Objective versus Subjective Performance Evaluations

Stefan Terstiege*

March 20, 2012

Abstract

Why does incentive pay often depend on subjective rather than objective performance evaluations? After all, subjective evaluations entail a commitment problem. While the common explanation for this practice assumes lack of adequate objective measures, I argue that subjective evaluations might be used to withhold information from the worker. I also argue that withholding information is particularly important under circumstances where the commitment problem is small. The statements are derived from a two-stage principal-agent model in which the stochastic relationship between effort and performance is unknown.

Keywords: Performance evaluation, compensation, principal-agent, moral hazard

JEL Codes: D83, D86, J41, M12, M52

---

*University of Bonn, Bonn Graduate School of Economics, Adenauer Allee 24-42, D-53113 Bonn (Germany), e-mail: sterstiege@uni-bonn.de. I thank Markus Fels, Daniel Krähmer, Matthias Kräkel, Matthias Lang and Sina Litterscheid for very helpful comments.
1 Introduction

Many employers try to boost employee morale by relating compensation—pay, career, power, etc.—to performance. This paper examines how a worker’s performance should be measured. One possibility is to establish in advance a comprehensible, visible evaluation procedure that disregards expertise of biased persons (objective evaluation). Alternatively, performance can be rated by the personal impression of the worker’s supervisor, which is unverifiable (subjective evaluation). As an illustration, consider the evaluation of a cook’s work. The restaurant owner may either ask the guests to rank the meal on a scale from one to ten, or taste herself.

If an employee’s compensation depends on subjective performance evaluations, it is at the discretion of his supervisor. Honest evaluations come at a cost. An owner-manager, for instance, who would generally be tempted to understate performance and save on labor costs, might be forced to acquire credibility through obligatory bonus pools (paid out to some third party in case of negative evaluation, e.g., to charity) or up-or-out career systems (employee lost in case of negative evaluation).\(^1\) On the other hand, if compensation depends on an objective evaluation, it is court-enforçable, and the employer does not incur such commitment costs. Yet, subjective evaluations are common practice.\(^2\)

A usual explanation for this fact holds that employers use subjective performance measures to obtain more accurate evaluations of their workers. More precisely, they complement objective measures which do not capture all tasks that are to be carried out, or which are subject to influences beyond the agent’s control. For if the evaluation is incomplete, the agent might “game”

\(^1\)The same problem occurs with an employed manager, whose interests should be aligned with the owner’s objectives. Note that supervisors can also be tempted to overstate performance so as not to disgustle employees, or when they are susceptible to bribery or currying favor by employees. See Prendergast (1999) for a discussion. Finally, discretion over compensation allows supervisors to discriminate workers based on sex, nationality, etc.

\(^2\)Due to the availability of data, most studies on performance reviews consider CEO compensation. E.g., Bushman, Indjejikian and Smith (1996) found that in 190 of 248 firms in the USA, bonuses were (at least partly) at the discretion of the board of directors. Murphy and Oyer (2003) examined 280 firms in the USA: 43% displayed discretion in determining the size of bonus pools, and 67% in allocating a bonus pool across participants.
the incentive contract and neglect tasks that are not included in the evaluation. In fact, empirical studies document that subjective evaluations are more likely in jobs which comprise many tasks (see Brown 1990 and MacLeod and Parent 1999). On the other hand, if the evaluation is noisy, a risk-averse agent might not respond to incentive pay.

However, employers conduct subjective evaluations often not out of necessity, as this theory suggests, but *eschew* to establish comprehensive objective performance measures. The following list provides some examples:

- In 2003, the National Research Council (NRC), a think tank, examined the quality of project management within the U.S. Department of Energy. It stated a “lack of objective measures that makes it difficult to assess progress in improving project management”, and “to build confidence within [...] Congress [...] and the public in the department’s ability to manage the money it spends on its projects. Evidence continues to be anecdotal rather than objective, quantitative, and verifiable.” A subsequent report by the NRC suggested a number of appropriate measures.

- A study on 17 U.S. investment banks revealed that employees in sales and trading divisions receive bonuses that largely depend on subjective performance appraisals despite “the ease with which the profitability of an individual trader can be measured each day.” All 17 banks had implemented bonus pools. Similarly, Prendergast (1999) notes that “there are many measures of the productivity of [...] a baseball player (batting average, home runs, ...

---

3See Baker, Gibbons and Murphy (1994), Prendergast (1999) and the general analysis of Bernheim and Whinston (1998). This argument draws on the multi-tasking problem described by Holmstrom and Milgrom (1991) and Baker (1992). E.g., as Bushman, Indjejikian and Smith (1996) note, a CEO has to plan the long-term strategy of the company. If it is concealed from the financial market, his performance should not only be measured with the stock price.

4See Rajan and Reichelstein (2009). This argument draws on the informativeness principle, established by Holmstrom (1979) and Harris and Raviv (1979).


6See National Research Council (2005).

etc.), yet explicit contracts are rarely written on those measures.”

- The standards for associates to become partner at professional service firms are usually highly intransparent, although many explicit performance measures are conceivable. (According to Morris and Pinnington (1998), who surveyed law firms in the UK, the most important promotion criteria are: getting new business, fee-earning ability, technical skill and getting on with clients.) At such organizations, up-or-out career systems are commonplace.

I offer an alternative explanation, according to which subjectivity itself can be a desirable property of a performance measure. It builds on the notion that evaluations generate information, and that only subjective evaluations can be concealed for a while. My theory applies to jobs with many tasks that are carried out sequentially (e.g., phases of project management), and, more generally, to long-term employment relationships in which performance can be assessed over time (e.g., after trading days, matches of season, or fiscal years until promotion decision). In a nutshell, subjective evaluations at early stages of such jobs are advantageous in terms of profit and surplus if details like productivity, ability or profitability are highly uncertain and can be inferred from performance. This is because the worker could manipulate the incentive scheme if he learned about these details too soon, and because commitment costs are actually low.

I derive my statements from a two-stage principal-agent model with hidden actions. Specifically, in each stage the wealth-constrained agent can exert effort to increase the likelihood of good performance, but effort is not observable. The principal can use incentive pay for profit maximization. Three assumptions are important: (1) the likelihood of good performance does not only depend on effort, but also on an unknown, persistent parameter, (2) surplus maximization requires effort in each stage, and (3) profit maximization inevitably requires effort in stage 2.

With this model, I compare profit and surplus in two scenarios that differ as to how the principal evaluates performance. In the objective scenario, she evaluates objectively in both stages. That means, it becomes public information at the end of each stage whether the agent was successful.

---

9 See Gilson and Mnookin (1989).
10 Put differently, concealed evaluations might become unverifiable. For example, restaurant guests cannot be asked to judge a meal again months after they ate it.
In the *subjective scenario*, the principal evaluates subjectively in stage 1. Here, only she herself learns the outcome of stage 1, and it cannot be verified.

The analysis yields that in each scenario, the principal cannot appropriate the entire surplus if the agent is to exert effort in both stages. In the objective scenario, incentive contracts must take into account that the evaluation in stage 1 provides information about the unknown parameter, and that the agent is in a better position to draw inference since he privately knows his effort choice. In particular, bad performance in stage 1 indicates an unfavorable state, and thus calls for amplified incentive pay in stage 2. But this scheme must not tempt the agent to produce a failure in stage 1 on purpose. Therefore, the principal provides the same high-powered incentive also when stage 1 was successful. As a consequence, the agent can secure a rent. In the subjective scenario, on the other hand, the principal optimally reveals her subjective evaluation of stage 1 only after stage 2, and only after particular histories. Incentive pay remains invariant over time, and the agent does not receive a rent. However, the principal’s discretion in determining the agent’s performance bonus for stage 1 implies a credibility issue. Ex post, the principal will always submit an evaluation which minimizes labor costs. She must therefore commit to pay the performance bonus regardless of her evaluation—to the agent if he was successful, and otherwise to a budget breaker. As a result, the subjective scenario entails commitment costs, which are due when the principal submits a negative rating for stage 1.

In each scenario, the principal might consequently induce the agent to work hard only in stage 2, which is inefficient. I show that greater uncertainty about the stochastic relationship between effort and performance leads to a higher rent but lower commitment costs; i.e., it renders the subjective scenario advantageous in terms of profit and surplus. Intuitively, this result can be explained as follows. The model captures greater uncertainty in form of more extreme posteriors about the unknown parameter. In particular, failure in stage 1 results in a more pessimistic posterior, so that the rent in the objective scenario increases. In the subjective scenario, in contrast, the principal benefits from a more pessimistic posterior after failure. This is because she only reveals her subjective rating of stage 1 when the agent performed well in stage 2. Conditional on that event, stage 1 is more likely to have been successful as well.

According to this paper, employers thus establish subjective performance measures to withhold
information from their workers, and objective measures if verifiability is the major concern. Several papers indicate similar pros and cons of subjective and objective evaluations, but do not explore the trade-off. Most closely related are recent papers by Bashkar (2009), DeMarzo and Sannikov (2011) and Kwon (2011), who study dynamic moral hazard problems in which performance is publicly observed. As in my model, both the principal and the agent do not know the stochastic relationship between effort and output, and learn about it from past performance. The papers show that incentive pay involves extra costs for the principal since just the agent knows past effort.\(^{11}\) It is not examined whether the principal would gain from subjective evaluations.

Closely related is also the analysis of interim feedback during long-term relationships by Lizzeri, Meyer and Persico (2002), who consider a dynamic moral hazard problem in which the principal privately observes the agent’s performance.\(^{12}\) They find that the principal should better not reveal the agent’s interim performance before the job is finished since she would otherwise have to pay more rent for the same effort provision. (Different from my model, an interim evaluation does not have any relevant informational content, but only determines the agent’s continuation payoff. If it is concealed, more effective carrot-and-stick schemes are possible.) However, the principal can verify her evaluation, i.e., it is not subjective.

Incentive contracts with subjective performance measures are studied in the seminal papers by MacLeod (2003), Levin (2003) and Fuchs (2007). They consider situations in which the principal does not benefit from private instead of public information about performance, but has no means to evaluate objectively. It is analyzed in detail how lack of verifiability exacerbates the contracting problem. Among these papers, Fuchs (2007) is most closely related to the present study. He also investigates a finitely repeated moral hazard problem and argues that subjective evaluations should only be revealed after the last stage. (As with Lizzeri, Meyer and Persico (2002), evaluations do not have relevant informational content in his model. If the agent is kept uninformed about his continuation payoff, incentive pay involves more compressed wages, so that the truth-telling constraints for the principal are less severe.)

\(^{11}\) Hirao (1993) and Manso (2011) consider similar settings in which the principal avoids the extra costs by either terminating the employment relationship or implementing a different technology after a failure.

\(^{12}\) For more recent references in this strand of literature, see footnote 21.
This paper is organized as follows. The next section presents the model. Section 3 first derives the optimal contract for the objective and the subjective scenario, respectively. Afterwards, the two scenarios are compared with respect to the principal’s profit and the generated surplus, and it is shown that more uncertainty renders the subjective scenario advantageous. Section 4 concludes.

2 The model

At some initial date 0, a principal in need for a project manager is matched to an agent. The quality of the match is determined by a random variable, whose distribution function is denoted by $F$ and whose realization $\theta$ lies in the interval $[0, 1]$. Both parties do not know which state obtains.

Project management comprises two stages, and the final return to the principal depends on the agent’s performance in each stage. At date 1, stage 1 begins, in which the agent can exert effort $e_1 \in \{0, 1\}$ at costs $e_1c_1$, where $c_1 > 0$. (For example, one may think of stage 1 as the planning stage, where the agent is to design a blueprint of the final good that meets the principal’s needs.) The agent’s effort provision is not observable (hidden action). However, the principal can evaluate the agent’s performance, which essentially means to acquire information about output that would remain unknown otherwise. I distinguish two scenarios, which differ as to how the agent’s performance in stage 1 is evaluated. In the objective scenario, an objective evaluation takes place. This means that performance is verifiable and commonly observed at the end of the stage. In the subjective scenario, on the other hand, performance is not verifiable and privately observed by the principal at the end of the stage (hidden information). Therefore, only her subjective assessment is available. To concentrate the analysis on the trade-off between subjectivity and objectivity, suppose that the stochastic relationship between effort and performance does not depend on the nature of the evaluation. Thus, the performance measures in the two scenarios do not differ in their accuracy, but only as to whether they are comprehensible for the agent and any third party. Formally, performance is good ($x_1 = 1$) with probability $e_1\theta$, and bad ($x_1 = 0$) with probability $1 - e_1\theta$.\(^\text{13}\)

\(^\text{13}\)The assumptions that the agent himself cannot assess performance and that shirking yields failure with certainty
At date 2, the project enters stage 2. Here, the agent can exert effort \( e_2 \in \{0, 1\} \) at costs \( e_2 c_2 \), where \( c_2 > 0 \). (E.g., stage 2 may stand for the execution stage, in which the agent is to produce the good according to the blueprint in a timely, cost-effective manner.) Again, the principal cannot monitor whether the agent shirks. In both scenarios, the agent’s performance in stage 2 is publicly observable at the end of that stage. It is again either good \( (x_2 = 1) \) or bad \( (x_2 = 0) \); the good outcome obtains with probability \( e_2 \theta \).

Finally, when stage 2 is completed, the project yields the principal a return of \( R(x_1, x_2) \). I normalize \( R(0, 0) = 0 \) and assume

\[
R(1, 1) - R(0, 1) = R(1, 0) - R(0, 0) = R_1 > 0
\]

\[
R(1, 1) - R(1, 0) = R(0, 1) - R(0, 0) = R_2 > 0.
\]

This parameterization captures that the principal prefers good performance, and that the two stages are technologically independent in that \( x_1 \) and \( x_2 \) are neither substitutes nor complements with respect to project return. The latter property implies that the evaluation of stage 1 cannot be replaced by an inspection of the return generated in stage 2. Note that in the subjective scenario, where just the principal observes the outcome of stage 1, return is only partially verifiable.

Both parties are risk neutral, do not discount future payoffs, and have an outside option of zero. In particular, the principal cannot hire a different agent for stage 2. At date 0, the principal, having full bargaining power, offers a contract to specify the agent’s wage. It may be contingent on any verifiable data concerning the agent’s performance, possibly involving messages. The agent does not dispose of own resources, so that only non-negative wages are feasible. In case the contract requires a budget breaker, a third party is available to the principal. All payments are made after stage 2 is completed.

According to the previous description, the timing of the model can be summarized as follows:

0. Match quality \( \theta \) realizes, and the principal offers a contract.

1. Stage 1: The agent exerts effort \( e_1 \). Then, performance \( x_1 \) realizes. In the objective scenario,

are not critical for the qualitative nature of the results. I could alternatively require that the principal’s evaluation generates some piece of information which is not accessible for the agent (e.g., a comparison with other workers) and that shirking yields success with a probability less than \( \theta \).
$x_1$ is commonly observed. In the subjective scenario, $x_1$ is privately observed by the principal.

2. Stage 2: The agent exerts effort $e_2$. Then, performance $x_2$ realizes. In both scenarios, $x_2$ is commonly observed. Finally, payments are made.

The agent’s performance in stage 1 conveys information about the unknown quality of the match, which can be used to update the initial expectation $\mu = E[\theta]$. However, one has to take the agent’s effort decision in stage 1 into account to interpret this information. If the agent exerted effort ($e_1 = 1$), the posterior expectation of match quality is either $\mu_1$ (in case $x_1 = 1$) or $\mu_0$ (in case $x_1 = 0$). Henceforth, assume $1 > \mu_1 > \mu > \mu_0 > 0$.\footnote{Precisely, $\mu = \int_0^1 \theta dF$, $\mu_1 = (1/\mu) \int_0^1 \theta^2 dF$ and $\mu_0 = [1/(1 - \mu)] \int_0^1 \theta(1 - \theta) dF$. The ordering of the posteriors holds for instance if the distribution function $F$ is discrete and assigns positive probability to at least three distinct realizations, or if $F$ has a density that is supported on any nondegenerate subinterval of $[0,1]$.} If the agent was lazy ($e_1 = 0$), his performance is bad irrespective of the quality of the match, so that no information is conveyed and the posterior expectation remains $\mu$. An important property of the model is that the two parties might have to base their posteriors on imperfect information: in both scenarios, the principal does not observe $e_1$; the agent, on the other hand, does not observe $x_1$ in the subjective scenario.

In the next section, I compare the two scenarios with respect to the principal’s profit and the generated surplus. As usual, the analysis involves a two step procedure to determine optimal contracts (cf. Grossman and Hart 1983). The first step is to compute for each effort plan $(e_1, e_2(x_1))$ the cheapest contract that implements it. The second step is to identify among these contracts the profit-maximizing one. Let $(e_1, e_2)$ denote a plan which specifies the same effort level for stage 2 after both possible outcomes of stage 1. To concentrate the analysis on the relevant circumstances, I make the following assumptions.

**Assumption 1.** In a first-best world (i.e., if $e_1$, $e_2$, $x_1$ and $x_2$ were verifiable), the principal would implement effort plan $(1,1)$.

**Assumption 2.** In both scenarios, the principal wants to implement $e_2 = 1$ regardless of the agent’s performance in stage 1.

In the following, I consider the sum of payoffs to principal, agent and budget breaker as the
generated surplus.\textsuperscript{15} By assumption 1, effort plan \((1, 1)\) maximizes this surplus. Straightforward calculations show that the assumption is equivalent to \(\mu R_1 - c_1 \geq 0\) and \(\mu_0 R_2 - c_2 \geq 0\). Assumption 2, on the other hand, requires a sufficiently high value \(R_2\) (the analysis will not need the precise threshold). It says that the optimal contract either implements \((1, 1)\) or the inefficient plan \((0, 1)\). In particular, the principal does not want to make \(e_2\) contingent on the agent’s past performance. However, the evaluation in stage 1 is required to create incentives for that stage. I want to find out whether subjective or objective evaluations are more appropriate.

3 The analysis

In each stage, performance is bad for sure when the agent shirks. If the model did not involve uncertainty about the quality of the match (or, more generally, about the stochastic relationship between effort and performance), this condition would imply that in the objective scenario, the principal implements the efficient effort plan \((1, 1)\) and appropriates the entire surplus.\textsuperscript{16} On the other hand, when the principal must resort to a subjective evaluation as in the subjective scenario, incentive contracts typically distribute some surplus to a budget breaker to overcome the problem of hidden information (see, e.g., MacLeod 2003). In that case, the principal might well implement an inefficient effort plan to maximize the share of surplus which she can retain; hence, both from the principal’s perspective and from a welfare perspective the objective scenario is preferred when the quality of the match is uncertain. The analysis will show that this ranking might be reversed under uncertainty.

3.1 Objective scenario

In the objective scenario, the agent’s performance is objectively evaluated in both stages, so that just the problem of hidden action occurs. Optimal contracts neither involve communication nor

\textsuperscript{15}Thus, the efficiency analysis focuses on production and ignores distribution.

\textsuperscript{16}If match quality was known to be \(p\), for instance, the principal could achieve this aim by paying \(c_1/p\) and \(c_2/p\) to the agent if he is successful in stage 1 and 2, respectively.
a budget breaker, but consist of a performance-contingent wage scheme.

I begin the two step procedure of finding the optimal contract by calculating the cheapest wage scheme which implements effort plan \((0, 1)\). In the objective scenario, this plan is somewhat delicate. Even if the principal does not reward good performance in stage 1, the agent might exert effort in order to acquire information about the unknown quality of the match for strategic reasons (i.e., to find out whether it pays off to work hard in the next stage).\(^{17}\) That is, the principal could be in a situation where she has to reward laziness. To rule out this issue, assume the principal has the means to cancel stage 1 altogether. The agent’s wage will then only depend on his performance in stage 2.

A wage scheme \(w = (w_x)_{x \in \{0, 1\}}\) that motivates the agent to work hard only in stage 2 yields the principal an expected profit of \(\mu R_2 - C(w)\), where \(C(w)\) denotes the principal’s expected payment.\(^{18}\) The agent is willing to exert effort whenever

\[
\mu w_1 + (1 - \mu)w_0 - c_2 \geq w_0.
\]

He accepts the contract if

\[
\mu w_1 + (1 - \mu)w_0 - c_2 \geq 0.
\]

In particular, he accepts if wages are non-negative, which has to be the case due to his limited liability. Hence, the cheapest wage scheme that implements effort plan \((0, 1)\) solves

\[
\min_{w \geq 0} C(w) = \mu w_1 + (1 - \mu)w_0 \quad s.t. \quad (1).
\]

It is straightforward to verify that the principal should offer \((w_0, w_1) = (0, c_2/\mu)\). With this wage scheme, the principal appropriates the entire surplus of stage 2, denoted by

\[
S_2 = \mu R_2 - c_2.
\]

**Lemma 1.** In the objective scenario, the cheapest contract to implement effort plan \((0, 1)\) yields the principal a profit of \(S_2\), while the agent’s payoff is zero, and the budget breaker is not involved.

\(^{17}\)In the context of contracting with endogenous information, strategic information acquisition has been studied extensively; see Crémer and Khalil (1992), Crémer, Khalil and Rochet (1998a) and Crémer, Khalil and Rochet (1998b).

\(^{18}\)In the following, all payoffs should be understood as expected payoffs.
I proceed with the efficient effort plan \((1,1)\). A wage scheme \(w = (w_{x_1 x_2})_{x_1, x_2 \in \{0, 1\}}\) that motivates the agent to work hard in both stages yields the principal an expected profit of \(\mu(R_1 + R_2) - C(w)\). There are now four incentive-compatibility constraints to satisfy. First, the agent has to exert effort in stage 2 if he was successful in the first stage:

\[
\mu_1 w_{11} + (1 - \mu_1) w_{10} - c_2 \geq w_{10}.
\] (2)

Second, he must also work hard if his performance was bad in stage 1, holding the pessimistic posterior \(\mu_0\):

\[
\mu_0 w_{01} + (1 - \mu_0) w_{00} - c_2 \geq w_{00}.
\] (3)

Third, the agent must exert effort in stage 1 given that he will do so in stage 2:

\[
\mu[\mu_1 w_{11} + (1 - \mu_1) w_{10} - c_2] + (1 - \mu)[\mu_0 w_{01} + (1 - \mu_0) w_{00} - c_2] - c_1 \geq \mu w_{01} + (1 - \mu) w_{00} - c_2.
\] (4)

Fourth, shirking in both stages has to be unprofitable:

\[
\mu[\mu_1 w_{11} + (1 - \mu_1) w_{10} - c_2] + (1 - \mu)[\mu_0 w_{01} + (1 - \mu_0) w_{00} - c_2] - c_1 \geq w_{00}.
\] (5)

Finally, the contract should be acceptable for the agent at date 0:

\[
\mu[\mu_1 w_{11} + (1 - \mu_1) w_{10} - c_2] + (1 - \mu)[\mu_0 w_{01} + (1 - \mu_0) w_{00} - c_2] - c_1 \geq 0.
\]

Since \(\mu > \mu_0\), this constraint is automatically satisfied with strict inequality by (4), (3) and the limited liability condition. Put differently, the agent can secure a rent if the principal wants to implement effort plan \((1,1)\) in the objective scenario. This is because, if he complies, his performance in stage 1 conveys information about the unknown match quality. More specifically, bad performance gives rise to the pessimistic posterior \(\mu_0\). Since the agent must continue to work hard, the principal has to offset the demotivating experience by a high reward for good performance in stage 2. But if the agent actually failed in stage 1 because he shirked, no information is conveyed via his performance, so that he is not pessimistic to be successful through hard work in stage 2. However, the principal deems the outcome of stage 1 to be informative since, in equilibrium, there is no shirking; she provides the same high-powered incentive whenever performance in stage 1 is bad. Accounting for limited liability, the agent must therefore get a strictly positive payoff if he
deviates from the effort plan \((1, 1)\) and works according to \((0, 1)\). Hence, it requires a rent to make \((1, 1)\) incentive-compatible.\(^{19}\)

For the same reason as the participation constraint, the incentive-compatibility condition (5) does not bind either. The cheapest wage scheme which implements effort plan \((1, 1)\) thus solves

\[
\min_{w \geq 0} C(w) = \mu[\mu_1 w_{11} + (1 - \mu_1)w_{10}] + (1 - \mu)[\mu_0 w_{01} + (1 - \mu_0)w_{00}] \quad \text{s.t.} \quad (2), (3) \text{ and } (4).
\]

**Lemma 2.** To implement effort plan \((1, 1)\) in the objective scenario as cheap as possible, the principal should pay wages according to

\[
(w_{00}, w_{10}, w_{01}, w_{11}) = \left(0, \frac{c_1}{\mu}, \frac{c_2}{\mu_0}, \frac{c_1}{\mu} + \frac{c_2}{\mu_0}\right).
\]

*Proof.* First, reformulate the constraints as

\[
\mu_1(w_{11} - w_{10}) \geq c_2, \quad (2')
\]

\[
\mu_0(w_{01} - w_{00}) \geq c_2, \quad (3')
\]

\[
\mu[\mu_1(w_{11} - w_{10}) + w_{10} - \mu_1(w_{01} - w_{00}) - w_{00}] \geq c_1. \quad (4')
\]

It is now routine to verify that \((3')\) and \((4')\) bind, and that the proposed wage scheme indeed solves the program. \(\Box\)

The wage scheme in Lemma 2 rewards performance in stage 2 independently of performance in stage 1. In particular, wages do not condition on the principal’s posterior expectation of match quality, so that the agent has no incentive to manipulate it. Relative to this posterior however, the wage scheme is too high-powered when stage 1 was successful (constraint \((2)\) holds with strict inequality). As a consequence, the principal does not get the entire surplus, but shares it with the agent. The following notation for the generated surplus in stage 1 and the agent’s payoff will be helpful in stating this result:

\[
S_1 = \mu R_1 - c_1,
\]

\[
A = \left(\frac{\mu}{\mu_0} - 1\right)c_2.
\]

\(^{19}\)See Bashkar (2009), DeMarzo and Sannikov (2011) and Kwon (2011) for similar observations about incentive costs in more general settings.
Lemma 3. In the objective scenario, the cheapest contract to implement effort plan $(1, 1)$ yields the principal a profit of $S_1 + S_2 - A$. The agent obtains as payoff $A$, and the budget breaker is not involved.

By assumption 2, the optimal contract implements effort in stage 2. Therefore, the second step of the principal’s contracting problem just requires to compare her profits as stated in lemmas 1 and 3. It turns out that the principal has to make a trade-off: while the inefficient plan $(0, 1)$ sacrifices surplus in stage 1 (worth $S_1$), the efficient plan $(1, 1)$ entails rent for the agent (worth $A$).

Proposition 1. In the objective scenario, the optimal contract is as follows:

1. For $S_1 \leq A$, the inefficient effort plan $(0, 1)$ is implemented. The principal makes a profit of $S_2$, the agent gets zero, and the budget breaker is not involved.
2. For $S_1 > A$, the efficient effort plan $(1, 1)$ is implemented. The principal makes a profit of $S_1 + S_2 - A$, the agent gets $A$, and the budget breaker is not involved.

3.2 Subjective scenario

Consider now the subjective scenario, where the agent’s performance in stage 1 is not verifiable and can only be evaluated subjectively by the principal. By assumption 2, it is again the case that the optimal contract either implements effort plan $(0, 1)$ or the efficient plan $(1, 1)$.

I begin the two step analysis by searching for the cheapest contract that implements $(0, 1)$. As in scenario 1, appropriate contracts consist of a wage scheme contingent on the agent’s performance in stage 2; they do not require a budget breaker or communication.\(^{20}\) The principal should therefore offer the same contract as in the objective scenario.

Lemma 4. In the subjective scenario, the cheapest contract to implement effort plan $(0, 1)$ yields the principal a profit of $S_2$, while the agent’s payoff is zero, and the budget breaker is not involved.

\(^{20}\)If the principal does not communicate $x_1$, such a wage scheme will always implement $e_1 = 0$ even if stage 1 cannot be cancelled. Thus, strategic information can be deterred more easily if the agent is evaluated subjectively, because the principal herself decides which information to communicate.
Next, I search for the cheapest contract that implements the efficient effort plan \((1, 1)\). Clearly, to provide an incentive to work hard repeatedly the agent needs to be rewarded depending on his performance in each stage. The principal must therefore make the wage scheme now also contingent on her subjective evaluation since no other performance measure is available for stage 1. As a consequence, the problem of hidden information becomes acute, and it leads to two difficulties. First, due to its subjective nature, the performance evaluation in stage 1 may lack credibility: ex post, the principal could prefer to submit a dishonest evaluation to save on wages. But the agent will shirk if he cannot be sure to receive a reward for good performance. To overcome the difficulty, the principal must commit to involve the budget breaker; transfers to this third party will solve her credibility problem. The second difficulty relates to the fact that this principal-agent relationship comprises two stages, rather than just one. More precisely, it is not clear whether the principal should reveal her evaluation already before stage 2 begins or afterwards, and in which form this should be done.\(^{21}\) I take the following approach to address these questions. First, I derive the cheapest payment scheme (i.e., wages to the agent and transfers to the budget breaker) that implements effort plan \((1, 1)\) under \textit{mediated} talk. The presence of a mediator allows for additional communication protocols, and can only benefit the principal. In a second step, I show that face-to-face communication after stage 2 works as well as mediated talk. The advantage of this approach is that the benchmark case with mediator can be analyzed using the revelation principle for multistage games (Myerson 1986).

So suppose for the moment, a mediator (i.e., an impartial person with whom each party can communicate confidentially) is available to coordinate communication. According to the revelation principle, the cheapest contract to implement effort plan \((1, 1)\) can be found in the class of contracts with the following communication protocol: before stage 2 the principal reports her evaluation

\(^{21}\)Exactly these questions are raised by the literature on interim feedback during long-term relationships. For example, Lizzeri, Meyer and Persico (2002), who consider a situation without uncertainty, find that the principal should not communicate the agent’s interim performance before the project is completed. On the other hand, Suvorov and van de Ven (2009) show that if the principal cannot commit to a wage scheme, coarse feedback in connection with a bonus payment is beneficial. Similarly, in dynamic tournaments it might be advantageous to announce an ordinal midterm ranking (Gershkov and Perry 2009), provide full feedback (Ederer 2010) or partially disclose the participants’ interim performance (Goltsman and Mukherjee 2011).
in form of a verifiable message $m \in \{0,1\}$ to the mediator, who reveals it publicly after stage 2. No further communication takes place; in particular, the agent does not receive interim feedback about his performance in stage 1.\(^{22}\) Moreover, the revelation principle asserts that contracts may without loss of generality specify payments which induce the principal to report her evaluation truthfully to the mediator. In the following, I will restrict attention to contracts with these properties.

To proceed with the analysis, let $w = (w_{m,x})_{m,x \in \{0,1\}}$ be a wage scheme and denote by $b = (b_{m,x})_{m,x \in \{0,1\}}$ a scheme of non-negative transfers from the principal to the budget breaker. Note that $m$ stands for the principal’s reported evaluation, which will be honest in equilibrium. A combination of $w$ and $b$ that implements effort plan $(1, 1)$ yields the principal a profit of $\mu(R_1 + R_2) - C(w + b)$, where $w + b$ specifies her total payment.

The wage scheme must make effort plan $(1, 1)$ incentive-compatible for the agent given that the principal reports truthfully to the mediator. In contrast to the objective scenario, the agent is now uninformed about past performance when choosing an effort level in stage 2. He can only use the prior expectation to estimate the quality of the match. Therefore, the incentive constraint for stage 2 pools the conditions (2) and (3):

$$\mu[\mu_1 w_{11} + (1 - \mu_1)w_{10} - c_2] + (1 - \mu)[\mu_0 w_{01} + (1 - \mu_0)w_{00} - c_2] \geq \mu w_{10} + (1 - \mu)w_{00}. \quad (6)$$

Furthermore, the incentive constraints (4) and (5) must be met to induce hard work in each stage. The agent’s participation is then guaranteed because of the limited liability condition.

To make the wage scheme credible, the principal must indeed be willing to report her evaluation of stage 1 truthfully to the mediator before stage 2 begins. Since the report is not transmitted to the agent until after the project is completed, it cannot affect effort but only payments. Whenever wages alone would tempt the principal to cheat, transfers to the budget breaker need to be specified such that the correct report leads to the lowest total payment for the principal. Formally, in case

---

\(^{22}\)The only interim messages that the agent might generally receive are recommendations by the mediator, which effort to provide. However, in the present case, such messages would be superfluous since the agent is always supposed to work hard.
the principal observes good performance, this requires
\[ \mu_1(w_{11} + b_{11}) + (1 - \mu_1)(w_{10} + b_{10}) \leq \mu_1(w_{01} + b_{01}) + (1 - \mu_1)(w_{00} + b_{00}), \] (7)
while for bad performance, it has to be
\[ \mu_0(w_{01} + b_{01}) + (1 - \mu_0)(w_{00} + b_{00}) \leq \mu_0(w_{11} + b_{11}) + (1 - \mu_0)(w_{10} + b_{10}). \] (8)
Finally, only non-negative transfers \( b \) are feasible.

I have now identified all conditions which a contract must satisfy to implement effort plan \((1, 1)\) under mediated communication. Conditions (4), (5) and (6) rule out deviations by the agent to the effort plans \((0,1)\), \((0,0)\) and \((1,0)\), respectively. Conditions (7) and (8) ensure that the principal submits honest evaluations for stage 1 to the mediator. The cheapest contract results in the following costs:

\[
\min_{w \geq 0, b \geq 0} C(w + b) = \mu[\mu_1(w_{11} + b_{11}) + (1 - \mu_1)(w_{10} + b_{10})] \\
+ (1 - \mu)[\mu_0(w_{01} + b_{01}) + (1 - \mu_0)(w_{00} + b_{00})] \\
s.t. \ (4), \ (5), \ (6), \ (7) \text{ and } (8).
\]

Lemma 5. To implement effort plan \((1, 1)\) in the subjective scenario with a mediator as cheap as possible, the principal should pay wages according to

\[
(w_{00}, w_{10}, w_{01}, w_{11}) = \left(0, 0, \frac{c_2}{\mu}, \frac{c_1}{\mu \mu_1} + \frac{c_2}{\mu}\right) \tag{9}
\]
and transfers to the budget breaker according to

\[
(b_{00}, b_{10}, b_{01}, b_{11}) = \left(0, 0, \frac{c_1}{\mu \mu_1}, 0\right). \tag{10}
\]

Proof. First, consider the choice variables \( w_{00}, b_{11} \) and \( b_{00} \). In terms of the objective and the principal’s truth-telling constraints, it is without loss of generality to relabel \( b_{00} + w_{00} \) as \( b_{00} \), \( w_{11} + b_{11} \) as \( w_{11} \) and \( b_{01} + [(1 - \mu_0)/(\mu_0)]b_{00} \) as \( b_{01} \). On the other hand, the agent’s incentive to provide effort in each stage is clearly stronger the smaller \( w_{00} \) and the larger \( w_{11} \). Hence, \( w_{00} = b_{11} = b_{00} = 0 \) is without loss of generality.
Next, note that the principal could appropriate the entire surplus if her truth-telling constraint (7) was slack. I now show that this constraint binds. To this end, reformulate (7) as

\[ w_{10} + b_{10} \leq \mu_1[w_{01} - (w_{11} - w_{10}) + b_{01} + b_{10}] \]  

(7')

and the agent’s incentive constraint (4) as

\[ \mu[w_{10} - \mu_1(w_{01} - (w_{11} - w_{10}))] \geq c_1. \]  

(4')

By (4'), \( w_{10} > \mu_1[w_{01} - (w_{11} - w_{10})] \), so that the principal must distribute surplus to the budget breaker to satisfy (7'). Hence, this constraint binds. For the moment, ignore (8), the second truth-telling constraint. The principal should then specify the transfers \( b_{10} = 0 \) and \( b_{01} = w_{11} + [(1 - \mu_1)/\mu_1]w_{10} - w_{01} \). Moreover, it is without loss of generality to relabel \( w_{11} + [(1 - \mu_1)/\mu_1]w_{10} \) as \( w_{11} \), so that \( w_{10} = 0 \) cannot be suboptimal. In fact, these choices satisfy constraint (8).

Accounting for the previous findings, the remaining choices are obtained from

\[
\min_{w_{11} \geq 0, w_{01} \geq 0} C(w + b) = \mu w_{11} \\
\mu \mu_1 w_{11} - w_{01} \geq c_1 \\
\mu \mu_1 w_{11} + (1 - \mu) \mu_0 w_{01} \geq c_1 + c_2 \\
\mu \mu_1 w_{11} + (1 - \mu) \mu_0 w_{01} \geq c_2
\]  

(4’’), (5’), (6’)

This program is solved by \( w_{01} = c_2/\mu \) and \( w_{11} = c_1/(\mu \mu_1) + c_2/\mu \).

The payment scheme stated in the proposition prescribes that wages are only due if stage 2 yields a success. Thus, the agent is exposed to the risk that he might not be rewarded for a success in stage 1. This risk is not a necessary property of the cheapest contract under mediation.

---

23This can be achieved with \( w = (0, c_1/\mu, c_2/\mu, (c_1 + c_2)/\mu) \) and \( b = (0, 0, 0, 0) \).

24(9) seems to use the agent’s continuation payoffs after stage 1 as carrot and stick to induce hard work in stage 1: good performance in stage 2 is better rewarded if stage 1 was successful as well (see Lizzeri, Meyer and Persico (2002) and Ohlendorf and Schmitz (2011) for the use of carrot and stick in dynamic moral hazard problems). But in fact, the extra reward just compensates for the risk that success in stage 1 will not be rewarded if the agent fails in stage 2.

25Specifically, there is a cheapest contract under mediation that rewards success in stage 1 independently of the agent’s performance in stage 2. It consists of \( (w_{00}, w_{10}, w_{01}, w_{11}) = (0, c_1/\mu, c_2/\mu, (c_1 + c_2)/\mu) \) and (10).
However, note that the payment scheme satisfies the principal’s truthtelling constraints pointwise for each performance outcome in stage 2. The principal would consequently be willing to reveal her subjective evaluation also after stage 2.

I now propose a contract that implements effort plan \((1, 1)\) in the original setting without mediator at the same costs as the cheapest contract under mediated talk. Since the presence of a mediator does not preclude any mode of face-to-face communication between principal and agent, it is impossible to find a better contract. Let the payments still be specified by (9) and (10), the cheapest payments under mediated talk. Communication should proceed as follows. After stage 2, the principal reports her evaluation of stage 1 in form of a verifiable message \(m \in \{0, 1\}\) to the agent if stage 2 was successful. No further communication takes place. From the agent’s perspective, nothing has changed as compared with mediation provided that the principal is still willing to report truthfully. And that is indeed the case since any report results in the same payment.\(^{26}\)

The following notation for the budget breaker’s payoff will be useful in stating Lemma 6:

\[
B = \left(\frac{1}{\mu_1} - 1\right) c_2.
\]

**Lemma 6.** In the subjective scenario, the cheapest contract to implement effort plan \((1, 1)\) yields the principal a profit of \(S_1 + S_2 - B\). The agent’s payoff is zero, and the budget breaker receives \(B\).

Without payments to the budget breaker, the principal would always report bad performance to save on wages. To gain credibility, she must commit to pay the reward for good performance in stage 1 after each evaluation—to the agent if he was successful, and otherwise to the budget breaker. The cost to overcome the problem of hidden information is therefore due when the principal truthfully reports a negative evaluation. Now the advantage of the risky wage scheme (9) becomes apparent. It only requires the principal’s evaluation if the agent performs well in

\(^{26}\)Without assumption 2, face-to-face communication does in general not work as well as mediated talk. Possibly, \(e_2\) should then depend on \(x_1\), which requires interim communication. But the analysis of the objective scenario implies that the agent should still learn as little as possible about \(x_1\) before choosing \(e_2\). Randomized feedback might solve the trade-off optimally, and a mediator typically allows for more favorable randomizations.
stage 2. This is more likely to occur with high match quality, where stage 1 should have been successful as well.

Similar to the objective scenario, the principal must make a trade-off to find the optimal contract: while the inefficient plan (0,1) sacrifices surplus in stage 1 (worth $S_1$), the efficient plan (1,1) now requires payments to the budget breaker (worth $B$).

**Proposition 2.** In the subjective scenario, the optimal contract is as follows:

1. For $S_1 \leq B$, the inefficient effort plan (0,1) is implemented. The principal makes a profit of $S_2$, the agent gets zero, and the budget breaker is not involved.
2. For $S_1 > B$, the efficient effort plan (1,1) is implemented. The principal makes a profit of $S_1 + S_2 - B$. The agent’s payoff is zero, and the budget breaker receives $B$.

### 3.3 Comparison of the scenarios

The previous findings can be used to compare the two scenarios with respect to the generated surplus and the principal’s profit. By Propositions 1 and 2, the inefficient effort plan (0,1) yields in both scenarios the same profit. The efficient plan (1,1), on the other hand, generates an additional surplus of $S_1$ relative to (0,1). However, this plan entails incentive costs, namely $A$ in the objective scenario and $B$ in the subjective scenario. Proposition 3 follows immediately from these considerations.

**Proposition 3.** Define $m = \min\{A, B\}$ and $M = \max\{A, B\}$.

1. Suppose $A = B$. The scenarios are equivalent with respect to surplus and profit.
2. Suppose $A < B$. For $S_1 \leq m$, the scenarios are equivalent with respect to surplus and profit. For $m < S_1 \leq M$, both surplus and profit are strictly larger in the objective scenario. For $M < S_1$, the scenarios are equivalent with respect to surplus. Profit is strictly larger in the objective scenario.
3. Suppose $A > B$. Compared to $A < B$, the rankings of the scenarios are reversed.

In both scenarios, the principal can appropriate the entire surplus if she implements the inefficient effort plan (0,1). But each scenario has its own problem if the agent is to work according to
the efficient plan \((1, 1)\). When performance in stage 1 is objectively evaluated (objective scenario),
the outcome of stage 1 cannot be concealed from the agent. It conveys information about the unknown quality of the match, which just the agent can interpret correctly with certainty. Therefore, the agent receives private information before the project is completed, and it secures him a rent. With a subjective evaluation (subjective scenario), on the other hand, this problem does not arise: it is the principal who gets informed privately, and she may communicate with the agent only after his job is done. But in contrast to the objective scenario, the principal must involve the budget breaker to make the evaluation credible. Each problem can be more substantial, and each ranking of the scenarios can arise.

Uncertainty about match quality causes the problem with effort plan \((1, 1)\) in the objective scenario. More uncertainty, in form of more extreme posteriors after stage 1, increases the rent the agent can secure by pretending a pessimistic belief. Perhaps surprisingly, more uncertainty diminishes the problem that arises with this effort plan in the subjective scenario. Recall that to reduce the probability of a negative rating for stage 1, which triggers a transfer to the budget breaker, the principal only submits the rating after stage 2 and only when that stage was successful. Conditional on this event, a negative rating gets less likely if stage 1 gets more informative about match quality, so that the principal retains more surplus from the efficient effort plan. In summary, more uncertainty is detrimental in terms of profit and surplus in the objective scenario, but beneficial in the subjective scenario. I state this result formally using the concept of mean preserving spread (MPS) as a criterion for differences in uncertainty (cf. Rothschild and Stiglitz 1970). Given two distribution functions \(F_1\) and \(F_2\) of match quality, \(F_2\) is an MPS of \(F_1\) if and only if \(\int_0^1 h(\theta)dF_2 \geq \int_0^1 h(\theta)dF_1\) for every concave function \(h\) over \([0, 1]\).\(^{27}\)

**Proposition 4.** Suppose \(F\) is replaced by an MPS.

1. In both scenarios, the profit from the inefficient effort plan \((0, 1)\) does not change.
2. In the objective scenario, the profit from the efficient effort plan \((1, 1)\) decreases.
3. In the subjective scenario, the profit from the efficient effort plan \((1, 1)\) increases.

**Proof.** By the definition of MPS, the prior \(\mu\) does not change. The posteriors \(\mu_1\) and \(\mu_0\), on the

\(^{27}\)Alternatively, \(F_2\) is an MPS of \(F_1\) if and only if \(\int_0^1 \theta dF_1 = \int_0^1 \theta dF_2\) and \(\int_0^1 F_2(\theta) - F_1(\theta) d\theta \geq 0\).
other hand, are expectations of a convex and a concave function, respectively (cf. footnote 14). It follows that $\mu_1$ is greater and $\mu_0$ smaller than given $F$. Thus, the MPS leaves $S_2$ unchanged, raises $A$ and lowers $B$. In light of Propositions 1 and 2, this finding concludes the proof.

4 Conclusion

This paper offers an explanation as to why incentive pay often depends on subjective performance evaluations even though comprehensive objective appraisal systems could be established. My analysis builds on two central assumptions. First, an evaluation is considered as a means to generate information about the worker’s performance that would remain unknown otherwise. Second, the acquired information is either private and unverifiable or public and verifiable, depending on whether the employer rates subjectively or objectively. Verifiability is clearly an important concern since discretionary compensation entails a commitment problem for the employer. I argue that, nevertheless, subjective evaluations can be advantageous in terms of profit and surplus because they allow to withhold information from the worker. Indeed, withholding information seems particularly important under circumstances where the commitment problem is small. According to my theory, subjective evaluations are used at early stages of employment relationships in which characteristics like ability, productivity or profitability are highly unknown and can be inferred from performance.

References


