Interim Performance Evaluation in Contract Design

Bin R. Chen and Y. Stephen Chiu^{*†}

December 21, 2011

Abstract

We study a principal-agent problem with sequential efforts and limited liability. An interim performance evaluation (IPE) allows the principal to learn the degree to which the early effort is successful. We find conditions under which it is desirable to conduct such an IPE. A trade-off is identified between the beneficial effect when the agent's morale is boosted by a positive IPE outcome and the harmful effect when his morale is damaged by a negative one. We study both objective or subjective IPEs and characterize in each scenario the optimal contract and compare the corresponding effort plan with the first best effort plan.

Keywords: Interim Performance Evaluation, Subjective Evaluation, Objective Evaluation, Contract Design

JEL: D82, L14, M20.

1 Introduction

In most real life situations, the successful production of a good often requires sequential investments. Building a bridge, for example, first requires planning and a feasibility analysis at an early stage and then the construction of the bridge at a later stage. Likewise, the completion of a doctoral degree requires the student to take course work, to pass in a qualifying exam, to conduct research and to write a thesis of reasonable quality. In many such examples, interim performance evaluations (IPEs)

^{*}We would like to thank the coeditor David Myatt and two anonymous referees for very useful comments and suggestions. We also thank Kalyan Chatterjee, Stephen Ching, Jacques Crémer, Jin Li, Ping Lin, Wing Suen, Zhigang Tao, Frances Xu and seminar participants in the 2010 Econometric Society World Congress (Shanghai), CUHK, HKU, and Seoul National University for helpful comments. All remaining errors are our own.

[†]Chen: School of Business, Sun Yat-sen University, Guangzhou, CHINA. Email: xiechenbin@gmail.com. Chiu (corresponding author): School of Economics and Finance, University of Hong Kong, Pokfulam Road, HONG KONG. Email: schiu@econ.hku.hk.

are performed.¹ In the case of bridge building, the feasibility analysis may turn out to suggest that building the bridge is impractical. In the case of doctoral studies, the qualifying exam may lead to the student being dropped from the program. The merit of such IPEs is clear, as optimal continuation actions vary dependent on IPE outcomes. A less clear issue is whether they are still valuable when the principal intends to continue with the same course of action regardless of the IPE outcome. These issues are also related to the two main purposes of conducting IPE — providing performance feedback and designing compensation package — as argued by Cleveland et al (1989).²

Despite its popularity in practice, warnings against the use of IPEs can be found among both management researchers and economists. As Cook (1975) pointed out, most companies experimented with reward systems using informative feedbacks have found to be failures. In his recent best seller, Roberts (2004) writes that "[s]ubjective measures and milestones may provide more effective incentives for innovation than do the accounting numbers, but using them to provide very intense incentives is certainly problematic."

The question of the value of IPEs also appears in recent research in contract theory. Lizzeri, Meyer and Persico (2002), Fuchs (2007), and Manso (2011) obtains results that support the aforementioned concern. Specifically, they show that, whenever the desired continuation action does not depend on the IPE outcome, it is better not to reveal the IPE outcome to the reviewees. Not revealing this information means that there is no need to provide differential incentives depending on the IPE outcomes (to put it another way, fewer incentive compatibility constraints need to be satisfied).

In this paper, we set out to study the issue through a simple two-stage principal-agent problem. The principal solicits the agent's help to produce a final good, which may turn out to be a success or a failure. The probability of success depends on two non-observable sequential efforts made by the agent. The analysis focuses on situations where the two efforts exhibit complementarity—that is, a higher level of effort 1 makes the effort 2 more productive and vice versa. We also assume that, while both parties are risk neutral, the agent suffers from limited liability, a reasonable assumption in many applications. The benchmark model is a so-called traditional contract, where the payment made to the agent is dependent on the quality of the final good.

Given this benchmark, we study a contract in which payments to the agent depend not only on

¹By surveying 400 organizations worldwide, Aberdeen group (2010) reports that 91% of employers around the world set up feedback system and evaluate the employees' performance regularly. Outside the organizational setting, IPE is also known as opinion polls in political elections and information feedback scheme in sports competitions, etc.

 $^{^{2}}$ According to empirical study of Cleveland et al (1989), 69% of their survey respondents considered salary administration and 53% considered performance feedback to be among the three main purposes of performance appraisals. Similar findings are obtained in other empirical studies (Levine 1986, Rendero 1980).

the quality of the final product but also on the outcome of an IPE conducted at the end of the first stage. For the sake of simplicity, we assume that the IPE is conducted at a low cost, and that the outcome of the IPE informs the principal about the quality of the interim product that is correlated with the probability of success of the final good. A positive (negative) IPE outcome means a high (low) probability of success and, given our assumption of effort complementarity, the increase in success probability due to the second-stage effort is large (small) and this thus boosts (damages) the agent's morale.

Our main results are as follows. Firstly, we show that, in case the same course of continuation effort is intended regardless of IPE outcome, the IPE contract might be more desirable than traditional contract, only if the aforementioned beneficial effect dominates the harmful effect; otherwise, conducting an IPE will worsen the principal's payoff in this setting. This result is consistent with the observation by two project management experts. As they found, the risk of conducting IPEs is that the self-confidence of reviewees is likely to be damaged if the IPE outcome turns out to be unfavorable, and so the reviewees are thus discouraged from working hard. "[D]espair is even worse because the project is permeated with an attitude that says, 'Why try when we are destined to fail?'"(Meredith and Mantel 1995).

Secondly, even if IPE outcomes are subjective,³ we still find parameter values over which conducting IPEs is more profitable than not conducting them. However, because the principal may have an incentive to lie about the findings, IPEs are less likely to be desirable. As a result, concealing feedbacks or revealing limited information in IPEs may become attractive. This result resonates well with "targeted inaccuracy" in rating, as documented by organizational behavior literature.⁴ Upon interviewing Navy officers, Bjerke et al. (1987) found that managers withhold evaluation information to secure esteem-building promotions for junior officers and do not wish to depress their morale.

Thirdly, we also characterize IPE contracts where the continuation efforts depend on the IPE outcomes. Naturally this makes that an IPE is more likely to be useful. While this is not surprising, it is still interesting to compare the difference between an optimal IPE contract where continuation action plan varies dependent on the IPE outcome with one where the continuation action plan remains the same.

 $^{{}^{3}}$ By a subjective IPE, we mean that the IPE outcome is not observable by the agent and the principal can lie about it and hence additional constraints are needed to ensure that she is telling the truth.

 $^{^{4}}$ As Murphy and Cleveland (1985) point out, "this does not mean that inaccuracy in rating is a good, but targeted inaccuracy might be a very good thing." In practice, many organizations go through periodic evaluation, but the supervisor (or manager) seldom reports the results or just provides the feedback of limited information (see Lazear and Gibbs 2009).

Finally, we find the notions such as morale and confidence, which are normally outside standard economic discourse, to be helpful in interpreting our results. With lower morale or self-confidence, the agent needs to be motivated by a greater-powered incentive. Bénabou and Tirole (2003) were the first to formulate self-confidence and to study the harmful effect of high-powered incentives. As the principal-agent problem we study is more standard, this paper echoes their call to use economic modeling to analyze the problems that psychologists are interested in.

The remainder of this paper is organized as follows. We will discuss related literature at the end of this section. Section 2 lays out the basic model. Section 3 analyzes objective IPE contracts and evaluates their value; an objective IPE is an IPE that is verifiable and so this is naturally the first step of our analysis. Sections 4 analyzes subjective IPE contracts in the same way and evaluates their value. Section 5 discusses and deals with some extensions. In Section 6 we give some concluding remarks.

Related Literature This paper is related to a nascent and growing literature on IPE and feedbacks. A couple of papers are concerned about sequential effort choices in a two-stage principalagent framework. Ray (2007a) finds that IPE enhances efficiency by providing the option of ending a project with early low returns. His results rely on the assumptions that production is indivisible and efforts across stages are perfect substitutes. Lizzeri, Meyer, and Persico (2002) examine whether the principal should tell the agent about an IPE outcome. They show that, given that the principal's intended continuation plan is independent of the IPE outcome, it is better not to give feedback. A similar result is also obtained by Manso (2011).

Notice that in these two papers, a key assumption is that a contract made contingent on the *IPE outcome is enforceable whether or not the feedback is revealed*. This contractibility assumption is supported by fact that, in both papers, output is produced in each stage and the IPE outcome in their context is nothing but the first stage output level. In contrast, we assume that the IPE outcome, if it is ever used in contracting, should be revealed to the agent in the interim stage. This assumption is supported by the fact that in our model the first stage product is only transient (the interim product may just be an input for the second stage production).

Yildirim (2005), Aoyagi (2008), Goltsman and Mukherjee (2009), Ederer (2010) characterize the optimal strategy of interim information disclosure in the context of a two-stage tournament. The results are somewhat mixed regarding whether the principal should disclose interim performance.

Like Lizzeri et al, it is assumed in their analysis that the principal has more information than the agent about the first-stage performance and that this performance measure is verifiable by the court and hence can be used in a contract.

The feature that the principal has more information than the agent does about his performance also appears in the recent literature on subjective evaluation in relational contracts (see pioneering work by Levin 2003 and MacLeod 2003). In such a setting, both the subjective evaluation and the output level are independently generated in each stage. Nonetheless, Fuchs (2007) shows a result similar to that of Lizzeri, Meyer, and Persico (2002) in which that interim feedbacks are not desirable given the same continuation action plan is intended.

2 The model

An agent (he) is hired by a principal (she) to complete a two-stage project. Both parties are risk neutral and the agent has no wealth and is subject to limited liability. There is no discounting between stages. In Stage 1, the agent chooses an unobservable effort $e_1 \in \{0, 1\}$ with a cost c_1e_1 , where $c_1 > 0$. At the end of this stage, an interim product of quality $x_1 \in \{0, 1\}$ is generated and the quality is high (i.e., $x_1 = 1$) with probability $r_0 + r_1e_1$ or low (i.e., $x_1 = 0$) with the remaining probability, where $r_0 \in (0, 1)$ and $r_1 \in (0, 1 - r_0)$. In Stage 2, the agent chooses another unobservable effort $e_2 \in \{0, 1\}$ with a cost c_2e_2 , where $c_2 > 0$. At the end of this stage, a final product is generated with quality $x_2 \in \{0, 1\}$.

If the interim product is of high quality, the final product is a "good" (i.e., $x_2 = 1$) with probability $t_0 + t_1e_2$ or a "bad" (i.e., $x_2 = 0$) with the remaining probability, where $t_0 \in (0, 1)$ and $t_1 \in (0, 1 - t_0)$. If the interim product is of low quality, however, the final product is a "good" with probability $t'_0 + t'_1e_2$ or a "bad" with the remaining probability, where $t'_0 \in (0, 1)$ and $t'_1 \in (0, 1 - t'_0)$. We assume that $t_0 > t'_0$ to accord with the definition of a high-quality interim product and that $t_1 > t'_1$ so that there is complementarity between efforts across stages.⁵ The nature of the final product is observable and verifiable, and the principal gains from the project a value of B > 0 in the case of a final "good" and a value of zero in the case of a final "bad."

Although not immediately observable, x_1 can be learned through an IPE. Specifically, subsequent

⁵If $t_1 < t'_1$, the efforts between the two stages are substitutes. If $t_1 = t'_1$, the two-stage efforts are independent. If $t_1 > t'_1$, the efforts between the two stages are complements. In our model, the complementarity may result from the fact that satisfactory completion of previous phases lays the foundation for proceeding to the next phase. Manso (2011) also argues that complementarity naturally arises when the principal tries to design a contract for "exploration."



Figure 1: The timeline of the game with the option of an IPE.

to its realization but prior to Stage 2, the principal has an option of arranging an evaluation through which she obtains a signal $\sigma \in \{H, L\}$ that *perfectly* reveals x_1 . Projects fitting our model can be found in many industries such as construction and software development. The potential safety problems of bridges under construction or possible bugs in an unfinished software package are not easily discovered unless the project manager undertakes interim checkups. We assume that the evaluation contains no noise. Later on in this paper (Section 5.1), we show that a noisy IPE does not affect the qualitative properties regarding the desirability of the IPE. We also assume that the cost of the evaluation is negligible; any positive cost will simply reduce the net gains of carrying out the IPE by the same amount.

The timing of this game is summarized in Figure 1. At the outset, the principal determines whether to conduct an IPE and proposes and signs a take-it-or-leave-it contract with the agent, whose reservation payoff is zero. The production process goes through Stage 1 and Stage 2 sequentially. When an IPE is not scheduled, the allowable contract will only be contingent on the verifiable nature of the final product, x_2 . So it is represented by a duple (w, b), where w is a wage rate that must be paid out and b is a bonus paid out if and only if $x_2 = 1$. The agent's limited liability dictates that⁶

$$w \ge 0 \text{ and } b + w \ge 0. \tag{1}$$

We call such a contract a "traditional contract." We use $\langle e_1; e_2 \rangle$ to denote the corresponding action plan implemented, where e_1 and e_2 are the agent's efforts in Stage 1 and Stage 2. In the absence of an IPE, our model is just a variant of the hidden-action model with limited liability, which is a building block of many recent papers on the agency problem (e.g., Crémer 1995, Che and Yoo 2001, Schmitz 2005, etc.). In the context of procurement and project management, the assumption of limited liability is quite reasonable because the agents are usually protected by bankruptcy laws (see, for instance, Calveras, Ganuza and Hauk 2004).

When an IPE is scheduled, the choice of contract is enriched because it can be written contingent not only on x_2 but also on the interim feedback m, disclosed by the principal. There are two formulations of the IPE: it may be objective so that the IPE outcome, σ , is publicly observable and verifiable and as a result $m = \sigma$; it may be subjective so that σ is known only by the principal. For both types of IPEs, without loss of generality, we assume that a set of all possible messages that the principal may utter equals $\{H, L\}$. Thus, the contract is represented by a quadruple (w_H, w_L, b_H, b_L) , where, given feedback m, w_m is the wage rate paid out to the agent unconditional on x_2 and b_m is the additional bonus paid out if $x_2 = 1$. The limited liability constraints dictate that

$$w_H, w_L \ge 0; \quad b_H + w_H \ge 0; \text{ and } b_L + w_L \ge 0.$$
 (2)

We call such a contract an "IPE contract." Note that the Stage-2 effort would be dependent on the feedback. We denote the corresponding action plan by $\langle e_1; e_2(H), e_2(L) \rangle$. To facilitate exposition, we make the following definition.

Definition 1 An IPE contract is called an effort-sorting scheme if it implements the action plan

⁶We assume that w is paid out along with b (if rewarded) at the very end. In this case, b may be negative as long as w + b is not negative. An alternative assumption is that w is paid out earlier, at the end of Stage 1. This is a more restrictive case because now b is not allowed to be negative due to limited liability. But the two assumptions turn out to be the same because under our assumption b is never negative.

 $\langle 1; 1, 0 \rangle$.

Given contract ϕ , we use V^{ϕ} to denote the expected revenue accruing to the principal, C^{ϕ} to denote her expected cost, and π^{ϕ} to denote her expected payoff; so we have $\pi^{\phi} \equiv V^{\phi} - C^{\phi}$.

Although our characterization of the IPE contract imposes little restriction on the parameter space, the contrast and comparison of it with the traditional contract becomes more focused with such a restriction. For this reason, we assume that (i) the optimal traditional contract implements action plan $\langle 1; 1 \rangle$,⁷ and (ii) among all of the incentive compatible (IC) constraints imposed on the agent by the contract, the one that prevents deviation from the action plan $\langle 1; 0 \rangle$ is binding. This can be expressed:⁸

A1
$$\frac{c_1}{c_2} \leq \frac{r_1(t_0 - t'_0)}{(r_0 + r_1)t_1 + (1 - r_0 - r_1)t'_1}$$
 and $B \geq \widehat{B}$.

 \widehat{B} is the minimum *B* such that the optimal traditional contract still implements $\langle 1; 1 \rangle$, instead of other action plans. See Appendix B for more details.

Given this assumption, the optimal traditional contract, denoted by T, is characterized by

$$w^{T} = 0 \text{ and } b^{T} = \frac{c_{2}}{(r_{0} + r_{1})t_{1} + (1 - r_{0} - r_{1})t'_{1}},$$
(3)

with corresponding implementation cost

$$C^{T} = (r_{0} + r_{1})(t_{0} + t_{1})b^{T} + (1 - r_{0} - r_{1})(t_{0}' + t_{1}')b^{T}.$$
(4)

3 Objective IPE contracts

Consider the scenarios in which the IPE is objective. In this case, the message conveyed by the principal will be the same as the IPE outcome, and there are no incentive problems whether the principal will lie or not. Before discussing any specific contract, we present the following property. (The proofs of lemmas and propositions are relegated to the Appendix unless otherwise stated.)

Lemma 1 Suppose that the IPE is objective. For any given action plan, the optimal contract is characterized, without loss of generality, by $w_L = w_H = 0$.

 $^{^{7}}$ Under the optimal traditional contract, the principal's payoff achieves the highest among all traditional contracts (not necessarily restricted to the same action plan).

 $^{^{8}}$ It is common to make such an assumption and to focus on the most interesting case. Our A1 serves the same purpose as Assumption 1 in Schmitz (2005) and Assumption 1 in Manso (2011).

It is easy to understand why $w_L = 0$ because an increase of this value will weaken the agent's incentive to exert the first-stage effort and is undesirable. Regarding w_H and b_H , both are equally effective at relaxing the first-stage IC constraint in our scenario. When the IPE outcome is a noisy signal about the quality of the interim product, using b_H is strictly more cost effective than using w_H . To see this, consider the case where the IPE outcome is completely uninformative so that an increase in w_H will have no effect in motivating the agent at the first stage, while an increase in b_H will do. (The reason is that now the Stage-1 effort will not change the realization of the IPE outcome at all but will affect the quality of final product.) Therefore, focusing on the case where $w_H = 0$ is not only without loss of generality but also robust to the introduction of noise to the IPE outcome.

In what follows, we turn to the optimal IPE contracts that implement $\langle 1; 1, 1 \rangle$ and $\langle 1; 1, 0 \rangle$, denoted by O1 and O0, respectively.

3.1 Implementing $\langle 1; 1, 1 \rangle$

We now characterize O1. Consider the agent's second-stage decision. Given signal H (hence the same message uttered by the principal), the agent prefers choosing $e_2 = 1$ to choosing $e_2 = 0$, if and only if

$$b_H \ge \frac{c_2}{t_1},\tag{5}$$

where t_1b_H is the extra benefit of exerting Stage-2 effort and c_2 is the cost of the effort. This condition is in fact a very familiar equation in the moral hazard problem with limited liability constraints. In a similar way, given signal L, the agent prefers choosing $e_2 = 1$ to choosing $e_2 = 0$, if and only if

$$b_L \ge \frac{c_2}{t_1'}.\tag{6}$$

Notice that, because $t_1 > t'_1$, the minimum bonus that motivates the agent to work hard in Stage 2 is greater when signal L is received. The intuition is that now that the agent has low morale, or is less confident about the project, he has to be given greater incentives.

The last IC constraint to check is whether the agent has any deviation motive at Stage 1.⁹ At the outset, foreseeing that he will for certain choose $e_2 = 1$, the agent prefers choosing $e_1 = 1$ to

 $^{^{9}}$ It can be shown that any other IC constraints are satisfied so long as these three IC constraints, (5), (6), and (7), are satisfied.

choosing $e_1 = 0$ if and only if

$$r_1(t_0 + t_1)b_H - r_1(t'_0 + t'_1)b_L + r_1(w_H - w_L) \ge c_1.$$
(7)

Notice that an increase in b_L makes this condition more difficult to maintain. Taken into account $w_H = w_L = 0$, and b_L satisfying (6) with equality, the condition can be rewritten as

$$b_H \ge \frac{c_1}{r_1 \left(t_0 + t_1 \right)} + \frac{\left(t_0' + t_1' \right) c_2}{\left(t_0 + t_1 \right) t_1'}.$$
(8)

In the case where this minimum b_H exceeds the one calculated in (5), the purpose of b_H is to motivate the Stage-1 effort, rather than to motivate the Stage-2 effort under signal H. It should be noted that assumption **A1** itself does not preclude the possibility of the right hand side (RHS) of (8) exceeding the RHS of (5). This suggests that there is a potential cost from using IPE because now additional constraints need to be satisfied. The following proposition summarizes the characterization of this IPE contract.¹⁰

Proposition 1 Suppose that the IPE is objective. O1 satisfies $w_H = w_L = 0$, $b_L = c_2/t'_1$ and

$$b_H = \max\left\{\frac{c_2}{t_1}, \frac{c_1}{r_1(t_0+t_1)} + \frac{(t'_0+t'_1)c_2}{(t_0+t_1)t'_1}\right\}.$$
(9)

Let us consider whether O1 is more profitable than traditional contracting. To this end, we assume **A1** so that the optimal traditional contract is to implement action plan $\langle 1; 1 \rangle$ and the contract T is characterized in (3). As both contracts involve implementing the same action plan, yielding the same expected revenue to the principal, O1 is more profitable than T if and only if C^{O1} is lower than C^{T} . For the moment, we assume that under O1 the first-stage IC constraint is nonbinding and hence $b_{H} = c_{2}/t_{1}$. In this case,

$$C^{O1} = (r_0 + r_1) \frac{t_0 + t_1}{t_1} c_2 + (1 - r_0 - r_1) \frac{t'_0 + t'_1}{t'_1} c_2.$$
(10)

 is only c_2 . Thus the extra cost of $(t_0/t_1) c_2$ is required to motivate the agent and this is known as the agent's limited liability rent conditional on signal H. In a similar way, the term $(t'_0/t'_1) c_2$ is the agent's limited liability rent conditional on signal L. It turns out that one factor that determines whether C^{O1} is lower than C^T is the comparison between these two rents. The following proposition formally states the result.

Proposition 2 Assume A1.

- 1. Suppose $\frac{t'_0}{t'_1} > \frac{t_0}{t_1}$. Then the implementation cost of O1 is strictly greater than that of T.
- 2. Suppose $\frac{t'_0}{t'_1} \leq \frac{t_0}{t_1}$. Then there exists $R^* > 0$ such that for all $c_1 < R^*c_2$, the implementation cost of O1 is lower than that of T.

Notice that t'_0/t'_1 (or t_0/t_1), which features as the coefficient in the limited liability rent discussed above, is the ratio of the default probability of success over the additional probability of success of Stage-2's effort given the interim product of low quality (or of high quality).¹¹ Result 1 states that if this ratio is higher under $x_1 = 0$ than under $x_1 = 1$, it is not worthwhile conducting an IPE. Intuitively, by comparing to traditional contracting, the IPE contract, if implementing the same action plan, can mitigate the agency problem after revealing H, but exacerbate the agency problem after revealing L. Given that the latter harmful effect dominates the former beneficial effect, revelation of feedback is less desirable in this case.¹² The findings here vividly support the observation by Meredith and Mantel (1995): when the project would be destined to fail upon unfavorable Stage-1 outcome (e.g., a very small t'_1), committing to not providing any feedback is optimal.

Result 2 states that, if the ratio is smaller under signal L than under signal H, it may be worthwhile conducting an IPE. The underlying reason is symmetrical to that of Result 1. Moreover, the restriction to $c_1 < R^*c_2$ is to make sure that b_H is indeed equal to c_2/t_1 and that the IC constraint that prevents deviation to $\langle 0; 1, 1 \rangle$ is nonbinding. In this case, conducting an IPE is beneficial. In case the IC constraint (7) is binding, however, the b_H exceeds c_2/t_1 and the cost C^{O1} calculated

¹¹Since we assume perfect evaluation, the quality of interim product x_1 is perfectly consistent with signal σ . But notice that all these probabilities are defined conditional on x_1 instead of σ . For more information on limited liability rent, see Laffont and Martimont (2003) and Schmitz (2005).

 $^{^{12}}$ The trade-off between such two effects is traced down to the comparison between the limited liability rents generated under the two different Stage-1 outcomes. The underlying reason is that, no matter whether the contract is T or O1, their implementation costs are weighted average of the aforementioned two rents, but they differ in the relative weights. Details are in the proof of Proposition 2.

in above by assuming $b_H = c_2/t_1$ is an underestimation of the true cost. A beneficial IPE is not guaranteed.

To summarize, there are two reasons why C^{O1} may exceed C^T . Firstly, the increase in agency rent under signal L is too high for the reduction in agency rent under signal H to fully compensate it. Secondly, the reduction in bonus under signal H may not be as large as initially expected because the Stage-1 IC constraint under O1 is more likely to be binding than under T. Under O1, a high b_L which is to ensure $e_2 = 1$ upon signal L undermines the agent's incentive to work hard at the first stage and this feature does not appear under T.

3.2 Implementing $\langle 1; 1, 0 \rangle$

We now characterize O0. Like O1, O0 still prescribes that $w_H = w_L = 0$. b_L should be made as low as zero because, given action plan $\langle 1; 1, 0 \rangle$, there is no need to motivate the agent to work hard given $\sigma = L$. b_H should still satisfy (5) so that the agent will not deviate to $e_2 = 0$ given $\sigma = H$. The last thing to check is whether at the outset the agent has any incentive to deviate to action plan $\langle 0; 1, 0 \rangle$. This condition, after plugging $w_H = w_L = 0$ (from Lemma 1) and $b_L = 0$ (obviously the case), is thus equivalent to

$$b_H \ge \frac{c_1 + r_1 c_2}{r_1 \left(t_0 + t_1 \right)} \tag{11}$$

Thus, b_H should be the maximum of this term and c_2/t_1 . As a result, O0 is characterized as follows:

Proposition 3 Suppose that the IPE is objective. O0 is characterized by $w_H = w_L = b_L = 0$ and

$$b_{H} = \max\left\{\frac{c_{2}}{t_{1}}, \frac{c_{1} + r_{1}c_{2}}{r_{1}\left(t_{0} + t_{1}\right)}\right\}$$

Notice that b_H is smaller under O0 than under O1 (more specifically, the second term inside the maximum operator is smaller under O0 than its counterpart under O1). In other words, the incentive to shirk during the first stage is lower under O0 than under O1. The reason is as follows. Under O1, the agent, even after shirking during Stage 1, will still be induced to work hard under signal L, due to an attractive bonus $b_L = c_2/t'_1 > c_2/t_1$. Under O0, no such incentive is available under signal L. As a result, it becomes more attractive for the agent to shirk under O1 than under O0. Next we compare O0 with traditional contracting in terms of profitability. For this purpose, we focus on a scenario where **A1** holds so that T is indeed the optimal traditional contract. Notice that in switching from T to O0, there is a reduction in the expected benefit, $(1 - r_0 - r_1)t'_1B$, and also a reduction of the expected cost, $C^T - (r_0 + r_1)(t_0 + t_1)b_H$. Since the latter is independent of B, there exists a threshold, denoted by B^O , such that the former is smaller than the latter. Hence, the following proposition is established.

Proposition 4 Assume A1. There exists a cutoff B^O such that O0 yields a higher (lower) level of profit to the principal than T does if $B < B^O$ ($B > B^O$). Moreover, this cutoff B^O is increasing in c_2, t'_0 , and t_0 , and decreasing in t'_1 .

The comparative statics state that O0 is more desirable, if the ratio of t'_0/t'_1 is higher. A larger ratio of t'_0/t'_1 is associated with greater limited liability rents, when inducing continuation efforts upon an interim product of low quality. The traditional contract T does not rely on an IPE, but indeed induces $e_2 = 1$ upon $x_1 = 0$. Hence, a rise in the ratio of t'_0/t'_1 increases the average rent paid to the agent. The effort-sorting scheme, however, by ceasing to induce any effort after revealing Ldoes not suffer from this problem. Thus, a rise in the ratio of t'_0/t'_1 favors the use of effort sorting.¹³ Likewise, a rise in t_0 incurs greater rents when inducing $e_2 = 1$ upon $x_1 = 1$. Although C^{O0} and C^T both increase in t_0 , the latter increases more rapidly in t_0 . The reason is that when working upon an interim product of high quality, the bonus, as well as limited liability rent, is reduced if positive feedback is revealed. (The comparative static with respect to t_1 is more complicated because it also determines the degree of complementarity between efforts across stages.)

Some discussion with the literature is in order. According to Ray (2007a) and Manso (2011), providing (objective) feedback is beneficial for the principal, if it helps her to screen out bad projects or provide proper incentives for "exploration." Such advantages of IPEs can be translated into the implementation of an "effort-sorting" schemes in our context: the conduct of IPE can make the Stage-2 effort contingent on the Stage-1 outcome. Moreover, we find that the effort-sorting scheme payoff dominates the traditional contract only under certain conditions, which are favored by a smaller *B* or a larger t'_0/t'_1 .

 $^{^{13}}$ A smaller t'_1 not only enlarges saving on the limited liability rent but also reduces the expected revenue loss of effort sorting.



Figure 2: The value of objective IPEs

3.3 The value of IPE contracts

Thus far, we have characterized O1 and O0 and compared each with respect to T. What remains to be done is to find out the optimal contract when an IPE option is available but need not be adopted in the contract. A second question concerns how this optimal contract performs with respect to the first-best outcome.

For both questions, we can put down our analysis using a graphical method (refer to Figure 2). The horizontal axis denotes t'_1 , where $t'_1 \in (0, t_1)$ so that efforts are complementary across stages. The vertical axis denotes B and we focus on the range where **A1** is satisfied and assume other parameters (such as c_1 and c_2) are unchanged in the exercise. The whole space in the figure thus represents the parameter range over which T is the optimal traditional contract. **Optimal contract** Given A1, we can verify that the optimal IPE contract is either O1 or $O0.^{14}$ Notice that between these two contracts, the principal will choose the former over the latter if its extra benefit more than offsets its extra cost, which is equivalent to

$$B > \widetilde{B} \equiv \frac{t_0' + t_1'}{(t_1')^2} c_2,$$

where \tilde{B} is decreasing in t'_1 , intersecting B^O from above when $t'_1 = (t'_0/t_0)t_1$.¹⁵ Thus, with the help of our earlier propositions, the space in panel a of Figure 2 can be partitioned into four regions. If $t'_1 < (t'_0/t_0)t_1$, O1 is dominated by T, so the principal will choose from either O0 or T. In the bottom left region, the optimal contract is O0, while in the top left region, the optimal contract is T; the threshold of B is B^O . On the other hand, if $t'_1 \ge (t'_0/t_0)t_1$, T is dominated by O1, so only O0 and O1 will be considered. In the top right region, the optimal contract is O1, while in the bottom right region, the optimal contract is O0; the threshold of B is \tilde{B} now. This thus completely characterizes the optimal contract when an IPE is available and need not be chosen. The general insight is that when B is sufficiently large, it is profitable to implement $e_2 = 1$ even when the IPE reveals less favorable feedback L. Moreover, the corresponding threshold of B is decreasing in t'_1 , because a rise in t'_1 will enlarge the extra benefit of investing $e_2(L) = 1$.

Comparison with the first best Now we turn to comparison with the first best, which is the solution to the social planner's welfare maximization problem when he observes the IPE signal. Note that given A1, the first best dictates that either $\langle 1; 1, 1 \rangle$ or $\langle 1; 1, 0 \rangle$ will be implemented. The former is chosen over the latter if and only if its extra benefit more than offsets the extra cost of effort. The condition is equivalent to

$$B > B^* \equiv \frac{c_2}{t_1'},$$

¹⁴Other action plans are not optimal. The reasons are as follows. Firstly, inducing $\langle 1; 0, 1 \rangle$ is less desirable than inducing $\langle 1; 1, 0 \rangle$. The latter allows the principal to not only reap a greater expected revenue but also curtail a greater expected cost because the agent is now motivated by an optimistic belief to choose a high Stage-2 effort. Secondly, inducing $\langle 1; 0, 0 \rangle$ is less desirable than inducing $\langle 1; 1, 1 \rangle$, simply because traditional contract implementing $\langle 1; 0 \rangle$ is dominated by that implementing $\langle 1; 1 \rangle$ given **A1**. Thirdly, inducing a low effort in Stage 1, e.g., the action plan $\langle 0; y, z \rangle$, is never optimal; this is further guaranteed by **A1**: when c_1/c_2 is sufficiently low and the project is sufficiently valuable, the principal will not benefit from a low Stage-1 effort.

¹⁵Here \tilde{B} is defined and calculated by assuming that only Stage-2 IC constraints are binding under O1 and O0 or that c_1/c_2 is small enough. If either of these Stage-1 IC constraints are binding, the above \tilde{B} is an underestimation of threshold, but all the features characterized in Figure 2 is without loss of generality. See the Appendix for more detail.

which is a downward slopping curve in panel b of Figure 2. (In the Appendix, we show that B^* is smaller than the thresholds B^O and \tilde{B} .) As a result, the space can be divided into three regions. In the top and bottom regions, the principal's choice is as efficient as the first best. In the middle region, the principal implements $\langle 1; 1, 0 \rangle$ while the first best calls for $\langle 1; 1, 1 \rangle$, meaning there is an under-investment problem. Intuitively, agency problem increases the marginal cost of inducing $e_2(L) = 1$, ensuring that the Stage-2 effort is under-invested especially subsequent to a bad IPE outcome.

Notice that in the top region, the availability of an IPE has no efficiency implication. In the middle region, it actually reduces efficiency compared with traditional contracting. In the bottom band, the principal will choose O0, which implements the first best so the availability of an IPE improves efficiency. Thus, the efficiency implications with regard to conducting an IPE are mixed.

4 Subjective IPE contracts

In this section, we study the scenario in which the IPE is subjective. The principal is now tempted to lie or to hide the signal from the agent and hence the contracting problem is complicated by this adverse selection problem. We focus on the optimal IPE contracts that implement $\langle 1; 1, 1 \rangle$ and $\langle 1; 1, 0 \rangle$. Denoting the two contracts by S1 and S0, respectively. We focus on IPE contracts with the following full-revealing property.

Definition 2 An IPE contract is said to satisfy the full-revealing property, if message m, announced by the principal, equals the true signal σ learnt in the subjective IPE.

In other words, the new contracting problem is now constrained by additional "truth-telling" conditions.

4.1 Implementing $\langle 1; 1, 1 \rangle$

To implement $\langle 1; 1, 1 \rangle$, the cost-minimization problem confronting the principal is the same as the case of the objective IPE, except that the following two truth-telling constraints are added.

$$(t_0 + t_1) b_H + w_H \leq (t_0 + t_1) b_L + w_L; \tag{12}$$

$$(t'_0 + t'_1) b_L + w_L \leq (t'_0 + t'_1) b_H + w_H.$$
(13)

(12) states that, given the agent's belief in the principal, the principal would not benefit from lying and claiming to have received signal L when she had actually received H. The LHS (left hand side) is the principal's expected payment to the agent when she tells the truth and the RHS is its counterpart when she lies. (13) is the corresponding constraint when the principal had actually received L. We can verify that, given contract O1, (13) is violated and (12) is not. In other words, the principal will have an incentive to lie if signal L is received but no incentive to lie when signal His received. Interestingly, such upward cheating motive of the principal is supported by observations in practice.¹⁶

It is easy to verify that the optimal contract still entails $w_L = 0$ and $b_L = c_2/t'_1$; and that an increase in w_L or b_L will not only increase the expected cost but will also worsen the truth-telling constraint (13). What may be different are the choices of w_H and b_H . We can represent their choices in a (b_H, w_H) diagram (see Figure 3). (12) is satisfied if and only if (b_H, w_H) is below or on the H line. (13) is satisfied if and only if (b_H, w_H) is above or on the L line. Notice that both lines have a horizontal intercept of c_2/t'_1 . An additional constraint, IC_H , is to ensure $e_2 = 1$ upon the agent receiving a high message. This constraint is simply $b_H \ge c_2/t_1$ if we assume non-bindingness of the agent's first-stage IC constraint under the O1 problem. The shaded region indicates where all three constraints are satisfied, and one can verify that the optimal (b_H, w_H) is the dotted point within the shaded region.

Proposition 5 Suppose c_1/c_2 is sufficiently low. Then action plan $\langle 1; 1, 1 \rangle$ is implementable and S1 satisfies $w_L = 0$, $b_L = \frac{c_2}{t_1'}$, $b_H = \frac{c_2}{t_1}$, and $w_H = (t'_0 + t'_1) \left(\frac{c_2}{t'_1} - \frac{c_2}{t_1}\right)$.

Proof. Omitted.

The condition on c_1/c_2 is to ensure that we can ignore the IC constraint that prevents deviation to action plan (0; 1, 1) so that the only relevant constraints are (5), (6), (12), and (13). This assumption not only simplifies the characterization of S1, but also ensures that (1; 1, 1) is indeed implementable. When c_1/c_2 is not sufficiently small, b_H may be made larger to ensure no deviation to action plan (0; 1, 1). This means the IC_H line will move rightwards and hence the existence of the shaded region in Figure 3 is no longer guaranteed.

 $^{^{16}}$ According to Murphy and Cleveland (1991), the ratings that supervisors report to workers are significantly higher and more skewed than the ratings they report to independent researchers. In a ten-year study of a thousand-member social service department, Milkovich, Newman, and Milkovich (2007) reports that only three of the possible ten thousand ratings were "below average."



Figure 3: Solving the subjective IPE contract S1

The characterization of S1 has interesting properties. The contract is the same as O1, except that w_H is made large enough to ensure that the principal has no incentive to lie when receiving signal L. In principle, her honesty can also be ensured with an increase in b_H (or a simultaneous increase of both b_H and w_H) but this is not the most cost-effective. The L line has a slope of $t'_0 + t'_1$ in absolute terms. That is, the principal will continue to have no incentive to lie under signal L when w_H is increased by one and b_H is reduced by $1/(t'_0 + t'_1)$. By doing so, the principal increases the expected cost by one (because of the increase in w_H) and reduces the expected cost by $(t_0 + t_1)/(t'_0 + t'_1) > 1$ (because of the decrease in b_H). Hence there is a net saving of expected cost. This explains why using w_H is more cost effective than using b_H to deter cheating.

Now that IPE is subjective, it is less likely that $\langle 1; 1, 1 \rangle$ is implemented at a lower cost under T. Despite its drawback, beneficial S1 is not impossible. The following proposition summaries this result.

Proposition 6 Suppose c_1/c_2 is sufficiently low. There exists parameters under which S1 exists and its implementation cost is lower than that under T.

Relative to the case of O1, the desirability of S1 over T is guaranteed under more restrictive conditions.¹⁷ It is worth noting that, our analysis of S1 is restricted to the IPE contract satisfying the full-revealing property. Of course the principal could choose an IPE contract such that $w_H =$ $w_L = 0, b_H = b_L = b^T$, and she pools or randomizes her messages. In this case the outcome simply replicates that of T.

4.2 Implementing $\langle 1; 1, 0 \rangle$

Similarly, to implement $\langle 1; 1, 0 \rangle$, the cost-minimization problem confronting the principal is the same as in the case of objective IPE, except that the following two truth-telling constraints are added.

$$(t_0 + t_1)(B - b_H) - w_H \ge t_0(B - b_L) - w_L; \tag{14}$$

$$t'_{0}(B - b_{L}) - w_{L} \ge (t'_{0} + t'_{1})(B - b_{H}) - w_{H}.$$
(15)

 $^{1^{7}}$ According to the proof of Proposition 6, S1 payoff dominates T if the following conditions hold: (1) the ratio of c_1/c_2 is sufficiently small; (2) t'_1 is sufficiently large (with the threshold greater than $(t'_0/t_0)t_1$ and close to t_1); (3) $(1 - r_0 - r_1)(t_0 - t'_0) > t_1 + t'_0$.

These two constraints are the counterparts of (12) and (13) for the problem of S1, ensuring the honesty of the principal when she has actually received H and L, respectively. Likewise, if contract O0 is to be used and the principal is trusted, she may have an incentive to lie when she receives signal L. As a result, a contract that differs from O0 is called upon in order to implement $\langle 1; 1, 0 \rangle$.

The analysis of S0 is similar to the case of S1. With the help of a graphical method, we obtain the following proposition.

Proposition 7 Consider the scenario where a subjective IPE is conducted. Assume that c_1/c_2 is sufficiently low.

- 1. Suppose $\frac{t'_0}{t'_1} < \frac{t_0}{t_1}$. The action plan $\langle 1; 1, 0 \rangle$ is implementable if and only if $B \ge \frac{\left(t_0 + t_1 t'_0 t'_1\right)}{\left(t_1 t'_1\right)} \frac{c_2}{t_1}$.
- 2. Suppose $\frac{t'_0}{t'_1} > \frac{t_0}{t_1}$. The action plan $\langle 1; 1, 0 \rangle$ is implementable if and only if $B \ge \frac{(t_0+t_1)}{t_1} \frac{c_2}{t_1}$.
- 3. Suppose $\langle 1; 1, 0 \rangle$ is implementable. The optimal contract S0 satisfies $b_L = w_L = 0$, $b_H = \frac{c_2}{t_1}$, and $w_H = \max\left\{t'_1 B - (t'_0 + t'_1)\frac{c_2}{t_1}, 0\right\}$.

Results 1 and 2 clarify the conditions under which the effort-sorting scheme is implementable. The implementability of the action plan is dependent on B. This can be understood by the fact that B appears in the principal's truth-telling constraints, (14) and (15). Intuitively, if B drops below some threshold, the principal would not gain from eliciting any continuation efforts (even though H is received); in this case, she always reports the signal which enables her to pay less, and hence the pooling equilibrium prevails. Given that action plan $\langle 1; 1, 0 \rangle$ is necessarily implemented by a separating equilibrium, it requires that B is large enough.

Result 3 characterizes S0. Like S1, w_H might be set to be positive for ensuring the principal's honesty in information revelation.¹⁸ We will not show subjective IPE contracts for other action plans (e.g., $\langle 0; 1, 0 \rangle$ and $\langle 0; 1, 1 \rangle$), but worth noting that all these contracts share similar feature that $w_H \ge 0$ and $w_L = 0$. The findings here shed light on the findings by Bewley (1995, 1999), in which a wage cut will impact on the worker's future productivity. This link is at odds with the traditional incentive theory, and it is usually attributed to "low morale." By stressing the "signalling" role of

¹⁸Note that under S0, w_H may still equal 0. The reason of prescribing a positive w_H is to prevent the principal mimicking H when receiving L. Notice such cheating motive is weaker when B is smaller, because the revenue gains of cheating are lowered under S0. Particularly, if $B \leq \frac{(t'_0+t'_1)}{t'_1} \frac{c_2}{t_1}$ and $\frac{t'_0}{t'_1} > \frac{t_0}{t_1}$, even though the principal lies and succeeds in using b_H to induce $e_2(L) = 1$, the principal would gain negative profit by following L. So in this case, there is no need to prescribe a positive w_H in order to deter cheating.

previous payments, however, our analysis identifies that a high wage rate paid out at the interim stage can better motivate the agent at the next stage, since it allows him to believe in success.

As with the comparison between O0 and T (Proposition 4), here the desirability of implementing S0 is driven by similar factors.

Proposition 8 Assume that c_1/c_2 is sufficiently low. There exists a cutoff B^S such that S0 yields a higher (lower) level of profit to the principal than T if $B < B^S$ ($B > B^S$). Moreover, this cutoff B^S is increasing in c_2 , t'_0 , t_0 , and decreasing in t'_1 .

Notice that a greater B not only enlarges the expected revenue loss of effort sorting, but also strengthens the cheating motive of the principal who observes L, so it demands more money to be paid out to "signal" her honesty. Thus, both forces imply that the effort-sorting scheme is desirable under subjective IPEs if B is sufficiently small. The other comparative statics have similar interpretation, which will be elaborated later.

4.3 The value of subjective IPE contracts

Like in Section 3.3, we would like to find the principal's optimal choice of contract when the option of subjective IPE is available but need not be chosen. Then we will compare it with the first best outcome. One potential difficulty is that subjective IPE contracts are not necessarily implementable and this restriction makes a rigorous comparison inviable in limited space. Thus, we only do a heuristic exploration here, providing some key insights. Formal analysis is relegated to the Appendix.

Optimal contract Panel a of Figure 4 depicts the optimal choice of the principal. There is a downward slopping curve B^S , above which T is preferred to S0 and below which T is less profitable than S0. This B^S plays a similar role as B^O does in the objective IPE. Worth noting that B^S may coincide with B^O when t'_1 is small enough, but depart from B^O as t'_1 goes larger. Intuitively, the divergency between B^O and B^S is attributed to subjective nature of IPE. A rise in t'_1 enlarges the distance between them, because it increases revenue loss of shutting down continuation effort by following L; under the effort-sorting scheme, the cheating motive of the principal who observes L is strengthened, hence more rents are paid for restoring truthful reporting. Another curve $B^{\#}$ is the minimum B to ensure the implementability of action plan $\langle 1; 1, 0 \rangle$.¹⁹ Because curve $B^{\#}$ approaches

¹⁹According to Proposition 7, for $t'_1 \leq (t'_0/t_0)t_1$, curve $B^{\#}$ is flat and equal to $\frac{(t_0+t_1)}{t_1}\frac{c_2}{t_1}$; for $t'_1 > (t'_0/t_0)t_1$, curve $B^{\#}$ is increasing in t'_1 . It is interesting to note that these two minimum B agree with each other when $t'_1 = (t'_0/t_0)t_1$.



Figure 4: The value of subjective IPEs

infinity as t'_1 approaches t_1 , it must intersect with the downward slopping curve B^S at some $t'_1 < t_1$. We denote by α the value of t'_1 at the intersection point. Thus, within the region bounded by $B^{\#}$ and B^S , the optimal contract is S0; outside such a region, the optimal contract is T.²⁰

The bottom line is clear. Due to additional truth-telling constraints, the optimal *subjective* IPE contract is less profitable than the optimal *objective* IPE contract. Nonetheless, there still exist circumstances in which the optimal subjective IPE contract payoff-dominates the optimal traditional contract.²¹

 $^{2^{0}}$ In Figure 4, we implicitly assume that $(1 - r_0 - r_1)(t_0 - t'_0) \le t_1 + t'_0$, which guarantees T payoff dominates S1 for all $t'_1 \in (0, t_1)$. Detailed analysis of the other case where such condition does not hold, see the Appendix.

 $^{^{21}}$ In our analysis, we ignore mixed equilibrium in which the principal uses mixed strategy when reporting. The primary reason for the omission is that we want to implement some deterministic action plan and mixed equilibrium is ruled out.

Comparison with the first best outcome Under subjective IPE, the comparison of the optimal contract with the first best differs from the one found under objective IPE. Basically, as illustrated by panel b of Figure 4, there are four regions of parameters. In the upper region, the principal's optimal contract implements the first best outcome (e.g., the action plan $\langle 1; 1 \rangle$). Moreover, the region in which S0 maximizes the principal's profit is partitioned by curve B^* into two parts. If max $\{B^*, B^\#\} \leq B \leq B^S$, there is under-investment in the optimal contract (while the first best implements $\langle 1; 1, 1 \rangle$ rather than $\langle 1; 1, 0 \rangle$). If $B^\# \leq B \leq B^*$, the principal's choice coincides with the first best (now equal to $\langle 1; 1, 0 \rangle$).²² These three regions resemble what we obtained in case of objective IPE. However, because of the nature of subjective IPE, $\langle 1; 1, 0 \rangle$ may not be implementable even though the first best calls for it. As a result, a new region occurs at the bottom, indicating over-investment under the optimal contract.

5 Discussions and extensions

5.1 Noisy IPE

Thus far we have assumed that the IPE is perfectly informative in the sense that, given the IPE outcome, the agent's private information about his own effort at Stage-1 — whether $e_1 = 1$ or $e_1 = 0$ — becomes irrelevant. It is realistic to assume that evaluation contains noises. For instance, given the interim product of high quality, there is a probability q > 0.5 of receiving signal H (and probability 1-q of receiving signal L); given the interim product of low quality, there is a probability 1-q of receiving signal H (and probability q of receiving signal H (and probability q of receiving signal H (and probability q of receiving signal L). In the Appendix, we redo the characterizations of O1 and T and find that $C^{O1} \leq C^T$ if and only if

$$\frac{qRt_0 + (1-q)(1-R)t'_0}{(1-q)Rt_0 + q(1-R)t'_0} \ge \frac{qRt_1 + (1-q)(1-R)t'_1}{(1-q)Rt_1 + q(1-R)t'_1},\tag{16}$$

where $R = r_0 + r_1$. If $t'_0 > 0$ and $t'_1 > 0$, (16) is equivalent to $t_0/t_1 \ge t'_0/t'_1$ and as a result assuming q = 1 is without loss of generality. In case $t'_0 = t'_1 = 0$, (16) must hold as an equality and so $C^{O1} = C^T$. In the latter case, conducting IPE is equivalent to not conducting one. Notice that this equivalence or irrelevance result is established in an early version of this paper (Chen and Chiu

 $^{2^{2}}$ In the Appendix, we show that $B^{S} \ge B^{*}$ for all $t'_{1} \in (0, t_{1})$, if some condition (related to r_{0} and r_{1}) holds. In case that such condition does not hold, there may exist an intermediate range of t'_{1} in which $B^{S} \le B^{*}$. This case introduces a complicated partition, and we ignore discussion here.

2011).

Proposition 9 Given the noisy IPE modeled in this subsection, Proposition 2 continues to hold.

Another dose of realism can be added to the subjective IPE case. Previously, we assumed that the agent does not observe the subjective IPE outcome (nor any correlated signal of it). Consequently, such an IPE is not as valuable as its objective counterpart. In fact, it is realistic to assume that the agent, as well as the court, also observes some positively correlated signal of the IPE outcome. This scenario can be seen as an intermediate case between the objective IPE case and the subjective IPE case that we have studied. We conjecture that, in this case, the IPE will be more useful than it is in the subjective IPE studied in this paper.

5.2 Other parameter values

In the previous analysis, we have imposed assumption A1 so that under traditional contracting the agent's Stage-2 IC constraint tends to be binding. We found the conditions under which conducting an IPE either enhances or lowers the principal's profit. Alternatively, similar analysis can be applied to other parameter settings, but slightly different results are obtained.

Suppose that the optimal traditional contract T stills implement $\langle 1; 1 \rangle$, but its binding IC constraint is the one that prevents deviation to the action plan $\langle 0; 1 \rangle$, instead of $\langle 1; 0 \rangle$. This scenario occurs when c_1/c_2 and B are both sufficiently large. We can characterize contract T again and denote it by (w^T, b^T) . It is interesting to compare such T with its counterpart under the objective IPE (which is denoted by O1 and characterized by $(w_H^{O1}, w_L^{O1}, b_H^{O1}, b_L^{O1})$). Unsurprisingly, according to Lemma 1, all the wage rates (e.g., w) are set to be zero; but the choices of bonuses, as well as the comparison between implementation costs, follows a different pattern, in contrast with what we found under assumption A1.

Proposition 10 Suppose that B is sufficiently large such that the optimal traditional contract T still implements $\langle 1; 1 \rangle$. If

$$\frac{c_1}{c_2} \ge \frac{r_1 \left[(t_0 + t_1) - (t'_0 + t'_1) \right]}{t'_1},\tag{17}$$

then it holds true that $b_L^{O1} \leq b_H^{O1} \leq b^T$ and $C^{O1} \leq C^T$.

Intuitively, if c_1 is sufficiently large relative to c_2 , bonuses are chosen as instruments to deter the agent shirking in Stage 1. With help of an IPE, the principal can increase b_H alone (while keeping

 b_L at the minimum of inducing $e_2(L) = 1$), so it enables the agent to enjoy a rise in rents only after a more favorable Stage-1 outcome is achieved. (However, setting a higher *b* under traditional contract increases his rents even after a less favorable Stage-1 outcome is achieved.) Given that b_H is a more cost-effective device than *b* in relaxing the binding Stage-1 IC constraints, the IPE contract tends to be more desirable.²³

5.3 The agent can destroy evidence

In reality, the principal needs cooperation from the agent — for instance, to provide details or documentation — in her IPE exercise. While falsification of details is less likely and easier to detect, the strategic withholding of details is more difficult to detect. Suppose the agent can choose to provide full details or partial details (and the principal is not able to distinguish between them). In case of full details, the principal learns $\sigma = H$ if and only if $x_1 = 1$. In the case of provision of partial details, she receives $\sigma = H$ with probability p (probability 1 - p) and $\sigma = L$ with the remaining probability if $x_1 = 1$ (if $x_1 = 0$). Consideration of such strategic action on the part of the agent will mean that an IPE is less likely to be beneficial.

To fix idea, consider this problem in the context of objective IPE with the goal of implementing action plan $\langle 1; 1, 1 \rangle$. Without aforementioned problem, the optimal IPE contract is simply O1, characterized by Proposition 1. Suppose now O1 was used and the principal thought the agent would provide full details. However, because $b_H < b_L$ and $w_H = w_L$, the agent would always like the principal to receive that $\sigma = L$. Therefore, he would provide partial details to misguide the principal. To prevent such strategic behavior, the contract needs to be adjusted to make the realization of signal H more profitable to the agent, through an increase in b_H or an increase in w_H or both. In any case, it is more costly to the principal. Notice that under a traditional contract, such strategic behavior on the part of the agent is not useful for him because his payoff depends only on the final product's quality which he cannot falsify. As a result, the possibility of the agent destroying evidence makes an IPE less likely to be desirable in this case.

Our discussion has been confined to action plan $\langle 1; 1, 1 \rangle$. If the action plan to be implemented is $\langle 1; 1, 0 \rangle$, the contract O0 shall remain feasible and optimal (given the action plan) because now $b_L = 0$

²³According to Proposition 2, when assumption A1 holds, desirability of O1 is further guaranteed by a small enough c_1/c_2 . It seems contrary to what we find here. But this is not true. The reason is that, when c_1/c_2 is in some medium range, introduction of b_L may trigger the Stage-1 IC constraint binding under O1, while the Stage-1 IC constraint is not binding under T. However, Proposition 10 says the benefit of using b_H when Stage-1 IC constraints are binding under both O1 and T.

and the agent has no incentive to misguide the principal that $\sigma = L$. Thus, this consideration makes $\langle 1; 1, 0 \rangle$ more profitable to be implemented by the principal than $\langle 1; 1, 1 \rangle$. We leave those issues for future studies.

5.4 Alternative benchmark

In our study of the desirability of IPEs, the benchmark we use for comparison is where neither party observes any signal about the interim product's quality. An alternative benchmark is where the agent has some additional signal about the interim product's quality. In this case, his private information includes not only his first-stage effort choice but also this additional signal. We argue that, under this new benchmark, conducting an IPE is even more likely to be desirable, and hence using the old benchmark is a prudent choice.

To be more specific, suppose the agent receives signal H(L) if and only if $x_1 = 1$ (0). Then the contract T stipulated by (3) is no longer able to implement action plan $\langle 1; 1, 1 \rangle$. The reason for this is that given signal L, the minimum b required for the agent to exert $e_2 = 1$ is c_2/t'_1 and unfortunately $b^T < c_2/t'_1$. Hence, to implement action plan $\langle 1; 1, 1 \rangle$, a more generous contract is required. In general, the new benchmark will favor the use of an IPE more than the old benchmark would.

5.5 The agent has unknown abilities

In this paper, we assume that there is only one type of agent and his productivity is commonly known. In reality, the agent may have different abilities and may not be certain about his exact productivity. To be more specific, consider the following modification to the model in Section 2. Suppose there are two types of agent: h (high) and l (low). It is commonly believed that an agent is of the high type with probability of $\omega \in (0, 1)$. Although both types have the same costs regarding their effort, the h-type agent is more productive than the l-type, in the following sense. The i-type agent's Stage-1 effort increases the probability of the interim product's success by $\theta_i r_1$; his Stage-2 effort increases the probability of the final product's success by $\theta_i t_1$ ($\theta_i t'_1$) when the interim product is high quality (low quality), where $\theta_h > \theta_l$. We continue to focus on the case where c_1/c_2 is sufficiently small and the optimal traditional contract implements action plan $\langle 1; 1 \rangle$.

Suppose an objective IPE is used. Upon a low IPE signal, the agent will now be less confident,

not only because the interim quality is lower but also because the probability of his being a low type of agent is higher than was previously thought and hence the probability of success of the Stage-2 effort is even slimmer. Thus, we conjecture that the main advantage of the IPE, compared with the traditional contract, is attributable more to the fact that it allows the principal to give different instruments to the agent conditional on the IPE outcome (to implement $\langle 1; 1, 0 \rangle$), than to the fact that it incentivizes with a lower cost under the same continuation action plan under both high and low signal (to implement $\langle 1; 1, 1 \rangle$). We leave this interesting issue for future studies.²⁴

5.6 The opportunity of correction

In reality, an early stage that has resulted in a poor outcome may be repeated. The opportunity to redo the first task is an additional reason why conducing an IPE is beneficial to the principal. The simplest model to consider is the same model as in Section 2, except that, upon receiving the IPE outcome, the principal is allowed to ask the agent or somebody else to redo the Stage-1 task. Assume that (i) a task can only be redone once, (ii) redoing a task incurs an extra cost I for the principal due to the delay in project completion, additional materials, etc., and (iii) redoing the first task is like starting it afresh and previous experience is of no use to the second attempt. The final assumption has two implications. Firstly, redoing the first task is valuable only if signal L obtained. Secondly, in redoing the task, it is better to hire a new agent. This is because if the original agent is retained, he would gain positive rent even though he failed to pass the milestone in the first attempt, hence the original moral hazard problem would be exacerbated.

The basic trade-off of imposing a correction is that subsequent to a poor IPE outcome, hiring a new agent to redo the previous task increases the probability of achieving final success although the likelihood of being fired makes it more costly to motivate the first agent, due to the worsened agency problem in Stage 1 as well as extra difficulties concerning information disclosure. Thus, correction is favored by a larger B, but disfavored by a larger c_1/c_2 or I.

 $^{^{24}}$ Notice that given our assumptions, the principal will never be certain about the type of agent, even with the help of the IPE. We can envision scenarios, however, in which the type of agent is truly confirmed. The bottom line is that the IPE allows the principal to enhance her knowledge about the innate ability of the agent, and this provides an additional rationale as to why an IPE may be contemplated.

5.7 Multiple agents

In practice, the evaluation of an agent's performance is sometimes done by another agent on behalf of the principal. The reviewer may be an outside expert who has the expertise to access the reviewee's performance, or an inhouse supervisor of the reviewee with whom he has a close working relationship. In either case, there is a trade-off in using such a third party reviewer. While this reviewer may have advantageous technology or expertise to access the IPE outcome, his presence also introduces another moral hazard problem. If such a third party agent is indispensable to the IPE, the IPE is worth conducting only if the moral hazard problem associated with him is small enough.

There is now a strand of literature that extends the usual two-layer principal-agent framework into a three-layer principal-supervisor-agent framework. In particular, according to Ishiguro and Itoh (2001), despite the potential moral hazard problem arising from the supervisor, the existence of the supervisor actually leads to the first best outcome, which is infeasible under a two-layer model. The result relies subtly on the possibility of collusion between the supervisor and the agent in risk sharing, followed by contract renegotiation between the principal and the supervisor (see Felli and Villas-Boas 2000 and Chiu and Chou 2006 for more detail). Despite differences between Ishiguro and Itoh (2001) (and this line of studies more generally) and our framework — most notably, the supervisor's ability to observe the agent's effort and contract being renegotiable — it suggests that introducing a third party supervisor need not lead to additional moral hazard problems. In this spirit, an IPE that requires a third party reviewer need not be less desirable than one that does not require such a third party.

5.8 Relationship with the literature on interim feedback

In a two-period principal-agent model, Lizzeri, Meyer and Persico (2002) show that it is undesirable for the principal to reveal the IPE outcome to the agent given that she intends to implement the same continuation effort. Manso (2011) obtains a similar result in another two-period principalagent model. In both papers, by assumption, the IPE outcome is verifiable at the end and can be used in contracting. The question is whether she should reveal the information to the agent before his stage 2 decision making. Importantly, it is assumed that a contract made contingent on the IPE outcome is enforceable even in the absence of feedbacks.

In our paper, the question is whether to conduct the IPE. We impose a restriction that, in case

no interim feedback is given to the agent, no contracts made contingent on the IPE outcome are enforceable, i.e., any enforceable contract must be of the format (w, b).

Despite other differences,²⁵ the aforementioned difference in contractibility is key to understanding our apparently conflicting results. We can indeed replicate Lizerri, Meyer, and Persico's result within our framework through just a change of the assumption.

Proposition 11 Suppose (1) the principal observes the IPE outcome and can commit not to give any feedback to the agent; (2) a contract of the format (w_H, w_L, b_H, b_L) is enforceable whether or not feedbacks are involved; and (3) the effort plan to implement is $\langle 1; 1, 1 \rangle$. Then the principal's payoff is weakly higher under the regime without feedbacks than the one with feedbacks.

Notice that the contract with feedback is just O1. We show that O1 is still feasible as a contract without feedback; i.e., the agent will accept such an offer and also implement the same action plan. Nonetheless, there is room for further improvement because the principal can re-adjust the two bonuses to save the implementation cost without jeopardizing the agent's incentive.

To see this, suppose only the second stage IC constraints are binding under the O1 problem. In this case $b_H = c_2/t_1$ and $b_L = c_2/t'_1$. Suppose for the contract without feedback, denoted by O_1^* and characterized by $(w_H^*, w_L^*, b_H^*, b_L^*)$, the only binding IC constraint is the second stage constraint, i.e.,

$$(r_0 + r_1) t_1 b_H^* + (1 - r_0 - r_1) t_1' b_L^* \ge c_2.$$

Although $b_H^* = b_H$ and $b_L^* = b_L$ satisfy this constraint, the principal can reduce the implementation cost further through the following bang-bang choices. Specifically, $b_H^* = 0$ and $b_L^* = c_2/((1 - r_0 - r_1)t_1')$ if $t_0/t_1 \ge t_0'/t_1'$, and $b_H^* = c_2/((r_0 + r_1)t_1)$ and $b_L^* = 0$ otherwise. Hence, except for the case where $t_0/t_1 = t_0'/t_1'$, the implementation cost under O_1^* is strictly lower than under O_1 .

5.9 Extrinsic and Intrinsic Motivation

The seminal work by Bénabou and Tirole (2003) studies the interplay between extrinsic incentive and intrinsic motivation. It introduces notions such as self-confidence, trust, etc., into a specific class

 $^{^{25}}$ Mostly noticably, Lizerri, Meyer, and Persico (2002) assume continuous efforts and outputs in both periods, and outputs are outcomes of the current stage efforts only. However, we assume that the interim product created in the first stage is only transient and has no independent value once the final output is created.

of principal-agent model. What is novel about their model is that the principal knows the agent's ability better than the agent does and the agent receives some utility from achieving the goal (hence the term of intrinsic motivation). In equilibrium the agent infers that his ability is lower when he is promised a higher reward. Bénabou and Tirole are therefore able to relate to notions, such as self-confidence, that are foreign to economics but popular in psychology and education.

Notice that our model of objective IPE suggests that the low-ability agent be given a stronger incentive, with the interpretation that "ability" corresponds to the quality of the interim product. In this sense, the result resembles that of Bénabou and Tirole and in fact we have attributed the necessity of a stronger incentive to non-economic notions such as low morale or lack of self-confidence of the agent. Due to the fact that in our model there is neither asymmetric information at the outset nor any intrinsic motivation on the part of the agent, it appears to us that notions such as low morale and self confidence have their justifiable places in even broader economic environments. (Notice that the subjective IPE case also has a relevant connection with Bénabou and Tirole (2003). There "ability", which corresponds to quality of the first step in the production process, is the principal's private information.)

6 Conclusions

In this paper, we have studied a 2-stage principal-agent model with different specifications. We have shown that using an IPE may boost the principal's profit even though he did not plan to condition the continuation action plan on the IPE outcome. We have clarified the exact conditions under which it occurs. Albeit weaker, the domination of the IPE contract over the traditional contract is still possible under subjective IPE where the principal's incentive to lie is a concern.

We end with some remarks relating the theoretical findings here to evaluation or contract design in practice. First, we found that using an IPE may either boost or worsen the principal's payoff even if the action plan is not intended to vary with the IPE outcome. The trade-off we identified may provide hint to understand mixed findings on the effectiveness of IPE in such a scenario. For example, it is conceivable that a teacher always wants his or her student to work hard even after a negative result in a midterm exam. The laboratory study of 90 undergraduate students, conducted by Podsakoff and Farh (1989), documents that such a feedback system can both enhance and lower task performance and goal setting.²⁶ Second, as our results show, the principal may gain from sorting the continuation effort contingent on the feedback. This result resonates well with observations that IPEs are used more extensively in project management when the purpose of setting up such a feedback system is to control or plan the project procedure (See Meredith and Mantel 1995). Third, in case only subjective feedbacks can be given, the principal must make the contract attached to it credible. The characterization of subjective IPE contract suggests that stage work or task will be compensated right after a milestone is passed. This progressive payment schedule is the critical feature of so-called "milestone based contracts", which are widely used in procurements.²⁷

There are two directions that future studies can pursue. One is to weaken the assumption that the IPE outcome always reveals perfectly the quality of the interim product. In fact, we relaxed this assumption in Section 5.1, arguing that the main quantitative properties regarding desirability of IPEs still hold true. It is interesting to go a step further to spell out additional implications. Another direction for future studies is re-visiting the issue when efforts are substitutes, rather than complements.

 $^{^{26}}$ Hattie and Timperley (2007) systematically investigated the influence of feedback in classrooms, also with mixed results. On the one hand, they found feedback can be very powerful in enhancing learning in some circumstances, but on the other hand, feedback would have negative impact on learning and achievement.

²⁷According to Ariba Online Help, for writing such a contract, it is necessary to specify "amount to be paid to the supplier upon successful completion of each milestone". Also, for real life examples of such contract, refer to the payment schedules used by the US Army Corps of Engineers (http://www.hnd.usace.army.mil/umcs).

References

- Aberdeen Group. 2010. Employee Performance Management: The Alpha and the Omega of Talent Strategy and Business Execution. Boston, MA.
- [2] Aoyagi, M., 2008. "Information Feedback in a Dynamic Tournament." Mimeo, The Institute of Social and Economic Research, Osaka University.
- [3] Bewley, T. F., 1995. "A Depressed Labor Market as Explained by Participants." American Economic Review 85, 250-254.
- [4] Bewley, T. F., 1999. Why Wages Don't Fall During a Recession, Harvard University Press.
- [5] Bénabou, Robert, and Jean Tirole, 2003. "Intrinsic and Extrinsic Motivation," *Review of Economic Studies* 70 (3), 489-520.
- [6] Bjerke, D., J. Cleveland, R. Morrison, and W. Wilson. 1987. Officer Fitness Report Evaluation Study. Navy Personnel Research and Development Center Report TR 88-4. San Diego, CA: Navy Personnel Research and Development Center.
- [7] Bolton, P., D. Scharfstein, 1998. "Corporate Finance, the Theory of the Firm, and Organizations." Journal of Economic Perspectives 12, 95–114.
- [8] Calveras, A., Ganuza, J., Hauk, E., 2004. "Wild Bids. Gambling for Resurrection in Procurement Contracts." *Journal of Regulatory Economics* 26, pp. 41-68.
- Che, Y.-K., Yoo, S.-W., 2001. "Optimal Incentive for Teams." American Economic Review 91, pp. 525-541.
- [10] Chen, Bin R. and Y. Stephen Chiu 2011. "Interim Performance Evaluation in Contract Design," Available at SSRN: http://ssrn.com/abstract=1862484
- [11] Chiu, Y. Stephen and Eric S. Chou 2006. "Collusion and Renegotiation in Principal-Supervisor-Agent Models" mimeo, HKU.
- [12] Cleveland, Jeanette N., Kevin R. Murphy, and Richard E. Williams, 1989. "Multiple Uses of Performance Appraisal: Prevalence and Correlates." *Journal of Applied Psychology*, 74(1), pp. 130-35.

- [13] Cook, F. W., 1975. "How to Ruin Pay With Motivation." Paper delivered at Executive Enterprise's Seminar on Current Developments in Executive Compensation Plans and Packages, New York.
- [14] Crémer, J., 1995. "Arm's Length Relationships." Quarterly Journal of Economics CX, pp. 275-295.
- [15] Ederer, F., 2010. "Feedback and Motivation in Dynamic Tournaments." Journal of Economics and Management Strategy vol 19(3), pp. 733-769.
- [16] Felli, L. and J. M. Villas-Boas 2000, "Renegotiation and Collusion in Organization," Journal of Economics and Management Strategy, 9 (4), 453-483.
- [17] Fuchs, W., 2007. "Contracting with Repeated Moral Hazard and Private Evaluations." American Economic Review 97, pp. 1432-1448.
- [18] Gershkov, A., M. Perry, 2009. "Tournaments with Midterm Reviews." Games and Economic Behavior 66, pp. 162-190.
- [19] Goltsman, M., Mukherjee, A., 2011. "Interim Performance Feedback in Multistage Tournaments: the Optimality of Partial Disclosure." *Journal of Labor Economics* vol 29(2), pp. 229-265.
- [20] Ishiguro, S. and H. Itoh 2001, "Moral Hazard and Renegotiation with Multiple Agents," *Review of Economic Studies*, 68, 1-20.
- [21] Hattie, J. A. and H Timperley. 2007. "The Power of Feedback." *Review of Educational Research* Vol. 77, No. 1, pp. 81–112
- [22] Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives. Oxford: Princeton University Press.
- [23] Levin, J., 2003. "Relational Incentive Contracts." American Economic Review 93, pp. 835-857.
- [24] Levine, H.Z., 1986. "Performance Appraisals at Work." Personnel, 63(6), pp. 63-71.
- [25] Lizzeri, A., M. Meyer., N. Persico, 2002. "The Incentive Effects of Interim Performance Evaluations." CARESS working paper #02-09.

- [26] MacLeod, W.B., 2003. "Optimal Contracting with Subjective Evaluation." American Economic Review 93, pp. 216-240.
- [27] Manso, G., 2011. "Motivating Innovation." forthcoming, Journal of Finance.
- [28] Meredith, J., S. Mantel, 1995. Project Management: a Managerial Approach. 3rd edition, New York: John Wiley & Sons, Inc, pp. 125-126 & 446.
- [29] Murphy, K., 1992. "Performance Measurement and Appraisal: Motivating Managers to Identify and Reward Performance." in William J. Burns, Jr., eds., *Performance Measurement, Evaluation, and Incentives.* Boston: Harvard Business School Press.
- [30] Murphy, K. M., and J. N. Cleveland. 1985. "Understanding Performance Appraisal: Social, Organizational, and Goal-Based Perspectives." SAGE Publications.
- [31] Murphy, K. M., and J. N. Cleveland, 1991. Performance Appraisal: An Organizational Persepctive. Allyn and Bacon, Needham Heights, Massachusetts.
- [32] Podsakoff, P. M., Farh, J. L. 1989. "Effects of Feedback Sign and Credibility on Goal Setting and Task Performance." Organizational Behavior and Human Decision Processes 44, pp. 45–67.
- [33] Prendergast, C., 1999. "The Provision of Incentives in Firms." Journal of Economic Literature 37, pp. 7-63.
- [34] Raith, M., 2009. "Optimal Incentives and the Time Dimension of Performance Measurement." Mimeo, University of Rochester.
- [35] Ray, K., 2007a. "Performance Evaluations and Efficient Sorting." Journal of Accounting Research 45, pp. 839-882.
- [36] Ray, K., 2007b. "The Retention Effect of Withholding Performance Information." The Accounting Review 82, pp. 389-425.
- [37] Reid, A., 1999. Project Management: Getting It Right. Cambridge: Woodhead Publishing Ltd.
- [38] Rendero, T., 1980. "Performance Appraisal Practices." Personnel, 57(6), pp. 4-12
- [39] Roberts, J., 2004. The Modern Firm: Organizational Design for Performance and Growth. New York: Oxford University Press.

- [40] Scharfstein, D., J. Stein, 2000. "The Dark Side of the Internal Capital Markets: Divisional Rent Seeking and Inefficient Investments." *Journal of Finance* 55, pp. 2537–2564.
- [41] Schmitz, P.W., 2005. "Allocating Control in Agency Problems with Limited Liability and Sequential Hidden Actions." *RAND Journal of Economics* 36, pp. 318-336.
- [42] Yildirim, H., 2005. "Contests with Multiple Rounds." Games and Economic Behavior 51, pp. 213-227.
- [43] Zhao, R., 2009. "Productive Low Morale." *Economics Letters* 103, pp. 18-22.

Appendix A: Proofs

A.1 Proof of Lemma 1

Proof. Suppose the principal wants to implement the action plan $\langle x; y, z \rangle$. No matter what x, y and z are, w_L should be zero in the cost-minimized contract, since it helps nothing except for diluting the incentive in Stage 1. Then we consider the choice of w_H . It incentivizes the Stage-1 effort, so it would be useful only if x = 1. However, we claim that w_H cannot be a more cost-effective instrument than b_H is for inducing $e_1 = 1$.

Consider the case where x = 1. If the principal increases one unit of w_H , the cost is $(r_0 + r_1)$, while the Stage-1 IC constraints would be relaxed by the amount of r_1 . However, if the principal increases one unit of b_H , the cost is $(r_0 + r_1)(t_0 + yt_1)$; the IC constraints for preventing deviation to $\langle 0; y, z \rangle$ would be relaxed by $r_1(t_0 + yt_1)$, while the IC constraints for preventing deviation to $\langle 0; 0, 0 \rangle$ is relaxed by the amount of $[r_0t_1 + r_1(t_0 + t_1)]$. So by comparing the cost-benefit ratio, we found that b_H is at least as cost-effective as w_H for relaxing relevant IC constraints.

A.2 Proof of Proposition 1

Proof. Under objective IPE, the cost-minimization problem for implementing $\langle 1; 1, 1 \rangle$ is as follows:

$$\min_{w_H, w_L, b_H, b_L} C^{O1} = (r_0 + r_1) w_H + (1 - r_0 - r_1) w_L
+ (r_0 + r_1) (t_0 + t_1) b_H + (1 - r_0 - r_1) (t'_0 + t'_1) b_L,$$

subject to the following IC constraints: the IC constraints that prevent deviation to $\langle 1; 0, 1 \rangle$ and $\langle 1; 1, 0 \rangle$, respectively (i.e., (5) and (6)); the IC constraint that prevents deviation to $\langle 0; 1, 1 \rangle$ (i.e., (7); and the IC constraint that prevents deviation to $\langle 0; 0, 0 \rangle$:

$$[r_0 t_1 + r_1 (t_0 + t_1)] b_H + [(1 - r_0) t'_1 - r_1 (t'_0 + t'_1)] b_L + r_1 (w_H - w_L) \ge c_1 + c_2;$$
(18)

The claim that $w_H = w_L = 0$ is obtained from Lemma 1. To find out the bonuses, let us define

$$\gamma \equiv \frac{r_1 t_0}{t_1} - \frac{r_1 t_0'}{t_1'}.$$

We can verify that that (i) If $c_1 < \gamma c_2$, only (5) and (6) are binding and others constraints are non-binding; as a result, $b_H = \frac{c_2}{t_1}$, $b_L = \frac{c_2}{t_1}$. (ii) If $c_1 \ge \gamma c_2$, only (6) and (7) are binding and others are non-binding; as a result b_H and b_L satisfy (6) and (8) as equalities. (Note that in any case, (18) is a non-binding constraint. It can be explained by one-deviation property.) (iii) $c_1 < \gamma c_2$ if and only if

$$\frac{c_2}{t_1} > \frac{c_1}{r_1 \left(t_0 + t_1 \right)} + \frac{\left(t_0' + t_1' \right) c_2}{\left(t_0 + t_1 \right) t_1'}.$$

Hence, (i) to (iii) establish the claim that $b_H = \max\left\{\frac{c_2}{t_1}, \frac{c_1}{r_1(t_0+t_1)} + \frac{(t'_0+t'_1)c_2}{(t_0+t_1)t'_1}\right\}$ and $b_L = \frac{c_2}{t'_1}$.

A.3 Proof of Proposition 2

Proof. To simplify notation, we define

$$R \equiv (r_0 + r_1);$$

$$\alpha_H \equiv R(t_0 + t_1), \ \alpha_L \equiv (1 - R)(t'_0 + t'_1);$$

then using (3) and (4) we reckon that $C^T = (\alpha_H + \alpha_L) b^T$. Suppose $c_1 < \gamma c_2$ where $\gamma \equiv \frac{r_1 t_0}{t_1} - \frac{r_1 t'_0}{t'_1}$. Then $b_H^{O1} = c_2/t_1$ and, with some manipulation, we have

$$C^{O1} = \left(\alpha_H \beta_H + \alpha_L \beta_L\right) b^T,$$

where

$$\beta_H \equiv \frac{[Rt_1 + (1 - R)t'_1]}{t_1}, \ \beta_L \equiv \frac{[Rt_1 + (1 - R)t'_1]}{t'_1}.$$

Hence

$$C^{T} - C^{O1} = (\alpha_{H} + \alpha_{L}) b^{T} - (\alpha_{H}\beta_{H} + \alpha_{L}\beta_{L}) b^{T}$$
$$= [\alpha_{H} (1 - \beta_{H}) + \alpha_{L} (1 - \beta_{L})] b^{T}.$$

Substituting the following into the expression

$$1 - \beta_H = 1 - \frac{[Rt_1 + (1 - R)t'_1]}{t_1} = \frac{(1 - R)(t_1 - t'_1)}{t_1}$$

and

$$1 - \beta_L = 1 - \frac{[Rt_1 + (1 - R)t_1']}{t_1'} = -\frac{R(t_1 - t_1')}{t_1'},$$

and with some manipulation, we obtain

$$\begin{split} C^{T} - C^{O1} &= \left(R \left(1 - R \right) \left(t_{1} - t_{1}^{\prime} \right) \frac{\left(t_{0} + t_{1} \right)}{t_{1}} - R \left(1 - R \right) \left(t_{1} - t_{1}^{\prime} \right) \frac{\left(t_{0}^{\prime} + t_{1}^{\prime} \right)}{t_{1}^{\prime}} \right) b^{T} \\ &= \left(R \left(1 - R \right) \left(t_{1} - t_{1}^{\prime} \right) \left(\frac{t_{0} + t_{1}}{t_{1}} - \frac{t_{0}^{\prime} + t_{1}^{\prime}}{t_{1}^{\prime}} \right) \right) b^{T} \\ &= R \left(1 - R \right) \left(t_{1} - t_{1}^{\prime} \right) \left(\frac{t_{0}}{t_{1}} - \frac{t_{0}^{\prime}}{t_{1}^{\prime}} \right) b^{T}. \end{split}$$

Since $b^T > 0$ and $t_1 > t_1', C^T - C^{O1} \ge 0$ if and only if

$$\frac{t_0}{t_1} \ge \frac{t_0'}{t_1'}.$$

Next, we consider the case where $c_1 \ge \gamma c_2$. In this case,

$$b_{H}^{O1} = \frac{c_{1}}{r_{1}\left(t_{0}+t_{1}\right)} + \frac{\left(t_{0}'+t_{1}'\right)c_{2}}{\left(t_{0}+t_{1}\right)t_{1}'} > \frac{c_{2}}{t_{1}},$$

and the above calculation underestimates the true C^{O1} . Hence, *a fortiori*, it must hold true that $C^{O1} > C^T$ when $\frac{t_0}{t_1} < \frac{t'_0}{t'_1}$. To conclude, given $\frac{t'_0}{t'_1} > \frac{t_0}{t_1}$, $C^T < C^{O1}$. Given $\frac{t'_0}{t'_1} \le \frac{t_0}{t_1}$, for sufficiently small c_1/c_2 , we have

 $C^T > C^{O1}.$ The proof is thus complete. \blacksquare

A.4 Proof of Proposition 3

Proof. When implementing (1; 1, 0), the principal's cost-minimization problem is as follows.

$$\min_{w_H, w_L, b_H, b_L} C^{O0} = (r_0 + r_1) (t_0 + t_1) b_H + (1 - r_0 - r_1) t'_0 b_L + (r_0 + r_1) w_H + (1 - r_0 - r_1) w_L,$$

The IC constraint that prevents deviations to $\langle 1; 0, 0 \rangle$ is (5). The IC constraint that prevents deviation to $\langle 0; 0, 0 \rangle$ is

$$[(r_0 + r_1)t_1 + r_1t_0]b_H - r_1t'_0b_L$$

$$+r_1(w_H - w_L) \ge c_1 + (r_0 + r_1)c_2.$$
(19)

The IC constraint that prevents deviation to (0; 1, 0) is

$$r_1(t_0 + t_1) b_H - r_1 t'_0 b_L + r_1(w_H - w_L) \ge c_1 + r_1 c_2.$$
⁽²⁰⁾

First, let note that a positive b_L or w_L only dilutes incentive in Stage 1, so both of them should be zero. Second, according to Lemma 1, it is without loss of generality to set $w_H = 0$. Thus, when O0 contract achieves optimum, only b_H is positive. Third, we further find that if $\frac{c_1}{c_2} \leq \frac{r_1 t_0}{t_1}$, only (5) is binding; if $\frac{c_1}{c_2} > \frac{r_1 t_0}{t_1}$, only (20) is binding. In any case, (19) would not be binding. Hence, the claim in Proposition 3 is established.

A.5 Proof of Proposition 4

Proof. Implementing O0 is more profitable than implementing T, if and only if

$$V^T - V^{O0} \le C^T - C^{O0}.$$
 (21)

Note that the LHS of the equation equals $(1 - r_0 - r_1) t'_1 B$, and its RHS equals $C^T - (r_0 + r_1) (t_0 + t_1) b_H$, where C^T is defined in (4) and the choice of b_H is characterized in Proposition 3. Therefore, (21) is equivalent to

$$B \le B^{O} \equiv \frac{1}{(1 - r_0 - r_1) t_1'} \left(C^T - (r_0 + r_1) (t_0 + t_1) b_H \right).$$

Substituting C^T and b_H into it, we obtain

$$B^{O} = \frac{1}{(1 - r_{0} - r_{1})t_{1}'}$$

$$\times \left(\frac{(r_{0} + r_{1})t_{0} + (1 - r_{0} - r_{1})t_{0}'}{(r_{0} + r_{1})t_{1} + (1 - r_{0} - r_{1})t_{1}'}c_{2} - (r_{0} + r_{1}) \times \max\left\{\frac{t_{0}c_{2}}{t_{1}}, \frac{c_{1}}{r_{1}}\right\}\right) + \frac{c_{2}}{t_{1}'}$$

$$(22)$$

It is routine to show that B^O defined in (22) is increasing in c_2, t'_0, t_0 , but is decreasing in t'_1 .

A.6 Proof of Proposition 6

Proof. For c_1/c_2 sufficiently small, we reckon that the implementation cost of S1 is

$$C^{S1} = (r_0 + r_1) \left(t_0 + t_1 - t'_0 - t'_1 \right) \frac{c_2}{t_1} + \left(t'_0 + t'_1 \right) \frac{c_2}{t'_1}.$$

Differentiating it with respect to t'_1 , we have

$$\frac{\partial C^{S1}}{\partial t_1'} = -\left(r_0 + r_1\right) \frac{c_2}{t_1} - \frac{t_0'}{\left(t_1'\right)^2} c_2.$$

Differentiating C^T , (4), with respect to t'_1 , we have

$$\frac{\partial C^T}{\partial t_1'} = -(1-r_0-r_1) \frac{\left[(r_0+r_1)t_0 + (1-r_0-r_1)t_0'\right]}{\left[(r_0+r_1)t_1 + (1-r_0-r_1)t_1'\right]^2} c_2.$$

Focusing at the point where $t'_1 = t_1$, we have

$$\frac{\partial C^{S1}}{\partial t'_1}\Big|_{t'_1=t_1} = -\left(r_0 + r_1 + \frac{t'_0}{t_1}\right)\frac{c_2}{t_1} < 0,$$

and

$$\frac{\partial C^T}{\partial t'_1}\Big|_{t'_1=t_1} = -\left(1-r_0-r_1\right)\frac{\left[\left(r_0+r_1\right)t_0+\left(1-r_0-r_1\right)t'_0\right]}{t_1^2}c_2 < 0.$$

It is easy to show that

$$\left. \frac{\partial C^{S1}}{\partial t_1'} \right|_{t_1'=t_1} > \left. \frac{\partial C^T}{\partial t_1'} \right|_{t_1'=t_1}$$

if and only if

$$(1 - r_0 - r_1)(t_0 - t'_0) > t_1 + t'_0.$$
⁽²³⁾

Notice that if $t'_1 = \frac{t'_0}{t_0}t_1$, $C^{S1} \ge C^{O1} = C^T$ and if $t'_1 = t_1$, $C^{S1} = C^T$. Altogether, this suggests that there exists $\tau \in ((t'_0/t_0)t_1, t_1)$ such that for all $t'_1 \in (\tau, t_1)$, the implementation cost under S1 is lower than under T. This completes the proof. (Notice that the condition (23) is very restrictive but is not impossible. For example, when $r_0 + r_1 = 0.4, t_0 = 0.5, t'_0 \doteq 0$, then the t_1 that satisfies this condition can be as high as 0.3. If t'_1 is close to 0.3 as well, then t'_1/t_0 is extremely high. But



Figure 5: Solving the subjective IPE contract S0.

this is precisely what makes S1 payoff-dominate T.)

A.7 Proof of Proposition 7

Proof. To characterize S0, let note that its cost-minimization problem is the same as that of O0 (described in the proof of Proposition 3), except that (14) and (15) are added. It is conceivable that $w_L = 0$ and $b_L = 0$ as in O0; an increase in either of them would only increase the implementation cost as well as making (15) more difficult to hold. Then the question boils down to finding a pair of b_H and w_H .

Using a graphical method as we did for S1, we depict three constraints (5), (14), and (15) by lines IC_H , H, and L, respectively.

Notice that now in general H line and L line do not have a common horizontal intercept. (i) When $\frac{t'_0}{t'_1} < \frac{t_0}{t_1}$ (refer to panel a of Figure 5), the H line's horizontal intercept is smaller. A shaded region satisfying all three constraints exists, if and only if the IC_H line is on the left hand of the interception point of the H line and L line. The corresponding condition is

$$B > \frac{(t_0 + t_1 - t_0' - t_1')}{(t_1 - t_1')} \frac{c_2}{t_1}.$$

(ii) When $\frac{t'_0}{t'_1} > \frac{t_0}{t_1}$ (refer to panel b of Figure 5), the *H* line's horizontal intercept is greater. A shaded region satisfying all three constraints exists if and only if the IC_H line is on the left hand of that intercept. The corresponding condition is

$$B > \frac{(t_0 + t_1)}{t_1} \frac{c_2}{t_1}.$$

In both cases, a minimum B is required for a solution to exist; if it does, there exists a shaded region and the optimal choice of (b_H, w_H) pair is the dotted point inside it. (Note that in panel b, there is a chance that the neither (14) and (15) is binding when the IC_H line just stands strictly in between the two intercepts.)

A.8 Proof of Proposition 8

Proof. We compare S0 with T. S0 payoff-dominates T if

$$V^T - V^{S0} \le C^T - C^{S0}.$$
 (24)

In case that $w_H > 0$, (24) is equivalent to

$$(1 - r_0 - r_1) t_1' B \leq \left[\frac{(r_0 + r_1) t_0 + (1 - r_0 - r_1) t_0'}{(r_0 + r_1) t_1 + (1 - r_0 - r_1) t_1'} c_2 - (r_0 + r_1) \frac{t_0}{t_1} c_2 \right] + (1 - r_0 - r_1) c_2 - (r_0 + r_1) \left[t_1' B - (t_0' + t_1') \frac{c_2}{t_1} \right],$$

which is rearranged to

$$B \leq B' \equiv \frac{1}{t_1'} \left[\frac{(r_0 + r_1) t_0 + (1 - r_0 - r_1) t_0'}{(r_0 + r_1) t_1 + (1 - r_0 - r_1) t_1'} - (r_0 + r_1) \frac{t_0}{t_1} \right] c_2 + \frac{1}{t_1'} (1 - r_0 - r_1) c_2 + (r_0 + r_1) \frac{(t_0' + t_1')}{t_1'} \frac{c_2}{t_1}.$$

We can easily verify that B' is increasing in c_2, t'_0 , and t_0 and decreasing in t'_1 .

On the other hand, in case that $w_H = 0$, (24) is equivalent to $B \leq B^O$. Thus, B^S is defined as follows.

$$B^S = \min\left\{B', B^O\right\}.$$

It is clear that S0 is implementable and yields a higher profit than T, if $B \in [B^{\#}, B^{S}]$. The remaining work is to prove that such a range of B does exist.

Claim 1 There exists $\alpha \in (0, t_1)$ such that if $t'_1 < \alpha$, the range of $[B^{\#}, B^S]$ exists; if $t'_1 \ge \alpha$, no range of B exists to support the optimality of choosing the effort-sorting scheme.

First, if $t'_1 = 0$,

$$B^S = \infty \ge \frac{1}{t_1} \frac{(t_0 + t_1)}{t_1} c_2 = B^{\#};$$

Second, if $t'_1 \to t_1$,

$$B^{\#} = \infty > B^S.$$

Third, it can be verified that B^S is decreasing in t'_1 , while $B^{\#}$ is weakly increasing in t'_1 . Thus, we can find such a cutoff of t'_1 described in the claim.

A.9 Proof of Proposition 9

Proof. Assume that under traditional contract T only the Stage-2 IC constraint is binding and under objective IPE contract O1 only the two Stage-2 IC constraints are binding. Define

$$\begin{aligned} R &\equiv (r_0 + r_1); \\ \alpha_H &\equiv R(t_0 + t_1) q + (1 - R) (t'_0 + t'_1) (1 - q), \\ \alpha_L &\equiv R(t_0 + t_1) (1 - q) + (1 - R) (t'_0 + t'_1) q; \\ f(H|1) &\equiv Rq + (1 - R) (1 - q), \\ f(L|1) &\equiv R(1 - q) + (1 - R) q. \end{aligned}$$

In this case, $C^T = (\alpha_H + \alpha_L) b$ where

$$b = \frac{c_2}{Rt_1 + (1 - R)t_1'}.$$

 $C^{O1} = \alpha_H b_H + \alpha_L b_L$ where

$$b_H = \frac{[Rt_1 + (1 - R)t'_1]}{Rqt_1 + (1 - R)(1 - q)t'_1} f(H|1)b,$$

$$b_L = \frac{[Rt_1 + (1 - R)t'_1]}{R(1 - q)t_1 + (1 - R)qt'_1} f(L|1)b.$$

Hence,

$$\begin{split} & C^{T} - C^{O1} \\ = & (\alpha_{H} + \alpha_{L}) \, b - \left(\alpha_{H} \frac{[Rt_{1} + (1 - R) \, t_{1}']}{Rqt_{1} + (1 - R) \, (1 - q) \, t_{1}'} \, f \, (H|1) + \alpha_{L} \frac{[Rt_{1} + (1 - R) \, t_{1}']}{R \, (1 - q) \, t_{1} + (1 - R) \, qt_{1}'} \, f \, (L|1) \right) \, b \\ = & \left[\alpha_{H} \left(1 - \frac{[Rt_{1} + (1 - R) \, t_{1}']}{Rqt_{1} + (1 - R) \, (1 - q) \, t_{1}'} \, f \, (H|1) \right) + \alpha_{L} \left(1 - \frac{[Rt_{1} + (1 - R) \, t_{1}']}{R \, (1 - q) \, t_{1} + (1 - R) \, qt_{1}'} \, f \, (L|1) \right) \right] b \\ = & \left[\alpha_{H} \left(\frac{(1 - R) \, R \, (2q - 1) \, (t_{1} - t_{1}')}{Rqt_{1} + (1 - R) \, (1 - q) \, t_{1}'} \right) + \alpha_{L} \left(- \frac{(1 - R) \, R \, (2q - 1) \, (t_{1} - t_{1}')}{R \, (1 - q) \, t_{1} + (1 - R) \, qt_{1}'} \right) \right] b \\ = & F \times b \end{split}$$

Because $b > 0, C^T - C^{O1} \ge 0$ if and only if

$$F = \alpha_H \left(\frac{(1-R)R(2q-1)(t_1-t_1')}{Rqt_1 + (1-R)(1-q)t_1'} \right) + \alpha_L \left(-\frac{(1-R)R(2q-1)(t_1-t_1')}{R(1-q)t_1 + (1-R)qt_1'} \right) \ge 0.$$
(25)

Substituting α_H and α_L into it,

$$\begin{split} F &= \left[R\left(t_0 + t_1\right)q + \left(1 - R\right)\left(t_0' + t_1'\right)\left(1 - q\right) \right] \frac{\left(1 - R\right)R\left(2q - 1\right)\left(t_1 - t_1'\right)}{Rqt_1 + \left(1 - R\right)\left(1 - q\right)t_1'} \\ &- \left[R\left(t_0 + t_1\right)\left(1 - q\right) + \left(1 - R\right)\left(t_0' + t_1'\right)q \right] \frac{\left(1 - R\right)R\left(2q - 1\right)\left(t_1 - t_1'\right)}{R\left(1 - q\right)t_1 + \left(1 - R\right)qt_1'} \\ &= \left[\left(1 - R\right)R\left(2q - 1\right)\left(t_1 - t_1'\right) \right] \times \\ &\left(\frac{Rqt_0 + \left(1 - R\right)\left(1 - q\right)t_0'}{Rqt_1 + \left(1 - R\right)\left(1 - q\right)t_1'} - \frac{R\left(1 - q\right)t_0 + \left(1 - R\right)qt_0'}{R\left(1 - q\right)t_1 + \left(1 - R\right)qt_1'} \right) \ge 0. \end{split}$$

Given $[(1-R) R (2q-1) (t_1 - t'_1)] \ge 0$, the above inequality is equivalent to

$$\frac{qRt_0 + (1-q)(1-R)t'_0}{(1-q)Rt_0 + q(1-R)t'_0} \ge \frac{qRt_1 + (1-q)(1-R)t'_1}{(1-q)Rt_1 + q(1-R)t'_1}.$$

Thus, in case $t'_0 > 0$ and $t'_1 > 0$, the equation becomes $t_0/t_1 \ge t'_0/t'_1$ for all $q \in (0.5, 1]$ and hence the treatment assuming q = 1 is without loss of generality. In case $t'_0 = t'_1 = 0$, the equation must hold as an equality and hence $C^{O1} = C^T$, which is a result shown in a previous version of this paper (Chen and Chiu 2011).

A.10 Proof of Proposition 10

Proof. First we characterize traditional contract T. Given the inequality (17) hold, only (30) is the binding IC constraint (note that (30) is binding if $c_1/c_2 \geq \frac{r_1[(t_0+t_1)-(t'_0+t'_1)]}{r_0t_1+(1-r_0)t'_1}$). Accordingly, Tsatisfies $w^T = 0$ and

$$b^{T} = \frac{c_{1}}{r_{1} \left[(t_{0} + t_{1}) - (t_{0}' + t_{1}') \right]}.$$

Second, given the inequality (17) hold, under objective IPE contract O1, only IC constraints (6) and (7) are binding and others are non-binding (note that it is the case if $c_1/c_2 \geq \frac{r_1 t_0}{t_1} - \frac{r_1 t'_0}{t'_1}$). As a result b_H and b_L satisfy

$$b_{H}^{O1} = \frac{c_{1}}{r_{1}(t_{0}+t_{1})} + \frac{(t_{0}'+t_{1}')c_{2}}{(t_{0}+t_{1})t_{1}'},$$

$$b_{L}^{O1} = \frac{c_{2}}{t_{1}'}.$$

Third, by using the above formula, we further reckon that

$$\frac{(t'_0+t'_1)}{(t_0+t_1)-(t'_0+t'_1)}\left(b_H^{O1}-b_L^{O1}\right)=b^T-b_H^{O1}.$$

We tend to claim that $b_H^{O1} \ge b_L^{O1}$, iff $b^T \ge b_H^{O1}$. With some rearrangement, we further find that $b^T \ge b_H^{O1}$ iff inequality (17) holds. Thus, if (17) holds, $b^T \ge b_H^{O1} \ge b_L^{O1}$ and obviously $C^T \ge C^{O1}$. The proof is completed.

A.11 Proof of Proposition 11

Proof. The no-feedback contract is solved by the following principal's cost-minimization problem.

$$\min_{w_H, w_L, b_H, b_L} (r_0 + r_1) w_H + (1 - r_0 - r_1) w_L + (r_0 + r_1) (t_0 + t_1) b_H + (1 - r_0 - r_1) (t'_0 + t'_1) b_L,$$

subject to the agent's limited liability contraints, as well as the following three IC constraints:

$$(r_0 + r_1) t_1 b_H + (1 - r_0 - r_1) t_1' b_L \ge c_2;$$
(26)

$$r_1(t_0 + t_1) b_H - r_1(t'_0 + t'_1) b_L + r_1(w_H - w_L) \ge c_1; \text{ and}$$
(27)

$$[r_0 t_1 + r_1 (t_0 + t_1)] b_H + [(1 - r_0) t_1' - r_1 (t_0' + t_1')] b_L + r_1 (w_H - w_L) \ge c_1 + c_2.$$
(28)

The three constraints are to prevent deviations to $\langle 1; 0 \rangle$, to $\langle 0; 1 \rangle$, and to $\langle 0; 0 \rangle$, respectively. The optimal feedback contract, which is just O1, is characterized by the same cost minimization problem except that (26) is replaced by (5) and (6). As (26) is implied by (5) and (6), the implementation cost (denoted by C^{O1^*}) under the optimal no-feedback contract (denoted by $O1^*$) is never greater than C^{O1} .

We argue further that, under some circumstances, C^{O1^*} is strictly lower C^{O1} . Consider the scenario in which in the problem of O1 only the second stage constraints are binding. Therefore, C^{O1} is characterized by $w_H = b_L = 0$ and $b_H = c_2/t_1$ and $b_L = c_2/t_{1'}$. Consider the following candidate contract for O1^*: (1) if $\frac{t_0}{t_1} \ge \frac{t'_0}{t'_1}$, then $w_H = w_L = 0$, $b_H = 0$ and $b_L = \frac{c_2}{(1-r_0-r_1)t'_1}$; (2) if $\frac{t_0}{t_1} < \frac{t'_0}{t'_1}$, then $w_H = w_L = 0$.

It is easy to verify that all of the three constraints (as well as the agent's limited liability contraints) are satsifed; i.e., the candidate contract is indeed feasible. We also note that the implementation cost under this contract is

$$C^{O1^*} = \min\left\{\frac{(t'_0 + t'_1)c_2}{t'_1}, \frac{(t_0 + t_1)c_2}{t_1}\right\}.$$

Clearly, whenever $\frac{t_0}{t_1} \neq \frac{t'_0}{t'_1}, C^{O1^*} < C^{O1} = (r_0 + r_1)(t_0 + t_1)\frac{c_2}{t_1} + (1 - r_0 - r_1)(t'_0 + t'_1)\frac{c_2}{t'_1}$.

Appendix B (not intended for publication):

B.1 Characterization of the optimal traditional contract

We first set up the cost-minimization problem of inducing action plan $\langle x; y \rangle$ without any IPE.

$$\min_{w,b} w + \left[(r_0 + xr_1) \left(t_0 + yt_1 \right) + \left(1 - r_0 - xr_1 \right) \left(t'_0 + yt'_1 \right) \right] b,$$

subject to the following IC constraints.

If y = 1, the IC constraint that prevents deviation to $\langle x; 0 \rangle$ is

$$\left[\left(r_0 + xr_1 \right) t_1 + \left(1 - r_0 - xr_1 \right) t_1' \right] b \ge c_2; \tag{29}$$

if x = 1, the IC constraint that prevents deviation to $\langle 0; y \rangle$ is

$$r_1\left[(t_0 + t_1y) - (t'_0 + t'_1y)\right]b \ge c_1; \tag{30}$$

if x = y = 1, the IC constraint that prevents deviation to $\langle 0; 0 \rangle$ is

$$\{r_1\left[(t_0+t_1)-(t_0'+t_1')\right]+r_0t_1+(1-r_0)t_1'\}b \ge c_1+c_2.$$
(31)

Obviously, in the optimal contract, w = 0. We then analyze the traditional contracts that implement $\langle 1; 1 \rangle$, $\langle 1; 0 \rangle$, $\langle 0; 1 \rangle$ and $\langle 0; 0 \rangle$, respectively, which are denoted by T, T10, T01 and T00.

Implementing $\langle 1;1 \rangle$ We assume that only IC constraint (29) is binding under T. In this case, it requires that

$$\frac{c_1}{c_2} \le \frac{r_1 \left(t_0 - t_0' \right)}{\left[\left(r_0 + r_1 \right) t_1 + \left(1 - r_0 - r_1 \right) t_1' \right]},\tag{32}$$

which is the first part of assumption A1. (The above condition is found when (29) and (31) bind simultaneously.)

Given (32), we find that T satisfies (3) and the implementation cost satisfies

$$C^{T} = \frac{\left[(r_{0} + r_{1}) \left(t_{0} + t_{1} \right) + \left(1 - r_{0} - r_{1} \right) \left(t_{0}' + t_{1}' \right) \right] c_{2}}{\left[(r_{0} + r_{1}) t_{1} + \left(1 - r_{0} - r_{1} \right) t_{1}' \right]}.$$

Implementing $\langle 1; 0 \rangle$ If the principal wants to implement $\langle 1; 0 \rangle$, the IC constraint should guarantee that the agent would not deviate to choose $\langle 0; 0 \rangle$ and $\langle 1; 1 \rangle$, where the former constraint is (30) and the latter constraint is

$$\left[\left(r_0 + r_1 \right) t_1 + \left(1 - r_0 - r_1 \right) t_1' \right] b < c_2.$$

We can verify that these two IC constraints are consistent given (32). Moreover, only (30) is binding. So the optimal choice of b and the implementation cost satisfy

$$\begin{aligned} b^{T10} &= \quad \frac{c_1}{r_1 \left(t_0 - t_0' \right)}, \\ C^{T10} &= \quad \frac{\left[\left(r_0 + r_1 \right) t_0 + \left(1 - r_0 - r_1 \right) t_0' \right] c_1}{r_1 \left(t_0 - t_0' \right)}. \end{aligned}$$

Implementing (0; 1) If the principal wants to implement (0; 1), the IC constraint should guarantee that the agent would not deviate to choose (0; 0) and (1; 1), where the former constraint is (29) and the latter constraint is

$$r_1 \left[(t_0 + t_1) - (t'_0 + t'_1) \right] b < c_1.$$

We can verify that given (32) these two IC constraints are inconsistent, so such action plan is not implementable.

Implementing (0; 0) It is easy to verify that the principal can implement (0; 0) by choosing b = 0 at a cost of 0.

Payoff Comparison Given (32), if T is the optimal traditional contract, it should yield a higher level of profit than T10 and T00, respectively. (Notice T01 is not implementable now.) The two corresponding conditions are $V^T - C^T \ge V^{T10} - C^{T10}$ and $V^T - C^T \ge V^{T00} - 0$, which are equivalent to

$$B \geq \widehat{B}_{1} \equiv \frac{\left[\left(r_{0}+r_{1}\right)\left(t_{0}+t_{1}\right)+\left(1-r_{0}-r_{1}\right)\left(t_{0}'+t_{1}'\right)\right]c_{2}}{\left[\left(r_{0}+r_{1}\right)t_{1}+\left(1-r_{0}-r_{1}\right)t_{1}'\right]^{2}} - \frac{\left[\left(r_{0}+r_{1}\right)t_{0}+\left(1-r_{0}-r_{1}\right)t_{0}'\right]c_{1}}{r_{1}\left(t_{0}-t_{0}'\right)\left[\left(r_{0}+r_{1}\right)t_{1}+\left(1-r_{0}-r_{1}\right)t_{1}'\right]},$$

and

$$B \geq \widehat{B}_{2} \equiv \frac{\left[(r_{0} + r_{1}) \left(t_{0} + t_{1} \right) + \left(1 - r_{0} - r_{1} \right) \left(t_{0}' + t_{1}' \right) \right] c_{2}}{\left[(r_{0} + r_{1}) t_{1} + \left(1 - r_{0} - r_{1} \right) t_{1}' \right]} \\ \times \frac{1}{\left\{ \left[(r_{0} + r_{1}) \left(t_{0} + t_{1} \right) + \left(1 - r_{0} - r_{1} \right) \left(t_{0}' + t_{1}' \right) \right] - \left[r_{0} t_{0} + \left(1 - r_{0} \right) t_{0}' \right] \right\}}$$

We further define $\hat{B} \equiv \max \{\hat{B}_1, \hat{B}_2\}$. Given (32), T is the optimal traditional contract iff $B \ge \hat{B}$. So assumption A1 summarizes these two conditions.

B.2 Setting $w_H = 0$ under objective IPE contract is without loss of generality

Here we show that for any optimal objective IPE contract, introduction of noise will make $w_H = 0$ strictly hold. Consider the same noisy IPE as described in Section 5.1. Then the expected cost of implementing action plan $\langle 1; 1, 1 \rangle$ equals

$$C = (r_0 + r_1) \left[q \left(w_H + (t_0 + t_1) b_H \right) + (1 - q) \left(w_L + (t_0 + t_1) b_L \right) \right] + (1 - r_0 - r_1) \left[(1 - q) \left(w_H + (t'_0 + t'_1) b_H \right) + q \left(w_L + (t'_0 + t'_1) b_L \right) \right].$$

Therefore, we reckon that

$$\frac{dC}{dw_H} = (2q-1)(r_0+r_1) + (1-q) > 0,$$

$$\frac{dC}{db_H} = (r_0+r_1)[q(t_0+t_1) - (1-q)(t'_0+t'_1)] + (1-q)(t'_0+t'_1) > 0.$$

To show that in general using b_H is strictly more cost effective than using w_H in relaxing Stage-1 IC constraint, here we consider the IC constraint that prevents deviation to action plan (0; 1, 1).

$$(r_{0} + r_{1}) \left[q \left(w_{H} + (t_{0} + t_{1}) b_{H} \right) + (1 - q) \left(w_{L} + (t_{0} + t_{1}) b_{L} \right) \right] + (1 - r_{0} - r_{1}) \left[(1 - q) \left(w_{H} + (t'_{0} + t'_{1}) b_{H} \right) + q \left(w_{L} + (t'_{0} + t'_{1}) b_{L} \right) \right] - c_{1} - c_{2} > r_{0} \left[q \left(w_{H} + (t_{0} + t_{1}) b_{H} \right) + (1 - q) \left(w_{L} + (t_{0} + t_{1}) b_{L} \right) \right] + (1 - r_{0}) \left[(1 - q) \left(w_{H} + (t'_{0} + t'_{1}) b_{H} \right) + q \left(w_{L} + (t'_{0} + t'_{1}) b_{L} \right) \right] - c_{2},$$

Rearranging, we have

$$r_1 \left[q \left(w_H + (t_0 + t_1) \, b_H \right) - (1 - q) \left(w_H + (t'_0 + t'_1) \, b_H \right) \right] - c_1$$

> $r_1 \left[q \left(w_L + (t'_0 + t'_1) \, b_L \right) - (1 - q) \left(w_L + (t_0 + t_1) \, b_L \right) \right],$

where w_H and b_H appear only in the LHS. Denote by dIC/dw_H and dIC/db_H the amount the IC constraint is relaxed by one unit increase of w_H and of b_H , respectively. It is straightforward to show that

$$\frac{dIC}{dw_H} = r_1 \left(2q - 1 \right) > 0$$

and

$$\frac{dIC}{db_H} = r_1 \left[q \left(t_0 + t_1 \right) - (1 - q) \left(t'_0 + t'_1 \right) \right] > 0.$$

The cost-benefit ratio of using w_H is

$$\frac{\frac{dC}{dw_H}}{\frac{dIC}{dw_H}} = \frac{r_0 + r_1}{r_1} + \frac{1}{2q - 1} \frac{1 - q}{r_1}$$

and the cost-benefit ratio of using b_H is

$$\begin{aligned} \frac{\frac{dC}{db_H}}{\frac{dIC}{db_H}} &= \frac{r_0 + r_1}{r_1} + \frac{t'_0 + t'_1}{q\left(t_0 + t_1\right) - \left(1 - q\right)\left(t'_0 + t'_1\right)} \frac{1 - q}{r_1} \\ &= \frac{r_0 + r_1}{r_1} + \frac{1}{q\left(\frac{t_0 + t_1}{t'_0 + t'_1} + 1\right) - 1} \frac{1 - q}{r_1} \le \frac{\frac{dC}{dw_H}}{\frac{dIC}{dw_H}}, \end{aligned}$$

where the equality holds if and only if q = 1. Therefore, for $q \in (0.5, 1)$, it is strictly better to use b_H instead of w_H to provide incentive for the agent in the first stage.

B.3 Comparison with the first best

In the first best, given A1, the social planner will implement the action plan (1; 1, 0) if and only if

$$(1 - r_0 - r_1) t_1' B < (1 - r_0 - r_1) c_2.$$

With some rearrangement, the above inequality is equivalent to $B < B^* \equiv \frac{c_2}{t'_1}$.

Objective IPE

We then compare the principal's contract choice under *objective* IPE with the first best. Given A1, there are three contracts the principal may consider: T, O1 and O0. In case $t'_1 < (t'_0/t_0)t_1$, O1 is dominated by T, so the principal will choose either O0 or T; O0 yields a greater profit than T if and only if $B \leq B^O$, where B^O is defined in (22). On the other hand, in case $t'_1 \geq (t'_0/t_0)t_1$, T is dominated by O1, so the principal will choose either O0 or O1, and the threshold of B is replaced by \tilde{B} , which is defined as the minimum B satisfying

$$V^{O1} - V^{O0} \le C^{O1} - C^{O0}.$$
(33)

To finish the welfare analysis, we compare the aforementioned thresholds B^O and \widetilde{B} with B^* .

Claim 2 Suppose that the IPE is objective. Given A1, the principal's threshold of implementing $\langle 1; 1, 0 \rangle$ is greater than B^* .

Proof. There are two cases to consider: (i) $t'_1 \leq \frac{t'_0}{t_0}t_1$ and (ii) $t'_1 > \frac{t'_0}{t_0}t_1$. Case (i). We compare the threshold B^O with B^* . If $\frac{c_1}{c_2} \leq \frac{r_1t_0}{t_1}$,

$$B^{O} = \frac{1}{(1-r_{0}-r_{1})t_{1}'} \left[\frac{(r_{0}+r_{1})t_{0}+(1-r_{0}-r_{1})t_{0}'}{(r_{0}+r_{1})t_{1}+(1-r_{0}-r_{1})t_{1}'}c_{2}-(r_{0}+r_{1})\frac{t_{0}}{t_{1}}c_{2} \right] + \frac{c_{2}}{t_{1}'}$$

$$\geq \frac{c_{2}}{t_{1}'} \equiv B^{*},$$

where the " \geq " is due to the fact that $\frac{t_0}{t_1} \leq \frac{t'_0}{t'_1}$ and hence the term inside the bracket in the RHS of the first line is positive. If $\frac{r_1t_0}{t_1} < \frac{c_1}{c_2} \leq \frac{r_1(t_0-t'_0)}{(r_0+r_1)t_1+(1-r_0-r_1)t'_1}$ (right boundary is imposed by assumption **A1**),

$$\begin{split} B^{O} &= \frac{1}{\left(1 - r_{0} - r_{1}\right)t_{1}'} \left(\frac{\left(r_{0} + r_{1}\right)t_{0} + \left(1 - r_{0} - r_{1}\right)t_{0}'}{\left(r_{0} + r_{1}\right)t_{1} + \left(1 - r_{0} - r_{1}\right)t_{1}'}c_{2} - \left(r_{0} + r_{1}\right)\frac{c_{1}}{r_{1}}\right) + \frac{c_{2}}{t_{1}'}\\ &\geq \frac{1}{\left(1 - r_{0} - r_{1}\right)t_{1}'} \left(\frac{\left(r_{0} + r_{1}\right)t_{0} + \left(1 - r_{0} - r_{1}\right)t_{0}'}{\left(r_{0} + r_{1}\right)t_{1} + \left(1 - r_{0} - r_{1}\right)t_{1}'}c_{2} - \frac{\left(r_{0} + r_{1}\right)\left(t_{0} - t_{0}'\right)}{\left(r_{0} + r_{1}\right)t_{1} + \left(1 - r_{0} - r_{1}\right)t_{1}'}c_{2}\right) + \frac{c_{2}}{t_{1}'}\\ &\geq \frac{c_{2}}{t_{1}'} \equiv B^{*}. \end{split}$$

The first " \geq " is due to $\frac{c_1}{c_2} \leq \frac{r_1(t_0-t'_0)}{(r_0+r_1)t_1+(1-r_0-r_1)t'_1}$, as guaranteed by **A1**.

Case (ii). we compare the threshold \tilde{B} with B^* .

$$\widetilde{B} = \frac{t'_0 + t'_1}{t'_1} \frac{c_2}{t'_1} \ge \frac{c_2}{t'_1} \equiv B^*.$$

 \tilde{B} is calculated by using (33) and assuming that only Stage-2 IC constraints are binding under O1 and O0. If either or both of the Stage-1 IC constraints are binding, the above \tilde{B} is underestimated. This argument is further supported by two facts: (i) the Stage-1 IC constraint under O1 is more likely to be made binding than that under O0 (this feature is discussed in Section 3.2); (ii) if any Stage-1 IC constraint is binding under either O1 or O0, $b_H^{O1} \ge b_H^{O0}$ and hence $C^{O1} - C^{O0} \ge (1 - r_0 - r_1) (t'_0 + t'_1) \frac{c_2}{t'_1}$. Hence we conclude that $\tilde{B} \ge B^*$.

Summarizing all results from above, we establish the claim.

Subjective IPE

We then compare the principal's contract choice under *subjective* IPE with the first best. Given **A1**, there are three contracts the principal may consider: T, S1 and S0. The welfare analysis is conducted under two cases.

Case 1 First, we consider the case where inequality (23) does not hold (i.e., $(1 - r_0 - r_1)(t_0 - t'_0) \leq t_1 + t'_0$). This case is illustrated by Figure 4. According to the proof of Proposition 6, in this case, T always yields a greater profit than S1 for any $t'_1 \in (0, t_1)$. Thus, the principal will choose either T or S0. According to Propositions 7 and 8, $\langle 1; 1, 0 \rangle$ is implementable and chosen by the principal if $t'_1 < \alpha$ and $B^{\#} \leq B \leq B^S$. In order to perform welfare analysis and to characterize panel b of Figure 4, we compare the threshold $B^{\#}$ with B^* , and then compare the threshold B^S with B^* .

Claim 3 Suppose that the IPE is subjective. Assume that c_1/c_2 is sufficiently low.

- (i) There exists a cutoff $\beta \in (0, t_1)$ such that $B^{\#}$ is smaller than B^* if and only if $t'_1 < \beta$.
- (ii) Suppose that $\frac{(r_0+r_1)}{(1-r_0-r_1)} \leq \frac{(r_0+r_1)t_0+(1-r_0-r_1)t'_0}{t_1}$. B^S is greater than B^* , for any $t'_1 \in (0, t_1)$.

Proof. Part (i).We first compare $B^{\#}$ with B^* , where $B^{\#}$ is defined as the minimum B satisfying implementation condition described in Proposition 7. Firstly, if $t'_1 = 0$,

$$B^{\#} = \frac{(t_0 + t_1)}{t_1} \frac{c_2}{t_1} \le \infty = B^*;$$

secondly, if $t'_1 = t_1$,

$$B^{\#} = \infty \ge \frac{c_2}{t_1} = B^*;$$

thirdly, $B^{\#}$ is weakly increasing in t'_1 , while B^* is decreasing in t'_1 . Thus, we can claim that $B^{\#} \leq B^*$ if and only if $t'_1 \leq \beta$.

Part (ii). We then compare B^S with B^* . As defined in the proof of Proposition 8,

$$B^S \equiv \min\left\{B', B^O\right\}.$$

Having shown in the last claim that $B^O \ge B^*$, we then focus on the comparison between B' and B^* . Notice that

$$B' \equiv \frac{c_2}{t_1'} \left[\frac{(r_0 + r_1) t_0 + (1 - r_0 - r_1) t_0'}{(r_0 + r_1) t_1 + (1 - r_0 - r_1) t_1'} - (r_0 + r_1) \frac{t_0}{t_1} + (1 - r_0 - r_1) + (r_0 + r_1) \frac{(t_0' + t_1')}{t_1} \right].$$

We define the term in the bracket as function $\Delta(t'_1)$. It is obvious that $B' \geq B^*$ if and only if $\Delta(t'_1) \geq 1$. We can verify that $\Delta''(t'_1) \geq 0$, so Δ is a convex function of t'_1 . We can further check that

$$\Delta'(t_1')|_{t_1'=t_1} = \left[-(1-r_0-r_1)\frac{(r_0+r_1)t_0 + (1-r_0-r_1)t_0'}{t_1^2} + (r_0+r_1)\frac{1}{t_1} \right] \le 0$$

if and only if

$$\frac{(r_0 + r_1)}{(1 - r_0 - r_1)} \le \frac{(r_0 + r_1)t_0 + (1 - r_0 - r_1)t'_0}{t_1}.$$

So under such a condition, $\Delta'(t'_1) \leq 0$ for all $t'_1 \in (0, t_1)$.

We then reckon that if $t'_1 \longrightarrow t_1$,

$$B' = \frac{c_2}{t_1'} \Delta(t_1') = \frac{c_2}{t_1'} \left(\frac{t_0'}{t_1} + 1\right) \ge \frac{c_2}{t_1'} = B^*$$

Taking all of above into consideration, we claim that if the precondition described in the claim holds, B^S is greater than B^* for all $t'_1 \in (0, t_1)$.

If such a precondition does not hold, the minimum $\Delta(t'_1)$ might be achieved at some middle point $\overline{t'_1} \leq t_1$ with $\overline{t'_1}$ implicitly defined by $\Delta'\left(\overline{t'_1}\right) = 0$. But whether $\Delta\left(\overline{t'_1}\right)$ is smaller or greater than 1 still depends on such parameters as r_0 , r_1 , t_0 , t'_0 , etc. In case that $\Delta\left(\overline{t'_1}\right) < 1$, it is possible that $B^S < B^*$ if $t'_1 \in \left(\widehat{t'_1}, \widehat{t'_1}\right)$, and $B^S \geq B^*$ if $t'_1 \notin \left(\widehat{t'_1}, \widehat{t'_1}\right)$. **Case 2** We next consider the case where inequality (23) holds (i.e., $(1 - r_0 - r_1)(t_0 - t'_0) > t_1 + t'_0$). As shown in the proof of Proposition 6, there exists $\tau \in ((t'_0/t_0)t_1, t_1)$ such that for any $t'_1 \in (\tau, t_1)$, the implementation cost under S1 is lower than under T; for any $t'_1 \in (0, \tau)$, the implementation cost under S1. Thus, for any $t'_1 \in (0, \tau)$, the principal chooses either T or S0, and the previous analysis still applies here; for any $t'_1 \in (\tau, t_1)$, the principal will choose either S0 or S1, and S0 yields a greater profit than S1 if and only if

$$(1 - r_0 - r_1) t'_1 B < C^{S1} - C^{S0}$$

= $(t'_0 + t'_1) \frac{c_2}{t'_1} - (r_0 + r_1) t'_1 B,$

where C^{S1} and C^{S0} are calculated by using Propositions 5 and 7 and assuming that c_1/c_2 is sufficiently low (note that under contract S0, $w_H = t'_1 B - (t'_0 + t'_1) \frac{c_2}{t_1} > 0$ since $t'_1 > (t'_0/t_0)t_1$ and $B \ge B^{\#}$). With some arrangement, the above condition is equivalent to

$$B < \widetilde{B} \equiv \frac{t'_0 + t'_1}{t'_1} \frac{c_2}{t'_1}.$$

To summarize, for any $t'_1 \in (\tau, t_1)$, $\langle 1; 1, 0 \rangle$ is implementable and chosen by the principal if $B^{\#} \leq B \leq \widetilde{B}$, and $\langle 1; 1, 1 \rangle$ is implemented otherwise. (Likewise, there exists $\alpha' \in (\tau, t_1)$ such that the range $\left[B^{\#}, \widetilde{B}\right]$ exists if and only if $t'_1 \leq \alpha'$.) So the characterization of optimal contract in this case is similar to panel a of Figure 4, except that there exists a strip that is close to line $t'_1 = t_1$ and S1 is the optimal contract. We established the relationship between $B^{\#}$ and B^* in the last claim, finding that $\widetilde{B} \geq B^*$. So the welfare analysis is similar to what is shown in panel b of Figure 4.