tasks.

**Proposition 4.** The optimal measurement intensity $m^{S*}(\bar{\phi}, \eta)$ is increasing in $\bar{\phi}$, and decreasing in $\eta$. Overall, the principal’s expected profit $\Pi^D(m^{S*}(\bar{\phi}, \eta), \bar{\phi}, \eta)$ is decreasing in $\bar{\phi}$ and in $\eta$.

**Proof.** See Appendix.

The first part of this result follows because a less congruent information system $\bar{S}$ would induce the agent to implement a more distorted effort allocation. To restrict this inefficiency, it is optimal to provide the supervisor with more powerful incentives aimed at motivating the implementation of a higher measurement intensity $m^{S*}$. However, improving the available information system imposes convex increasing costs $B(\cdot)$. As a consequence, $\Pi^D(\cdot)$ is decreasing in $\bar{\phi}$. Moreover, it is more costly for the principal to induce an arbitrary intensity $m^{S}$ when the supervisor provides a lower comparative advantage in generating the desired measures about the agent’s performance (higher $\eta$). Thus, a higher $\eta$ leads to the provision of less powerful incentives with the aim of inducing a lower measurement intensity $m^{S}$.

### 6. When is delegation profitable?

According to previous observations, one crucial factor for the principal’s delegation decision is the supervisor’s relative measurement efficiency as parameterized by $\eta$. If $\eta$ is sufficiently low, one can expect that the principal favors delegation over a centralized investment, and vice versa. However, it is not obvious how the congruity of the costless information system measured by $\bar{\phi}$ affects the principal’s delegation decision. To unravel the fundamental forces, I focus in this section on the principal’s delegation decision in case she is forced to decide between a central investment and delegation. Then, in Section 7, I elaborate on the optimal combination of supervision and a central investment.

**Proposition 5.** There exists a cut off cost parameter $\eta^*(\bar{\phi}) \in (0, 1)$ such that the principal utilizes

(i) a delegated performance measurement, if $\eta \leq \eta^*(\bar{\phi})$.

(ii) a centralized performance measurement, if $\eta > \eta^*(\bar{\phi})$.

The cut off $\eta^*(\bar{\phi})$ is strictly increasing in $\bar{\phi}$ if

$$
\frac{\rho(m^*(\bar{\phi}))}{\rho'(m^*(\bar{\phi}))} \left[ \frac{C''(m^*(\bar{\phi}))}{C'(m^*(\bar{\phi}))} \cdot \frac{\rho''(m^*(\bar{\phi}))}{\rho'(m^*(\bar{\phi}))} \right] < \frac{1 - \eta}{\eta},
$$

and decreasing, otherwise.

**Proof.** See Appendix.

The results emphasized in Proposition 5 are illustrated in Fig. 1, where the supervisor’s relative measurement efficiency $\eta$ is depicted on the horizontal axis, and the principal’s expected profits on the vertical axis. Since $\Pi^C(\cdot)$ is independent of $\eta$, its curve is parallel to the horizontal axis, whereas $\Pi^D(\cdot)$ is decreasing in $\eta$, see Proposition 4. The intersection of both curves therefore characterizes the cut off $\eta^*(\bar{\phi})$. If $\eta \leq \eta^*(\bar{\phi})$, the principal is apparently better off by delegating the information acquisition to the supervisor. In this case, the supervisor’s relative measurement efficiency – characterized by $\eta$ – suffices to compensate the additional agency costs associated with delegation. In contrast, if $\eta > \eta^*(\bar{\phi})$, centrally investing in the information acquisition is more cost efficient because the supervisor’s rent would outweigh his relative measurement efficiency.

Next, suppose that the costless information system $\bar{S}$ becomes more incongruent, i.e. $\bar{\phi}$ increases to $\bar{\phi}'$. As a result, it is optimal to conduct a higher measurement intensity under a centralized and delegated performance measurement, see Propositions 2 and 4. However, enhancing the respective measurement intensities imposes convex increasing costs such that $\Pi^C(\cdot)$ and $\Pi^D(\cdot)$ decrease. As can be deduced from Fig. 1, the decline of $\Pi^D(\cdot)$ in $\bar{\phi}$ particularly depends

---

14 See proof of Proposition 3 in the Appendix for a formal verification that $B(\cdot)$ is convex.
on the value of \( \eta \). Suppose for a moment that (25) is satisfied, i.e. the measurement \( M \) of the supervisor’s performance is sufficiently precise. If \( \eta \leq \eta^*(\tilde{\varphi}) \), the marginal costs for enhancing the respective measurement intensities are lower under a delegated than a centralized performance measurement. Consequently, \( \Pi^D(\cdot) \) is less decreasing in \( \bar{\varphi} \) than \( \Pi^C(\cdot) \) such that \( \eta^*(\tilde{\varphi}) \) increases to \( \eta^{**}(\tilde{\varphi}) \), see Fig. 1. Generally speaking, a less congruent information system \( \bar{S} \) implies that supervision can be the dominant strategy even for a lower relative measurement efficiency provided by the supervisor (higher \( \eta \)). However, if the evaluation of the supervisor’s performance \( M \) is sufficiently imprecise such that (25) is violated, the reverse can be observed such that \( \eta^*(\tilde{\varphi}) \) decreases to \( \eta^{***}(\tilde{\varphi}) \).

To illustrate the relevance of the preceding observations to the design of organizations, consider for instance a sales department and a human resources department. Suppose that the following individual performance measures are costlessly available: (i) the achieved sales in the sales department; and (ii), the performed office-hours in the human resources department. The achieved sales can thus be inferred to be a more congruent performance measure than the performed office-hours in the human resources department. For subsequent inferences, assume that the evaluation of a potential supervisor’s performance is sufficiently precise in both departments.\(^{15}\) According to previous results, a supervisor’s minimum required relative measurement efficiency is higher for the sales, than for the human resources department. This eventually leads to the conclusion that supervision is more likely to be a firm’s superior strategy for the human resources than for the sales department.

7. The optimal performance measurement

So far, I have considered the principal’s choice of either centrally investing in the information acquisition, or delegating these tasks to a supervisor. However, since the associated measurement costs are convex, it is optimal from the principal’s perspective to utilize both mechanisms simultaneously.\(^{16}\) In this case, the principal chooses her and the supervisor’s measurement intensities \( m \) and \( m^S \) to maximize the difference between expected firm value and aggregated costs. The principal’s maximization problem can be achieved by combining (13) and (22):

\[
\max_{m, m^S} \Pi = \frac{1}{4} \| \mu \|^2 \cos^2 \varphi (m + m^S, \bar{\varphi}) - C(m) - B(m^S, \eta).
\]

The optimal measurement intensities \( m^* \) and \( m^{S*} \) are characterized by the following first-order conditions:

\[
\frac{1}{4} \| \mu \|^2 \sin(2\varphi (m + m^S, \bar{\varphi})) \left[ -\frac{\partial \varphi}{\partial m} \right] = C'(m),
\]

\[
\frac{1}{4} \| \mu \|^2 \sin(2\varphi (m + m^S, \bar{\varphi})) \left[ -\frac{\partial \varphi}{\partial m^S} \right] = \frac{\partial B(m^S, \eta)}{\partial m^S}.
\]

\(^{15}\) Formally, (25) is satisfied.

\(^{16}\) I would like to thank an anonymous referee for pointing this out.
Clearly from above, the principal allocates the measurement intensities such that the marginal expected firm value equals the respective marginal costs. The subsequent proposition clarifies how the relation between the principal’s and the supervisor’s measurement intensities is affected by the congruity measure $\bar{\phi}$.

**Proposition 6.** There exists a cut off congruity measure $\bar{\phi}^*(\eta)$ such that the difference between the principal’s and the supervisor’s measurement intensities, $\Delta m(\bar{\phi}, \eta) = m^*(\bar{\phi}, \eta) - m^S(\bar{\phi}, \eta)$, is

(i) decreasing in $\bar{\phi}$, if $\bar{\phi} \leq \bar{\phi}^*(\eta)$,

(ii) increasing in $\bar{\phi}$, if $\bar{\phi} > \bar{\phi}^*(\eta)$.

The cut off $\bar{\phi}^*(\eta)$ is decreasing in $\eta$.

**Proof.** See Appendix.

Since measurement costs are convex, the information acquisition aimed at improving the agent’s performance evaluation is allocated between the principal and the supervisor. Their respective measurement intensities however, respond differently to an increase of the congruity measure $\bar{\phi}$. As long as $\bar{S}$ is still sufficiently congruent ($\bar{\phi} \leq \bar{\phi}^*(\eta)$), the principal enhances the supervisor’s measurement intensity $m^S(\cdot)$ relatively more than her own intensity $m^*(\cdot)$. In this case, increasing the supervisor’s measurement intensity by one unit imposes lower marginal costs as a result of his relative cost advantage as characterized by $\eta$. In contrast, the supervisor’s rent – which is increasing in $m^S$ – leads to higher marginal costs for $\bar{\phi} > \bar{\phi}^*(\eta)$. Thus, it is optimal from the principal’s perspective to enhance her own measurement intensity relatively more than that of the supervisor. Finally, a lower relative cost advantage for the supervisor (higher $\eta$) imposes higher costs for inducing an arbitrary intensity $m^S$. This in turn is reflected by a lower cut off $\bar{\phi}^*(\eta)$.

8. Conclusion

For economic relationships that are subject to moral hazard, objective performance measurement is frequently utilized to provide agents with incentives. However, applying incongruent performance measures in incentive contracts motivates agents to implement inefficient effort allocations across relevant tasks. Hence, it is beneficial from firms’ perspective to utilize mechanisms which can mitigate this inefficiency. This paper investigates and compares two alternatives for alleviating effort distortion: (i) to centrally invest in the generation of additional measures about the agent’s performance; and (ii), to delegate this task to a supervisor.

The analysis in this paper reveals how costly performance measurement in multi-task agencies can be efficiently organized. Suppose that the principal is compelled to decide between a central investment and delegation, and the evaluation of the supervisor’s performance is sufficiently precise. For this case, the preceding analysis yields two important implications. First, for a sufficiently incongruent information system, the principal generally favors delegation. The rationale behind this observation is that delegation imposes low requirements on the supervisor’s relative measurement efficiency. This in turn is more likely to be satisfied by a potential supervisor. Second, the more congruent the costless information system, the more likely is a centralized investment in the information acquisition. This is because delegation would impose high requirements on the supervisor’s relative measurement efficiency, which is less likely to be achieved by a potential supervisor. Nevertheless, whenever the principal can utilize a central investment as well as supervision to improve the agent’s performance evaluation, charging a supervisor with this task is always beneficial regardless of his relative measurement efficiency. His relative efficiency however, determines the allocation of the information acquisition activity between the principal and the supervisor.

In the preceding analysis, the supervisor was assumed to be financially constraint. To briefly discuss the effect of this assumption on the results, suppose that the supervisor has unlimited access to financial resources. In this case, the principal can avoid the supervisor’s rent extraction by making him the residual claimant, and charging him a fee equivalent to his prospective profit. This also empowers the supervisor to directly contract with the agent. ‘Selling the firm’ to the supervisor is therefore the principal’s dominant strategy as long as the supervisor exhibits a relative cost advantage in generating additional performance measures ($\eta < 1$). This in turn ensures the implementation of a higher measurement intensity, and leads thus to a more congruent evaluation of the agent’s performance.
Appendix

Proof of Proposition 1. Note first that \(e_i > 0, i \in \{1, \ldots, n\}\), requires that \(\beta^A > 0\). Thus, (7) is satisfied as long as (8) holds, and can therefore be omitted. Let \(\lambda\) and \(\xi\) denote Lagrange multipliers. Since \(e^* = \beta^A \bar{\omega}\), the Lagrangian becomes

\[
\mathcal{L}(\alpha^A, \beta^A) = \mu^T \bar{\omega} \beta^A - \alpha^A - (\beta^A)^2 \bar{\omega}^T \bar{\omega} + \lambda \left[ \alpha^A + \frac{1}{2} (\beta^A)^2 \bar{\omega}^T \bar{\omega} \right] + \xi \alpha^A.
\]

(29)

The corresponding first-order conditions are

\[
-1 + \lambda + \xi = 0, \quad \mu^T \bar{\omega} + \beta^A \bar{\omega}^T \bar{\omega}(\lambda - 2) = 0.
\]

(30) (31)

Due to complementary slackness, we need to consider two cases: (i) \(\lambda > 0\); and (ii), \(\lambda = 0\). First, suppose that \(\lambda > 0\). In this case, \(\alpha^A + (\beta^A)^2 \bar{\omega}^T \bar{\omega}/2 = 0\). Since the agent’s liability limit requires that \(\alpha^A \geq 0\), \(\lambda > 0\) would imply that \(\alpha^{A_t} = \beta^{A_t} = 0\) and consequently, \(e^* = (0, \ldots, 0)^T\). Therefore, \(\lambda > 0\) cannot be a solution. Thus, \(\lambda = 0\), which contemporarily implies that \(\xi = 1\), see (30). Hence, \(\alpha^{A_t} = 0\) due to complementary slackness. By solving (31) for \(\beta^A\) with \(\lambda = 0\), and applying the relations \(\sum_i \mu_i^2 = ||\mu||\) and \(\sum_i \mu_i \bar{\omega}_i = ||\mu|| ||\bar{\omega}|| \cos \bar{\psi}\), we obtain the optimal bonus

\[
\beta^{A_t}(\bar{\psi}) = \frac{||\mu||}{2 ||\bar{\omega}||} \cos \bar{\psi}.
\]

(32)

Furthermore, the optimal bonus contract \((\alpha^{A_t}, \beta^{A_t}(\bar{\psi}))\) provides the agent with the expected utility \(U^A(\bar{\psi}) = ||\mu||^2 \cos^2 \bar{\psi}/8\). Since \(2 \sin \bar{\phi} \cos \bar{\psi} = \sin(2\bar{\phi})\),

\[
\frac{\partial U^A(\bar{\psi})}{\partial \bar{\psi}} = -\frac{1}{4} ||\mu||^2 \sin (2\bar{\phi}),
\]

(33)

which is strictly negative for all \(\bar{\psi} \in (0, \pi/2)\). Due to the definition of the cosine, \(U^A(\bar{\psi})\) is maximized for \(\bar{\psi} = 0\), and zero for \(\bar{\psi} = \pi/2\). Finally, the principal’s expected profit can be obtained by substituting \(\alpha^{A_t}, \beta^{A_t}, \bar{\psi}\) in the principal’s objective function. \(\square\)

Proof of Proposition 2. Implicitly differentiating \(m^*\) with respect to \(\bar{\psi}\) yields

\[
\frac{dm^*}{d\bar{\psi}} = -\frac{1}{4} ||\mu||^2 \left[ 2 \cos (2\bar{\phi})(-\bar{\psi}_m)\bar{\psi}_{\bar{\phi}} + \sin (2\bar{\phi})(-\bar{\psi}_m) \right]
\]

\[
= \frac{\beta}{\alpha} \left[ \frac{1}{4} ||\mu||^2 \sin (2\bar{\phi} (m, \bar{\psi})) (-\bar{\psi}_m) - C'(m) \right].
\]

(34)

The denominator is strictly negative due to the second-order condition, whereas the numerator is strictly positive for all \(\bar{\phi} \in (0, \pi/4)\). Consequently, \(dm^*/d\bar{\psi} > 0\). Moreover, applying the Envelope Theorem yields

\[
\frac{d\Pi^C(m^*(\bar{\psi}), \bar{\psi})}{d\bar{\psi}} = -\frac{1}{4} ||\mu||^2 \sin (2\bar{\phi}(\cdot)) \bar{\psi}_{\bar{\phi}}.
\]

(35)

Hence, \(d\Pi^C(\cdot)/d\bar{\psi} < 0\) for all \(\bar{\phi} \in (0, \pi/4)\). \(\square\)

Proof of Proposition 3. First observe that (19) is equivalent to \(\beta^S(m^S, \eta) = \eta C'(m^S)/\rho'(m^S)\). Hence, the expected bonus \(B(m^S, \eta) \equiv \beta^S(m^S, \eta) \rho(m^S)\) becomes

\[
B(m^S, \eta) = \frac{\eta C'(m^S) \rho(m^S)}{\rho'(m^S)}.
\]

(36)

To induce \(m^S > 0\), \(\beta^S(\cdot)\) needs to be strictly positive. Consequently, (20) is satisfied as long as (21) holds, and can therefore be omitted. Let \(\lambda\) and \(\xi\) denote Lagrange multipliers. Thus, the Lagrangian becomes

\[
\mathcal{L}(\alpha^S, m^S) = \frac{1}{4} ||\mu||^2 \cos^2 \varphi(m^S, \bar{\psi}) - \alpha^S - B(m^S, \eta) + \lambda [\alpha^S + B(m^S, \eta) - \eta C(m^S)] + \xi \alpha^S.
\]

(37)
The first-order conditions with respect to $\alpha^S$ and $m^S$ are

$$-1 + \lambda + \xi = 0,$$

$$\frac{1}{4} ||\mu||^2 \sin(2\varphi(m^S, \tilde{\varphi}))(-\varphi_{m^S}) - \frac{\partial B(\cdot)}{\partial m^S} + \lambda \left[ \frac{\partial B(\cdot)}{\partial m^S} - \eta C'(m^S) \right] = 0. \tag{38}$$

Suppose for a moment that $\lambda > 0$ such that $\alpha^S + B(\cdot) - \eta C(m^S) = 0$ due to complementary slackness. Since the supervisor’s liability limit requires that $\alpha^S \geq 0$, $\lambda > 0$ would imply that $B(\cdot) \leq \eta C(m^S)$. In this case, the supervisor would choose $m^S = 0$. Hence, $\lambda > 0$ cannot be a solution. As a result, $\lambda = 0$, which further implies that $\xi = 1$, see (38). Thus, $\alpha^{S*} = 0$ due to complementary slackness. Since $\lambda = 0$, (39) implies that $m^{S*}$ solves

$$\frac{1}{4} ||\mu||^2 \sin(2\varphi(m^S, \tilde{\varphi}))(-\varphi_{m^S}) = \frac{\partial B(m^S, \eta)}{\partial m^S}. \tag{40}$$

Next, it is necessary to verify that the first-order approach is also sufficient. Recall that Assumption 2 ensures that the principal’s benefit $\hat{V}(m^S)$ is concave in $m^S$. Thus, it is sufficient to show that $B(\cdot)$ is convex increasing in $m^S$. The first derivative of $B(\cdot)$ with respect to $m^S$ is

$$\frac{\partial B(\cdot)}{\partial m^S} = \eta C'(\cdot) + \frac{\rho(m^S)\eta C''(m^S)}{\rho'(m^S)} + \frac{\rho(m^S)\eta C'(m^S)(-\rho''(m^S))}{\left[\rho'(m^S)\right]^2}. \tag{41}$$

which is strictly positive for all $m^S > 0$. Moreover, the second derivative is

$$\frac{\partial^2 B(\cdot)}{\partial (m^S)^2} = \eta C''(\cdot) + \frac{\left[\rho'(\cdot)\eta C''(\cdot) + \rho(\cdot)\eta C'''(\cdot)\right] \rho'(\cdot) - \rho(\cdot)\eta C''(\cdot)\rho''(\cdot)}{\left[\rho'(\cdot)\right]^2}
+ \frac{\left[\rho'(\cdot)\eta C''(\cdot) + \rho(\cdot)\eta C'''(\cdot)\right] \rho'(\cdot)}{\left[\rho'(\cdot)\right]^2}
- \frac{2\rho(\cdot)\eta C'(\cdot)(-\rho''(\cdot))\rho'(\cdot)\rho''(\cdot)}{\left[\rho'(\cdot)\right]^4}. \tag{42}$$

Since $\rho'(\cdot) > 0$ and $\rho''(\cdot) \leq 0$, $\partial^2 B(\cdot)/\partial (m^S)^2 > 0$. Accordingly, $B(\cdot)$ is convex increasing in $m^S$, which implies that the first-order approach is sufficient.

The supervisor extracts a rent if $R(m^{S*}, \eta) \equiv B(m^{S*}, \eta) - \eta C(m^{S*}) > 0$. Suppose for a moment that he implements $m^S = 0$. Since $\rho(0) = 0$ and consequently, $B(0, \eta) = 0$, it follows that $R(0, \eta) = 0$. Moreover, the first derivative of $R(\cdot)$ with respect to $m^S$ yields

$$\frac{\partial R(\cdot)}{\partial m^S} = \frac{\rho(m^S)\eta C''(m^S)}{\rho'(m^S)} + \frac{\rho(m^S)\eta C'(m^S)(-\rho''(m^S))}{\left[\rho'(m^S)\right]^2}, \tag{44}$$

which is strictly positive for all $m^S > 0$. In total, $R(0, \eta) = 0$ and $\partial R(\cdot)/\partial m^S > 0$ imply that $R(\cdot) > 0$ for all $m^{S*} > 0$. Finally, substituting $\alpha^{S*} = 0$ and $m^{S*}$ in the principal’s objective function gives $I^D(m^{S*}, \tilde{\varphi}, \eta).$ \hfill \Box

**Proof of Proposition 4.** Implicitly differentiating $m^{S*}$ with respect to $\tilde{\varphi}$ yields

$$\frac{dm^{S*}}{d\tilde{\varphi}} = -\frac{1}{4} ||\mu||^2 \left[ 2 \cos(2\varphi(\cdot))(-\varphi_{m^S})\varphi_{\tilde{\varphi}} + \sin(2\varphi(\cdot))(-\varphi_{m^S}) \right]
- \frac{\partial B(\cdot)}{\partial m^S} \frac{1}{4} ||\mu||^2 \sin(2\varphi(m^S, \tilde{\varphi}))(-\varphi_{m^S}) - \frac{\partial B(\cdot)}{\partial m^S}. \tag{45}$$
The denominator is strictly negative due to the second-order condition, whereas the numerator is strictly positive for all $m^S > 0$ and $\phi \in [0, \pi/2)$. Hence, $dm^{S*}/d\phi > 0$. Next, implicitly differentiating $m^{S*}$ with respect to $\eta$ yields

$$\frac{dm^{S*}}{d\eta} = -\frac{C'(m^S)\rho(m^{S*})}{\rho'(m^S)} - \frac{\rho(m^S)C'(m^S)(-\rho''(m^S))}{[\rho'(m^S)]^2}.$$  

(46)

Since the numerator and the denominator are strictly negative, $dm^{S*}/d\eta < 0$. Moreover, applying the Envelope Theorem yields

$$\frac{d\Pi_D(\cdot)}{d\eta} = -\frac{\partial B(\cdot)}{\partial \eta} = -\frac{C'(m^S)\rho(m^S)}{\rho'(m^S)},$$

(47)

which is strictly negative for all $m^{S*} > 0$. □

**Proof of Proposition 5.** Note first that $\eta \geq 1$ implies $B(m^S) > \eta C(m^S) \geq C(m)$ for all $m^S \geq m$. Thus, a central investment is superior. In contrast, if $\eta = 0$, the supervisor can be induced to implement $m^S \to \infty$ by providing $B(m^S, 0) = 0$. In this case, delegation is strictly superior. Moreover, recall from **Proposition 4** that $d\Pi_D(\cdot)/d\eta < 0$. Thus, there exists a cutoff $\eta^* \in (0, 1)$ such that the principal prefers delegation for $\eta \leq \eta^*$, and a central investment, otherwise.

Next, observe that $\eta^*$ implies $F = \Pi_D(m^{S*}(\tilde{\phi}, \eta^*)), \tilde{\phi}, \eta \Pi_C(m^{*}(\tilde{\phi}), \tilde{\phi}) = 0$. Applying the Implicit Function Theorem gives $d\eta^*/d\tilde{\phi} = -(\partial F/\partial \tilde{\phi})/\partial F/\partial \eta)$. To obtain $\partial F/\partial \tilde{\phi}$, I apply the Envelope Theorem to $\Pi_C(\cdot)$ and $\Pi_D(\cdot)$ separately. Thus,

$$\frac{\partial \Pi_D(\cdot)}{\partial \tilde{\phi}} = -\frac{\partial B(\cdot)}{\partial \phi} = -\frac{C'(m^S)\rho(m^S)}{\rho'(m^S)},$$

(48)

$$\frac{\partial \Pi_C(\cdot)}{\partial \tilde{\phi}} = -\frac{\partial B(\cdot)}{\partial \phi} = -\frac{C'(m^S)\rho(m^S)}{\rho'(m^S)}.$$  

(49)

Since $\partial F/\partial \tilde{\phi} = \partial \Pi_D(\cdot)/\partial \tilde{\phi} - \partial \Pi_C(\cdot)/\partial \tilde{\phi}$,

$$\frac{\partial F}{\partial \tilde{\phi}} = -\frac{\mu^2}{4}[\sin(2\varphi(m^*(\cdot), \cdot)]\varphi_{m^*} - \sin(2\varphi(m^{S*}(\cdot), \cdot)]\varphi_{m^{S*}}],$$

(50)

which is strictly positive if

$$\frac{\sin(2\varphi(m^{S*}(\cdot), \tilde{\phi}))}{\sin(2\varphi(m^{S*}(\cdot), \tilde{\phi}))} > \frac{\varphi_{m^{S*}}}{\varphi_{m^{S*}}}. 

(51)

Suppose for a moment that $m^{S*}(\cdot) = m^*(\cdot)$, which implies $\varphi(m^{S*}(\cdot), \tilde{\phi}) < \varphi(m^*(\cdot), \tilde{\phi})$. Thus, $\sin(2\varphi(m^{S*}(\cdot), \tilde{\phi})) < \sin(2\varphi(m^*(\cdot), \tilde{\phi}))$ for $\varphi(\cdot) \in (0, \pi/2)$. Moreover, observe that $m^{S*}(\cdot) = m^*(\cdot)$ implies $\varphi_{m^*} > \varphi_{m^{S*}}$. As a result, (51) is satisfied for $m^{S*}(\cdot) > m^*(\cdot)$, which further implies that $\partial F/\partial \tilde{\phi} > 0$. In contrast, $\partial F/\partial \tilde{\phi} \leq 0$ if $m^{S*}(\cdot) \leq m^*(\cdot)$.

Furthermore, applying the Envelope Theorem yields

$$\frac{\partial F}{d\eta} = -\frac{\partial B(\cdot)}{d\eta} = -\frac{\rho(\cdot)C'(\cdot)}{\rho'(\cdot)},$$

(52)

which is strictly negative for all $m^{S*}(\cdot) > 0$. As a result, $d\eta^*/d\tilde{\phi} > 0$ if $m^{S*}(\cdot) > m^*(\cdot)$, and $d\eta^*/d\tilde{\phi} \leq 0$, otherwise. Finally, it remains to identify a condition implying $m^{S*}(\cdot) > m^*(\cdot)$. Suppose for a moment that $m^S = m^*(\cdot)$, i.e. the supervisor would implement the same measurement intensity as the principal does in the optimum. Then, $m^{S*}(\cdot) > m^*(\cdot)$ if the slope of $B(\cdot)$ in $m^S = m^*(\cdot)$ is strictly smaller than the slope of $C(m^*(\cdot))$:

$$\frac{\partial B(\cdot)}{dm^S}_{m^S=m^*(\cdot)} < C'(m^*(\cdot)).$$

(53)
For parsimony, let \( m^* \equiv m^*(\tilde{\phi}) \). By utilizing the marginal bonus derived in proof of Proposition 3, one can show that (53) is equivalent to

\[
\frac{\rho(m^*)C''(m^*)}{\rho'(m^*)} + \frac{\rho(m^*)C'(m^*)}{[\rho'(m^*)]^2} \left[ -\rho''(m^*) \right] < \frac{1 - \eta}{\eta} C'(m^*) \quad (54)
\]

\[
\Leftrightarrow \quad \frac{\rho(m^*)}{\rho'(m^*)} \left[ \frac{C''(m^*)}{C'(m^*)} - \frac{\rho''(m^*)}{\rho'(m^*)} \right] < \frac{1 - \eta}{\eta}. \quad (55)
\]

Consequently, \( m^{S*}(\cdot) > m^*(\cdot) \) if (55) is satisfied, and \( m^{S*}(\cdot) \leq m^*(\cdot) \), otherwise. \( \square \)

**Proof of Proposition 6.** For parsimony, define

\[
A \equiv \frac{1}{4} \mu^2 \sin \left( 2\phi \left( m + m^S, \tilde{\phi} \right) \right) \left[ -\frac{\partial \phi}{\partial m} \right] - C'(m), \quad (56)
\]

\[
B \equiv \frac{1}{4} \mu^2 \sin \left( 2\phi \left( m + m^S, \tilde{\phi} \right) \right) \left[ -\frac{\partial \phi}{\partial m^S} \right] - \frac{\partial B(m^S, \eta)}{\partial m^S}. \quad (57)
\]

Applying Cramer’s Rule to (56) and (57) yields

\[
\frac{\partial m}{\partial \tilde{\phi}} = \frac{\frac{\partial A}{\partial \tilde{\phi}} \frac{\partial B}{\partial m^S} - \frac{\partial B}{\partial \tilde{\phi}} \frac{\partial A}{\partial m^S}}{\det(H)}, \quad (58)
\]

\[
\frac{\partial m^S}{\partial \tilde{\phi}} = \frac{-\frac{\partial B}{\partial \tilde{\phi}} \frac{\partial A}{\partial m} + \frac{\partial A}{\partial \tilde{\phi}} \frac{\partial B}{\partial m}}{\det(H)}. \quad (59)
\]

where

\[
H = \begin{pmatrix}
\frac{\partial^2 C}{\partial m^2} & \frac{\partial^2 C}{\partial m \partial m^S} \\
\frac{\partial^2 C}{\partial m \partial m^S} & \frac{\partial^2 C}{\partial m^S} \\
\frac{\partial^2 C}{\partial m^2} & \frac{\partial^2 C}{\partial m^S} \\
\end{pmatrix} . \quad (60)
\]

Since \( II \) is concave, \( H \) must be negative definite, and thus \( \det(H) > 0 \). It can be inferred from (56) and (57) that \( \frac{\partial A}{\partial \tilde{\phi}} = \frac{\partial B}{\partial \tilde{\phi}} \) and \( \frac{\partial A}{\partial m^S} = \frac{\partial B}{\partial m} \). Hence,

\[
\frac{\partial \Delta m}{\partial \tilde{\phi}} = \frac{\partial \left( m - m^S \right)}{\partial \tilde{\phi}} = \frac{-\frac{\partial A}{\partial m} \frac{\partial B}{\partial m^S} + \frac{\partial A}{\partial \tilde{\phi}} \frac{\partial B}{\partial m}}{\det(H)}. \quad (61)
\]

Since

\[
\frac{\partial A}{\partial m} = \frac{\partial a}{\partial m} - C''(m), \quad \frac{\partial B}{\partial m} = \frac{\partial b}{\partial m} - \frac{\partial^2 B(\cdot)}{\partial (m^S)^2} ,
\]

where \( \partial a/\partial m = \partial b/\partial m^S \), it follows

\[
\frac{\partial \Delta m}{\partial \tilde{\phi}} = \frac{\frac{\partial A}{\partial \tilde{\phi}} \left[ \frac{\partial^2 B(\cdot)}{\partial (m^S)^2} - C''(m) \right]}{\det(H)}. \quad (62)
\]

It can be verified that \( \partial A/\partial \tilde{\phi} > 0 \). Hence, \( \partial \Delta m/\partial \tilde{\phi} > 0 \) if \( \partial^2 B(\cdot)/\partial (m^S)^2 > C''(m) \), which in turn depends on the magnitudes of \( m^*(\tilde{\phi}, \eta) \) and \( m^{S*}(\tilde{\phi}, \eta) \) in response to \( \tilde{\phi} \). Since the supervisor’s rent is increasing in \( m^S \), there exists a cut off \( \phi^*(\eta) \) such that \( \partial^2 B(\cdot)/\partial (m^S)^2 > C''(m) \) for \( \phi > \phi^*(\eta) \), and vice versa. Finally, since the supervisor’s rent is increasing in \( \eta \), \( \partial \phi^* / \partial \eta < 0 \). \( \square \)
References