Input, Output or Mixed Monitoring in Teams?*

Parimal Bag[†] Peng Wang[‡]

March 29, 2019

Abstract

In team problems is it better for the principal to reward players based on their individual efforts or should they be rewarded based on collective performance or even a combination of the two? The answer will depend on the importance of two familiar challenges in team works – coordination and free riding. With perfectly complementary efforts, free-riding incentives are absent, so the principal prefers output monitoring over input monitoring but sometimes both may be dominated by mixed-wage incentives. When efforts are perfect substitutes in production, either of the two polar mechanisms may dominate the other but sometimes mixed-wages may dominate both. For more general technologies only output and input monitoring mechanisms are compared. It is shown that when the team production technology is supermodular, coordination becomes the primary concern and monitoring agents through collective output is mostly the better protocol. On the other hand, if the technology is submodular, output monitoring encourages free riding and so input monitoring may be a more attractive alternative.

***JEL* Classification:** D82; D86; J33; J41; M52.
Key Words: Team, monitoring, free riding, coordination.

* **Acknowledgements:** We thank the associate editor and two anonymous referees of this journal for helpful comments and suggestions. Any remaining mistakes are ours.

† Department of Economics, National University of Singapore, Faculty of Arts and Social Sciences, AS2 Level 6, 1 Arts Link, Singapore 117570; E-mail: ecsbpk@nus.edu.sg

‡ Institute of Economics and Finance, 86 Yu Shan Xi Lu, Nanjing Audit University, Nanjing, Jiangsu, P.R.China; E-mail: melon_wang@outlook.com

1 Introduction

Free riding and coordination are the two key challenges in implementing collective goals in many teams and organizations. For any new project the team performance will depend on the match between project requirements and the team members' individual experience in related past projects. Sometimes even the project participants might not know their types (i.e., productivity or skills) until they start working on the project. With this uncertainty, the principal has to choose ex ante an appropriate monitoring mechanism – whether to monitor individual efforts, team output or both – to maximize his own residual gain. This may sometimes mean prioritizing the goal of team coordination over problems of free riding or vice versa.

In team games, when coordination between different participants' efforts is likely to make a big difference to a project's outcome, controlling moral hazard through individualized input based incentives might not be the best choice. It might be better instead to rely on collective incentives, i.e., team output, to determine team members' rewards. On the other hand, giving incentives based on team output may encourage too much free riding, so sometimes the principal may still turn to individualized incentives.

In order to address the question of optimal monitoring and how the answer might depend on team production technology, we study a two-player team setting where each player can be a high or low productivity type and the types are learnt by the players at the production stage after the principal and the players have signed the contracts. The incentive contracts are linear: wages linear in individual input, team output, or both. Our analysis is carried out, first, with respect to two polar technologies: efforts are perfect complements or perfect substitutes. Perfect substitution technology is the workhorse model for studying voluntary contributions to public goods with applications in teams. Complementarity, on the other hand, is another typical assumption to study team problems. The analysis is then extended to other forms of substitution and complementarity, i.e., submodular and supermodular technologies.[1]

Our main observations are as follows. If the agents' efforts are perfect complements, the principal prefers output monitoring to input monitoring (Proposition 1). With perfectly complementary efforts free-riding incentives are absent. An agent of high productivity type will see his good effort translate into high output provided the other agent is also of high productivity type and chooses similarly good effort. If the other agent is of low productivity type, then the incentive for the high productivity type in putting in good effort will be less if

[1] We consider only input- or output-based wages, and not a combination of the two, under submodular and supermodular technologies.

the rewards are based on total output. But this is no bad an outcome for the principal because if the rewards were based on individual input instead, the high productivity agent would choose effort to maximize his own utility without considering whether that would maximize output given the other agent's effort and the low type. This last response would have damaged the principal's interests. Inducing ideal effort coordination via peer information is possible under output monitoring but not so under input monitoring. Besides principal's profit, output monitoring is shown to dominate input monitoring in terms of ex-ante social surplus generated (Proposition 2). In addition, we show that sometimes a mixed-wage contract linear in both input and output can be better for the principal than both input and output monitoring (Proposition 3). Although input monitoring in isolation fails to achieve the coordination goal, giving a positive weight to input in the mixed-wage incentives can encourage the low-type agent to put in more effort, which in turn induces the high-type agent to increase effort since their efforts are perfect complements.

When efforts are perfect substitutes in production, one might expect free-rider problem to be an important concern under output monitoring. However, because of the linear incentives each agent's effort decision is independent of the other agent's decision, thus neutralizing the free-rider problem. This tends to place output monitoring on par with input monitoring except that the former still suffers due to the moral hazard problem. On the other hand, under output monitoring agents' effort decisions are responsive to their types compared to input monitoring: high-type agent has a higher marginal benefit from an additional effort; under input monitoring agents' effort decisions are independent of their types. On balance, either monitoring may prevail. The exact choice of the incentive mechanism will depend on the distribution of agents' types and the ratio of the two types' (high and low) productivity. Our result will exhibit a *U-shape optimal monitoring* (Proposition 4), when we compare only input and output monitoring; the intuitions are discussed in Section 4. We also show that either mechanism can dominate in terms of ex-ante social surplus, and the result exhibits a similar pattern to profit comparison (Proposition 5). In addition, mixed monitoring can again be favored by the principal due to the different advantages associated each with input and output monitoring (Proposition 6).

To bring back the free-rider problem explicitly, we consider another form of substitution technology – one with diminishing returns to scale. Formally, the team production technology is submodular.[2] For linear incentives, now the agents' efforts are strategic substitutes under output monitoring: an increase in one agent's effort will decrease the effort of the other agent. This tends to pull input monitoring back strongly in contention: often input monitoring will

[2] A technology is submodular (resp. supermodular) if the marginal productivity of an agent's effort is decreasing (increasing) in the other agent's effort.

be the dominant mechanism (Simulation result 1). For this specific submodular technology (see Section 5.1), we provide a clearer link between returns to scale and the extent of free riding/choice of monitoring. Typically, when the degree of strategic substitutability under output monitoring is large, input monitoring often becomes the favored choice.

We also study a Cobb-Douglas production technology which is supermodular. Output monitoring based on linear incentives makes the agents' efforts strategic complements. Intuitively, the agent interactions are cooperative rather than one of free riding. In contrast to the perfectly complementary technology, input monitoring can sometimes be the preferred mechanism. The reason is, unlike in the perfectly complementary case, now by relying on input rather than output monitoring there might not be as sharp a loss from miscoordination if agent types and the corresponding efforts differ. In addition, under input monitoring there is no problem of moral hazard in teams. All in all, under Cobb-Douglas technology the optimal mechanism can either be input or output monitoring (Simulation result 2). In Section 5.2, we provide a sharper prediction relating returns to scale and the extent of coordination/choice of monitoring. What we want to highlight is that when the degree of strategic complementarity under output monitoring increases, output monitoring dominates with increasing frequency.[3]

The coverage of our analysis is thus quite broad: input, output or mixed monitoring for a whole range of technologies. With this reach, we aim to provide some intuitive guidance to the type of forces at play in influencing principal's decisions, by establishing a number of analytical results and sometimes numerical simulations where analytical results are not tractable. The lack of very general conclusions should neither be surprising nor should it be considered a shortcoming of our analysis. To the best of our knowledge, the existing literature on optimal monitoring in teams specialize to either (i) output-based incentives (Holmstrom, 1982), or (ii) output monitoring vs. individualized contributions monitoring but with a restriction to only complementary production technology (McAfee and McMillan, 1991), or (iii) output-based incentives in a sequential production chain with each agent either exerting one unit of effort or shirking (Winter, 2006).

■ **Literature review.** Our work is a follow-up of McAfee and McMillan (1991). The authors consider monitoring in a team setting and shows the equivalence between monitoring individual contributions and team output. For contributions monitoring, the principal is able to measure individual contributions but cannot disentangle effort from ability. In many real-world applications it might be plausible to assume that the principal can only observe

[3]Though we do not analyze mixed monitoring under submodular and supermodular technologies, we believe that mixed monitoring should still sometimes be better than monitoring on any single dimension, due to the different merits of input and output monitoring.

team members' efforts, e.g., the number of hours put in, but not the real contributions. McAfee and McMillan's analysis, in particular their optimal mechanism for output monitoring, cannot be applied to our setting for four reasons: (i) they consider only complementary technology (refer their footnote 9), (ii) their mechanism fails to satisfy limited liability, (iii) their mechanism relies on type reporting, and (iv) they compare output vs. contributions monitoring. In contrast, (i) we consider both supermodular and submodular technologies, (ii) the optimal (linear) contracts in our analysis satisfy limited liability, (iii) we do not make the contracts depend on agent types,[4] and (iv) our comparison is between output and input, rather than output vs. contributions.

An early work analyzing monitoring in a team setting is by Itoh (1991).[5] In a multi-agent, multi-task setting, he shows that under output-based incentives either complete specializations or team-work (with each agent doing more than one task) could be optimal. In an empirical work, Bushman et al. (1995) show that aggregate performance measures perform better than individualized incentives with greater intrafirm interdependencies. More recently Rahman (2012) analyzes the problem of monitoring of shirking by the monitor himself in a group-work environment. We side-step Rahman's issue as the principal himself is the monitor.

Winter (2006) studies a team project with component tasks executed in a pre-determined sequential order by the team members, each exerting zero or one unit of effort. In his core model, the project succeeds if and only if all tasks are successful. Team members are rewarded according to their positions in the production chain but only when the grand project has been successful, i.e., through output monitoring. The author's main interest is how to induce all agents to exert effort at minimal incentive costs.

We also want to distinguish our work from the efficient partnership question that concern with how players in a team can overcome the moral hazard and free-rider problems by designing clever sharing rules (Legros and Matthews, 1993; Nandeibam, 2002). The monitoring mechanisms that we apply are restricted to linear incentives, so the power the principal can exert to control agent incentives are limited. In addition, privacy of (agent) types would make the analysis of general optimal contracts based on output or input extremely hard. Not only we need to solve the agent incentives explicitly, the principal's payoffs under the two monitoring mechanisms have to be calculated in closed forms for comparability. So our goal here is not to offer a general sophisticated mechanism design solution to efficiency

[4]We consider only contracts that depend on team output or agents' input, and not on agent types. Agents learn their types after the contract has been signed. The timing of agents' learning of types is more like in Sappington (1983).

[5]See also Drago and Turnbull (1988).

issues, as the above authors have done, but instead derive an understanding of the simple intuitions behind the two dominant monitoring mechanisms within the special class of linear incentives.

Finally, the question of input or output monitoring has been studied in a principal-agent model by Khalil and Lawarrée (1995). They analyze whether the principal should prefer to be the residual claimant if he could choose which performance measure to monitor – the agent's input or output. Different from them, ours is not an adverse selection model and the challenge of monitoring teams, instead of a single agent, calls for a very different type of analysis. We also consider only linear incentives (as opposed to more general incentives) and, by default, our principal is the residual claimant.

The remainder of the paper is organized as follows. Starting with the formal model in Section 2, we analyze various technologies in Sections 3–5. In Section 6 we discuss two variants of the main model – type-dependent contracts and private information of agent types. A Supplementary file contains additional derivations, and supporting Matlab and Mathematica commands.

2 Model

A principal hires two agents, indexed by $j = 1, 2$, to work in a team on a joint project. Both the principal and the agents are risk neutral. Each agent can be of low or high ability type with productivity parameter $\theta_j \in \{\theta_L, \theta_H\}$, where $\theta_H > \theta_L > 0$. It is common knowledge that the probability that an agent is of low ability type θ_L is p, where $0 < p < 1$. Both the principal and the agents know the distribution of types before signing the contract. Each agent learns about his own type as well as the other agent's type only after contracting with the principal, during the project's implementation phase. The principal does not observe any of this information.

Each agent j exerts an effort level $e_j \in \mathfrak{R}_+$. The agents face the same convex effort cost $C(e_j) = d \cdot \frac{e_j^2}{2}$, $d > 0$, which is known by the principal. The cost functions are known to the principal. Agents' role in production $y = f(\theta_1, \theta_2, e_1, e_2)$ are symmetric, i.e., y remains the same after switching the agent indices.[6]

Depending on the nature of the project, the production function takes different forms. We analyze separately the cases where agents' efforts are perfect complements or perfect substitutes, and two other technologies, submodular and supermodular.

[6]Output can be allowed to be stochastic, e.g., $y = \epsilon f(\theta_1, \theta_2, e_1, e_2)$, where random variable ϵ is independent of θ_j and e_j, with mean 1. Since all parties are risk neutral, our results will not change qualitatively.

The particular timing, that the agents receive private information after contracting but before choosing actions, is similar to the one in Sappington (1983). Given this timing and the application we have in mind, it might not be realistic for the principal to write a sophisticated menu contract that depends on agent types. This could be because agents learn their type profile gradually as they work through the various parts of the project and adjust their overall efforts as and when necessary. Despite the static or one-shot presentation of agent interactions, the agents work in a span of calendar time which offers them sufficient scope to learn about their respective skill-suitability for the work being undertaken: they see or talk to each other from time to time about the ongoing team tasks.[7] We therefore look at mechanisms that do not require type reporting.[8]

Three main contract forms are compared. Under *output monitoring*, agents' wages are functions of only team output. Under *input monitoring*, principal is able to observe the agents' efforts and can specify wages based on individual efforts. A more powerful mechanism specifying wages based on joint efforts, while possible, is not very plausible for realistic applications. Under *mixed monitoring*, principal can observe individual efforts as well as the team output, thus can design wage on both dimensions.[9]

Both agents are offered uniform, non-discriminatory contracts. All contracts must satisfy agent limited liabilities – wages cannot be negative.

The timing of the game is as follows.

1. Each agent first decides whether or not to accept a contract offered by the principal.

2. On participation, each agent learns his own type as well as his partner's type.

3. Then they choose effort levels simultaneously.

4. The output is then realized and the payments are given accordingly.

■ Examples of different technologies and monitoring

[7] When a team of researchers start working on an R&D project for example, it takes a while of trial-and-error before they truly learn their knack for the various tasks. The effort choices are therefore over the full project span during which time the agents will be able to adjust efforts according to their types. This description is plausible even when production can be split into multiple time slots so long as the agents choose their respective efforts without observing other agents' effort decisions. That is, the efforts over multiple rounds are carried out under the principal-chosen veil of secrecy, denying agents to make strategic commitments to low efforts in the early stages in a bid to free ride. See Mohnen et al. (2008) or Bag and Pepito (2012), where the authors consider two-stage team production by two players under secrecy as strategically equivalent to a one-shot game.

[8] Herold (2010) explains why relying on high-powered incentive contract can be a signal of distrust and may thus dampen an agent's effort incentive.

[9] Note that when we compare different monitoring mechanisms, we do not consider monitoring costs, though in reality monitoring both dimensions should be more costly than monitoring any single one.

Consider the problem of a multinational company: the parent division wants to open a branch in a foreign country. Its success depends on the performance by the company's top management in two dimensions. One of the personnel in charge of the foreign division must secure a license from the host government by greasing the wheels of the bureaucracy. Also, knowing the idiosyncratic tastes of the foreign customers is important. The first task – successful lobbying with the key bureaucrats – does not come naturally to all managers. The second job of having a good sense of the market is also critical to the project's success. Failing in either of the two dimensions could seriously dampen the project's chance of success or profitability. Loosely speaking, the project exhibits a perfectly complementary technology. Incentivizing managers in charge of the foreign division based on the new initiative's success is a good example of output monitoring. Alternatively, the managers' remuneration could be based on how long they are willing to stay in their foreign posts away from their comfort zones at home – an example of input monitoring. Mixed monitoring incentives could take into account the duration of their assignment away from home and the project's profitability.

Examples of both perfect and imperfect substitution team technologies where monitoring could be an issue are plenty. Both perfect and imperfect substitutions feature dominantly in the workhorse models of voluntary contribution to public goods that is extensively applied to team problems; see, for example, Bergstron, Blume and Varian (1986) and Marx and Matthews (2000) on various strategic aspects of public good contributions, and the classic work by Holmstronm (1982) for an application of the problem of free riding and monitoring in teams. Supermodularity (submodularity) means one team member's marginal productivity of effort is increasing (respectively, decreasing) in another team member's effort. The canonical monitoring in voluntary contribution to public goods relies on aggregate output, although one can also consider rewarding team members according to individual efforts where the efforts are verifiable.[10]

3 Perfect complements

3.1 Wage linear in either input or output

In this section, we analyze the case of perfectly complementary technology, $y = \min\{\theta_1 e_1, \theta_2 e_2\}$, and compare input and output monitoring mechanisms. We will restrict to only linear con-

[10] Again, verifiability of efforts may be plausible if team members' task assignments require fixed time commitments and the team members choose how many tasks of a grand project they are willing to take on. We can interpret the multinational company's foreign operations to cover both perfect complementarity, supermodular technology and submodular technology (see sections 3–5 for the formal definitions), depending on the synergies between the two personnel's skills and efforts.

tracts here and in the rest of the paper. Linear incentives are quite relevant due to their extensive use, for instance piece-rate contracts and hourly wage rates in employment; Itoh (1991), Nandeibam (2002), Mohnen et al. (2008), all use some form of linear contracts.[11]

Since the fixed part of the wage does not affect marginal effort incentives, and limited liability restriction rules out negative wages, the principal would economize on effort implementation costs by setting the fixed wage component to zero.

■ **Input monitoring.** Suppose the principal offers wages $W^{in} = \alpha_{in} e_j^{in}$, $\alpha_{in} > 0$ to agent j for his effort e_j^{in}. Since agents' types affect only the team productivity and not the utility functions, their payoff functions are identical and independent of realized types. Agent j's payoff function is $\pi_j^{in} = \alpha_{in} e_j^{in} - d \cdot \frac{(e_j^{in})^2}{2}$, $j = 1, 2$, so he will choose effort $e_j^{in} = \frac{\alpha_{in}}{d}$.

The expected profit function for the principal is

$$\begin{aligned} E[\pi_p^{in}] &= p^2(\theta_L \frac{\alpha_{in}}{d}) + 2p(1-p)(\theta_L \frac{\alpha_{in}}{d}) + (1-p)^2(\theta_H \frac{\alpha_{in}}{d}) - 2\alpha_{in} \cdot \frac{\alpha_{in}}{d} \\ &= [p(2-p)\theta_L + (1-p)^2\theta_H] \frac{\alpha_{in}}{d} - \frac{2\alpha_{in}^2}{d}. \end{aligned}$$

The principal would choose α_{in} to satisfy

$$\begin{aligned} \frac{\partial E[\pi_p^{in}]}{\partial \alpha_{in}} &= \frac{1}{d}[p(2-p)\theta_L + (1-p)^2\theta_H - 4\alpha_{in}] = 0, \\ \text{i.e.,} \quad \alpha_{in} &= \frac{p(2-p)\theta_L + (1-p)^2\theta_H}{4}. \end{aligned}$$

Therefore, the agents' equilibrium efforts are

$$e_H^{in} = e_L^{in} = \frac{p(2-p)\theta_L + (1-p)^2\theta_H}{4d}$$

giving rise to the expected profit for the principal,

$$E[\pi_p^{in}] = \frac{[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{8d}. \tag{1}$$

■ **Output monitoring.** Now, suppose the principal offers wages based on output: $W^{out} =$

[11] Admittedly linear contracts will involve some loss of efficiency relative to more general non-linear contracts. Besides its extensive use, we rely on linear incentives to make the analysis tractable – to be able to solve the optimal wages explicitly in order to derive the principal's payoffs in closed form for comparison between output monitoring, input monitoring and mixed monitoring.

$\alpha_{out} y$, $\alpha_{out} > 0$. Agent j's payoff from choosing effort e_j^{out} while agent k chooses e_k^{out} is

$$\pi_j^{out} = \alpha_{out} \min\{\theta_j e_j^{out}, \theta_k e_k^{out}\} - d \cdot \frac{(e_j^{out})^2}{2}, \qquad j, k = 1, 2, \quad k \neq j.$$

Since the agents know each other's types, they will respond differently according to the type profile. We need to analyze two cases:

1. $\theta_j = \theta_k$. Then $\pi_j^{out} = \alpha_{out} \theta_j \min\{e_j^{out}, e_k^{out}\} - d \cdot \frac{(e_j^{out})^2}{2}$. Agent j's best response is

$$e_j^{out} = \begin{cases} \frac{\alpha_{out}\theta_j}{d}, & \text{if } e_k^{out} \geq \frac{\alpha_{out}\theta_j}{d} \\ e_k^{out}, & \text{if } e_k^{out} < \frac{\alpha_{out}\theta_j}{d}. \end{cases}$$

By symmetry, agent k's best response is similar. Thus, there is a continuum of Nash equilibria of the simultaneous move efforts game: $(e_j^{out}, e_k^{out}) = (e^*, e^*)$ where $e^* \in [0, \frac{\alpha_{out}\theta_j}{d}]$.

2. $\theta_j \neq \theta_k$. Thus, agent j's best response is

$$e_j^{out} = \begin{cases} \frac{\alpha_{out}\theta_j}{d}, & \text{if } e_k^{out} \geq \frac{\alpha_{out}\theta_j^2}{d\theta_k} \\ \frac{\theta_k}{\theta_j} e_k^{out}, & \text{if } e_k^{out} < \frac{\alpha_{out}\theta_j^2}{d\theta_k}. \end{cases}$$

By symmetry, agent k's best response is similar. We can see that the agents should choose effort levels such that $\theta_j e_j^{out} = \theta_k e_k^{out}$. Thus, the high-type agent's effort will be restricted by the low type's effort. Without loss of generality, assume $\theta_j = \theta_L$ and $\theta_k = \theta_H$. The Nash equilibria are: $(e_j^{out}, e_k^{out}) = (e^*, \frac{\theta_L}{\theta_H} e^*)$ where $e^* \in [0, \frac{\alpha_{out}\theta_L}{d}]$.

Since there are multiple equilibria, we assume that for any given α_{out}, the agents always choose the equilibrium that maximizes their payoffs. Thus, in the symmetric case, the agents will choose $e_j^{out} = e_k^{out} = \frac{\alpha_{out}\theta_j}{d}$. In the asymmetric case, effort levels are $e_L^{out} = \frac{\alpha_{out}\theta_L}{d}$ for the low-type agent and $e_H^{out} = \frac{\alpha_{out}\theta_L^2}{d\theta_H}$ for the high-type agent.

Then, the principal's expected profit function can be written as:

$$\begin{aligned} E[\pi_p^{out}] &= [1 - (1-p)^2][\theta_L^2 \frac{\alpha_{out}}{d}(1 - 2\alpha_{out})] + (1-p)^2[\theta_H^2 \frac{\alpha_{out}}{d}(1 - 2\alpha_{out})] \\ &= \frac{\alpha_{out}}{d}(1 - 2\alpha_{out})[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2]. \end{aligned}$$

The principal will choose α_{out} such that

$$\frac{\partial E[\pi_p^{out}]}{\partial \alpha_{out}} = \frac{1}{d}[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2](1 - 4\alpha_{out}) = 0, \quad \text{i.e.,} \quad \alpha_{out} = \frac{1}{4}.$$

If both agents are low types, the equilibrium effort of each agent is $e_L^{out} = \frac{\theta_L}{4d}$; if both are high types, the effort level for each agent is $e_H^{out} = \frac{\theta_H}{4d}$; if the agents are of different types, the effort levels are $e_L^{out} = \frac{\theta_L}{4d}$, $e_H^{out} = \frac{\theta_L^2}{4\theta_H d}$. The principal's expected profit can now be calculated as:

$$E[\pi_p^{out}] = \frac{p(2-p)\theta_L^2 + (1-p)^2\theta_H^2}{8d}. \tag{2}$$

■ **Comparison.** For the input-based pay, since the effort costs are the same for the two agents regardless of their types, they should exert the same level of efforts. This means when only one of them has higher productivity, part of his effort will be wasted. For the output-based pay, when the two agents are of different types, the low-type agent's effort becomes the key determinant of output with high type lowering his effort appropriately to coordinate with the low-type's effort. This saves on high type's effort costs, keeping the principal's incentive costs down and ex-ante profits high relative to input monitoring.

Proposition 1. *Suppose efforts are perfect complements. Then output monitoring is better than input monitoring.*

Overall, output monitoring outperforms input monitoring by tailoring agents' efforts to their respective productivity. The agents coordinate, according to their types, under output monitoring but not so under input monitoring.

■ **Incentive slopes.** An important point to take away so far is in the different approaches to incentive provision. Under output monitoring agent's effort choice is responsive to his marginal contribution, whereas under input monitoring effort decision depends only on marginal reward-marginal (effort) cost comparison.[12] Thus, under input monitoring the principal will have to align agent's equilibrium effort with his own objective which is profit, and considerations of profit must take into account the agents' expected productivity. So, the resulting α_{in} is the principal's ex-ante balancing act in aligning agents' interests with his own goal. In contrast, under output monitoring the principal lets the agents choose at what intensity they should work, depending on their realized types. Thus the resulting incentive slope α_{out} is kept independent of types as eventually the agents will internalize according to their types.

3.2 Social surplus

Now we compare the ex-ante social surplus from input monitoring with output monitoring to see the distortions from the efficient level. The ex-ante social surplus is simply the principal's expected profit plus the two agents' payoffs aggregated according to the distribution of the

[12] An agent's type influences only the productivity and not the effort costs.

type profile. Let V^{in} be the ex-ante social surplus under input monitoring and V^{out} under output monitoring.

Under input monitoring, agents' payoffs are

$$\pi_{LL}^{in} = \pi_{HH}^{in} = \pi_{LH}^{in} = \pi_{HL}^{in} = \frac{[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{32d},$$

and we know that the principal's expected profit is shown in equation (1). Thus, the ex-ante social surplus is

$$\begin{aligned} V^{in} &= \frac{[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{8d} + 2 \times \frac{[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{32d} \\ &= \frac{3[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{16d}. \end{aligned}$$

Under output monitoring, agents' payoffs are

$$\pi_{LL}^{out} = \frac{\theta_L^2}{32d}, \quad \pi_{HH}^{out} = \frac{\theta_H^2}{32d}, \quad \pi_{LH}^{out} = \frac{\theta_L^2}{32d}, \quad \pi_{HL}^{out} = \frac{2\theta_H^2\theta_L^2 - \theta_L^4}{32d\theta_H^2},$$

and we know that the principal's expected profit is shown in equation (2). Thus, the ex-ante social surplus is

$$\begin{aligned} V^{out} &= \frac{p(2-p)\theta_L^2 + (1-p)^2\theta_H^2}{8d} + 2p^2 \times \frac{\theta_L^2}{32d} + 2(1-p)^2 \times \frac{\theta_H^2}{32d} + 2p(1-p)\left[\frac{\theta_L^2}{32d} + \frac{2\theta_H^2\theta_L^2 - \theta_L^4}{32d\theta_H^2}\right] \\ &= \frac{3(1-p)^2\theta_H^4 + p(7-4p)\theta_L^2\theta_H^2 - p(1-p)\theta_L^4}{16d\theta_H^2}. \end{aligned}$$

Therefore,

$$V^{out} - V^{in} = \frac{p(1-p)\left[3(1-p)(2-p)\theta_H^4 - 6(1-p)(2-p)\theta_L\theta_H^3 + (7-9p+3p^2)\theta_L^2\theta_H^2 - \theta_L^4\right]}{16d\theta_H^2}.$$

Proposition 2. *Suppose efforts are perfect complements. Then the ex-ante social surplus under output monitoring is always higher than that under input monitoring.*

Thus, output monitoring not only generates higher expected profit for the principal but is also more efficient than input monitoring. Basically through coordination of agents' efforts, output monitoring aligns the principal's interest with that of the planner under perfect complementary technology. The alignment is not perfect, however, because the principal will not know in advance the agents' type realizations. There will still be inefficiency because the principal will have to balance the probabilities of various agent type combinations, with one incentive coping with all.[13]

3.3 Wage linear in both input and output

Instead of relying exclusively on input or output, we now consider mixed wages that are linear in input and output, i.e., $W = \alpha_{in} e_j + \alpha_{out} y$, $\alpha_{in}, \alpha_{out} > 0$. Agent j's payoff from choosing effort e_j while agent k chooses e_k is

$$\pi_j = \alpha_{in} e_j + \alpha_{out} \min\{\theta_j e_j, \theta_k e_k\} - d \cdot \frac{(e_j)^2}{2}, \qquad j, k = 1, 2, \quad k \neq j.$$

Since the agents know each other's types, they will respond differently according to the type profile. We need to analyze two cases:

1. $\theta_j = \theta_k$. Then $\pi_j^{out} = \alpha_{in} e_j + \alpha_{out} \theta_j \min\{e_j, e_k\} - d \cdot \frac{(e_j)^2}{2}$. Agent j's best response is

$$e_j = \begin{cases} \frac{\alpha_{in} + \alpha_{out}\theta_j}{d}, & \text{if } e_k \geq \frac{\alpha_{in} + \alpha_{out}\theta_j}{d} \\ e_k, & \text{if } e_k < \frac{\alpha_{in} + \alpha_{out}\theta_j}{d}. \end{cases}$$

By symmetry, agent k's best response is similar. Thus, there is a continuum of Nash equilibria: $(e_j, e_k) = (e^*, e^*)$ where $e^* \in [0, \frac{\alpha_{in} + \alpha_{out}\theta_j}{d}]$. We will choose the best equilibrium from the agents' point of view similar to the one under output monitoring: $e^* = \frac{\alpha_{in} + \alpha_{out}\theta_j}{d}$.

2. $\theta_j \neq \theta_k$. This case is more complicated than case (2) under output monitoring. Because agents are incentivized not only by total output but also individual efforts, there is no special reason to always coordinate one's effort with the other agent's effort in the production and avoid wastage of efforts. To see this, first, it can be verified that if $e_k \geq \frac{(\alpha_{in} + \alpha_{out}\theta_j)\theta_j}{d\theta_k}$, then the best response for agent j is $e_j = \frac{\alpha_{in} + \alpha_{out}\theta_j}{d}$ for which k's effort does not bind in production and thus involves no wastage of effort by j.

However, if $e_k < \frac{(\alpha_{in} + \alpha_{out}\theta_j)\theta_j}{d\theta_k}$, agent k's effort e_k might or might not matter in how agent j chooses her effort. We consider the following two scenarios.

Scenario 1: no wastage of effort (by the high-type agent).

In this scenario, the agents choose effort levels such that $\theta_j e_j = \theta_k e_k$. Without loss of generality, assume $\theta_j = \theta_L$ and $\theta_k = \theta_H$. The Nash equilibria are: $(e_j, e_k) = (e^*, \frac{\theta_L}{\theta_H} e^*)$ where $e^* \in [0, \frac{\alpha_{in} + \alpha_{out}\theta_L}{d}]$. Again, we will choose the best equilibrium efforts for the agents. Thus, effort levels are $e_L = \frac{\alpha_{in} + \alpha_{out}\theta_L}{d}$ for the low-type agent and $e_H = \frac{(\alpha_{in} + \alpha_{out}\theta_L)\theta_L}{d\theta_H}$ for the high-type agent.

Scenario 2: wastage of effort (by the high-type agent).

[13] The social planner's problem is included in the Supplementary materials for all the four technologies considered.

The agents will choose effort levels such that $\theta_H e_H > \theta_L e_L$ provided the incentive on effort, α'_{in}, is relatively large.[14] Following our equilibrium selection criterion as before, we will choose $e_L = \frac{\alpha'_{in}+\alpha'_{out}\theta_L}{d}$ and $e_H = \frac{\alpha'_{in}}{d}$ as the equilibrium efforts, and the condition required to ensure wastage of effort is $\alpha'_{out} < \frac{\theta_H-\theta_L}{\theta_L^2}\alpha'_{in}$.

Therefore, if we assume that there is no wastage of effort when agents' types are different, then the principal's ex-ante profit function can be written as:

$$
\begin{aligned}
E[\pi_p] = \frac{1}{d} \quad \{ & \quad [p(2-p)\theta_L + (1-p)^2\theta_H]\alpha_{in} - 2[1-p+p^2+p(1-p)\frac{\theta_L}{\theta_H}]\alpha_{in}^2 \\
+ & \quad [p(2-p)\theta_L^2 + (1-p)^2\theta_H^2]\alpha_{out}(1-2\alpha_{out}) \\
- & \quad [2p(3-p)\theta_L + 4(1-p)^2\theta_H + 2p(1-p)\frac{\theta_L^2}{\theta_H}]\alpha_{in}\alpha_{out}\}.
\end{aligned}
$$

The principal will choose α_{in} and α_{out} such that

$$
\begin{aligned}
\frac{\partial E[\pi_p]}{\partial \alpha_{in}} = \frac{1}{d} \quad \{ & \quad [p(2-p)\theta_L + (1-p)^2\theta_H] - 4[1-p+p^2+p(1-p)\frac{\theta_L}{\theta_H}]\alpha_{in} \\
- & \quad 2[p(3-p)\theta_L + 2(1-p)^2\theta_H + p(1-p)\frac{\theta_L^2}{\theta_H}]\alpha_{out}\} = 0, \\
\frac{\partial E[\pi_p]}{\partial \alpha_{out}} = \frac{1}{d} \quad \{ & \quad [p(2-p)\theta_L^2 + (1-p)^2\theta_H^2](1-4\alpha_{out}) \\
- & \quad 2[p(3-p)\theta_L + 2(1-p)^2\theta_H + p(1-p)\frac{\theta_L^2}{\theta_H}]\alpha_{in}\} = 0.
\end{aligned}
$$

Solving the above two first-order conditions we obtain:

$$
\alpha_{in} = \frac{[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2]p(1-p)\theta_L(1-\frac{\theta_L}{\theta_H})}{8[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2][1-p+p^2+p(1-p)\frac{\theta_L}{\theta_H}] - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H + p(1-p)\frac{\theta_L^2}{\theta_H}]^2}
$$

and

$$
\alpha_{out} = \frac{2[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2][1-p+p^2+p(1-p)\frac{\theta_L}{\theta_H}] - [p(3-p)\theta_L + 2(1-p)^2\theta_H + p(1-p)\frac{\theta_L^2}{\theta_H}][p(2-p)\theta_L + (1-p)^2\theta_H]}{8[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2][1-p+p^2+p(1-p)\frac{\theta_L}{\theta_H}] - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H + p(1-p)\frac{\theta_L^2}{\theta_H}]^2},
$$

and the parameter values need to satisfy

$$
\alpha_{in} > 0 \quad \text{and} \quad \alpha_{out} > 0. \tag{3}
$$

The principal's profit can finally be derived by substituting the expressions of α_{in} and

[14] We use α'_{in} and α'_{out} to represent the incentives in scenario 2 to distinguish with those used in scenario 1. Also, $E[\pi'_p]$ will be used to represent principal's expected profit in scenario 2.

α_{out} back to the expression of $E[\pi_p]$.

On the other hand, if there is wastage of effort by the high-type agent when his partner is low type, then the principal's expected profit function can be written as:

$$\begin{aligned} E[\pi'_p] = \frac{1}{d} \quad \{ & \quad [p(2-p)\theta_L + (1-p)^2\theta_H]\alpha'_{in} - 2(\alpha'_{in})^2 + [p(2-p)\theta_L^2 + (1-p)^2\theta_H^2]\alpha'_{out}(1-2\alpha'_{out}) \\ - & \quad [2p(3-p)\theta_L + 4(1-p)^2\theta_H]\alpha'_{in}\alpha'_{out}\}. \end{aligned}$$

The principal will choose α'_{in} and α'_{out} such that

$$\begin{aligned} \frac{\partial E[\pi'_p]}{\partial \alpha'_{in}} &= \frac{1}{d}\{[p(2-p)\theta_L + (1-p)^2\theta_H] - 4\alpha'_{in} - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H]\alpha'_{out}\} = 0, \\ \frac{\partial E[\pi'_p]}{\partial \alpha'_{out}} &= \frac{1}{d}\{[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2](1-4\alpha'_{out}) - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H]\alpha'_{in}\} = 0. \end{aligned}$$

Solving the above two first-order conditions, we obtain:

$$\begin{aligned} \alpha'_{in} &= \frac{[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2]p(1-p)\theta_L}{8[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2] - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H]^2}, \\ \alpha'_{out} &= \frac{2[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2] - [p(3-p)\theta_L + 2(1-p)^2\theta_H][p(2-p)\theta_L + (1-p)^2\theta_H]}{8[p(2-p)\theta_L^2 + (1-p)^2\theta_H^2] - 2[p(3-p)\theta_L + 2(1-p)^2\theta_H]^2}, \end{aligned}$$

with the parameter restrictions additionally imposed such that

$$\alpha'_{in} > 0, \quad \alpha'_{out} > 0 \quad \text{and} \quad \alpha'_{out} < \frac{\theta_H - \theta_L}{\theta_L^2}\alpha'_{in}. \tag{4}$$

The principal's profit can then be derived by substituting the expressions of α_{in} and α_{out} back to the expression of $E[\pi'_p]$.

Whether the optimal mixed wage involves wastage of effort or not depends on whether $E[\pi_p]$ or $E[\pi'_p]$ is bigger, when the parameter values satisfy the conditions (3) and (4). Since the expressions are complicated, characterizing conditions under which $E[\pi_p]$ or $E[\pi'_p]$ is bigger is intractable. Based on many rounds of simulations, $E[\pi_p]$ always turns out to be higher than $E[\pi'_p]$, i.e., optimal mixed wage should not involve wastage of effort whenever both conditions (3) and (4) hold.[15] If conditions (3) and (4) cannot be satisfied at the same time, then one of the scenarios will prevail as the mixed monitoring equilibrium according to the respective parameter values.

■ Comparison with output monitoring

[15] Sample graphs are provided in the Supplementary file.

Earlier in Proposition 1 we have shown that output monitoring dominates input monitoring. It is natural to ask whether mixed monitoring can outperform output monitoring.

Proposition 3. *Suppose efforts are perfect complements. Then mixed monitoring dominates output monitoring when either condition (3) or (4) is satisfied.*

The proof is simple. If either condition (3) or (4) holds, then the principal's maximization problem does not have corner solution. Thus, mixed wage performs better than both output-based wage and input-based wage.

One possible reason why mixed monitoring may dominate output monitoring is that in the latter the high-type agent's effort is going to be constrained by the low-type agent's effort, if the two agents happen to be of different types. When the two agents are very likely to be of different types and their productivity differs sufficiently, giving a positive weight to input in the determination of wages would make the low-type agent to increase her effort that, in turn, would induce the high-type agent to increase her effort as well. This is likely to improve the overall profitability of the principal.

4 Perfect substitutes

4.1 Wage linear in either input or output

We next consider monitoring incentives for another important class, the perfect substitution technology: $y = \theta_1 e_1 + \theta_2 e_2$. This technology differs from another variant, $y = f(\theta_1 e_1 + \theta_2 e_2)$, where $f'(.) > 0$ and $f''(.) < 0$, that will be discussed in Section 5.1.

■ **Input monitoring.** Suppose the principal offers wages $W^{in} = \alpha_{in} e_j^{in}$, $\alpha_{in} > 0$. Agent j will respond by choosing an effort $e_j^{in} = \frac{\alpha_{in}}{d}$. The principal's expected profit function then is given by:

$$\begin{aligned} E[\pi_p^{in}] &= p^2(2\theta_L)\frac{\alpha_{in}}{d} + 2p(1-p)(\theta_L+\theta_H)\frac{\alpha_{in}}{d} + (1-p)^2(2\theta_H)\frac{\alpha_{in}}{d} - 2\frac{\alpha_{in}^2}{d} \\ &= 2[p\theta_L + (1-p)\theta_H]\frac{\alpha_{in}}{d} - 2\frac{\alpha_{in}^2}{d}. \end{aligned}$$

The principal will now choose α_{in} such that

$$\begin{aligned} \frac{\partial E[\pi_p^{in}]}{\partial \alpha_{in}} &= \frac{2}{d}[p\theta_L + (1-p)\theta_H - 2\alpha_{in}] = 0, \\ \text{i.e.,}\ \ \alpha_{in} &= \frac{p\theta_L + (1-p)\theta_H}{2}. \end{aligned}$$

So the agents' equilibrium efforts are $e_H^{in} = e_L^{in} = \frac{p\theta_L+(1-p)\theta_H}{2d}$, giving rise to the following expected profit for the principal:

$$E[\pi_p^{in}] = \frac{[p\theta_L + (1-p)\theta_H]^2}{2d}. \tag{5}$$

■ **Output monitoring.** Now, suppose the principal offers wages based on output: $W^{out} = \alpha_{out} y$, $\alpha_{out} > 0$. Agent j's payoff function is

$$\pi_j^{out} = \alpha_{out}(\theta_j e_j^{out} + \theta_k e_k^{out}) - d \cdot \frac{(e_j^{out})^2}{2}, \qquad j, k = 1, 2, \quad k \neq j,$$

and his best-response is $e_j^{out} = \frac{\alpha_{out}\theta_j}{d}$. Thus, the agents' efforts are *independent* of each other, and the resulting Nash equilibrium will be unique. The neutrality result is due to linear wages.

The principal's expected profit function is

$$\begin{aligned} E[\pi_p^{out}] =& p^2[2\theta_L^2 \frac{\alpha_{out}}{d}(1-2\alpha_{out})] + 2p(1-p)[(\theta_L^2+\theta_H^2)\frac{\alpha_{out}}{d}(1-2\alpha_{out})] + (1-p)^2[2\theta_H^2 \frac{\alpha_{out}}{d}(1-2\alpha_{out})] \\ =& 2\frac{\alpha_{out}}{d}(1-2\alpha_{out})[p\theta_L^2 + (1-p)\theta_H^2]. \end{aligned}$$

So the principal chooses α_{out} such that

$$\frac{\partial E[\pi_p^{out}]}{\partial \alpha_{out}} = \frac{2}{d}[p\theta_L^2 + (1-p)\theta_H^2](1-4\alpha_{out}) = 0, \quad \text{i.e.,} \quad \alpha_{out} = \frac{1}{4}.$$

The agents' equilibrium effort levels, $e_H^{out} = \frac{\theta_H}{4d}$, $e_L^{out} = \frac{\theta_L}{4d}$, lead to the following expected profit for the principal:

$$E[\pi_p^{out}] = \frac{p\theta_L^2 + (1-p)\theta_H^2}{4d}. \tag{6}$$

■ **Comparison.** Ideally the principal would like the more productive agent to put in more effort: it benefits him in terms of team output. Output monitoring incentivizes agents better with effort proportional to the productivity parameter θ, while under input monitoring both types would put in identical efforts. This, however, is not everything. Under output monitoring the principal has to cope with the (team) moral hazard problem, which is not encountered in input monitoring: output monitoring allows the low-type agent to enjoy a higher payoff by putting in less effort than the high-type agent because agents' efforts cannot be directly observed. Input monitoring addresses this problem of the low type "riding" on the high-type agent under output monitoring.[16] Ultimately, the choice over the two monitoring

[16] Although we cannot strictly talk about 'free riding' under output monitoring as the two agents' efforts

mechanisms will rest on the probability of a player being low type, p, and the productivity ratio for low and high types, θ_L/θ_H. The following proposition gives a precise characterization of the optimal mechanism.

Proposition 4. [17] *Suppose the efforts are perfect substitutes. Input monitoring is better than output monitoring if and only if any one of the following conditions is satisfied:*

(a) $\frac{\theta_L}{\theta_H} \geq (3 - 2\sqrt{2})$;

(b) $\frac{\theta_L}{\theta_H} < (3 - 2\sqrt{2})$, *<u>and</u>*
either $p \leq \frac{(3\theta_H - \theta_L) - \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}$ *or* $p \geq \frac{(3\theta_H - \theta_L) + \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}$.

Another way to read the above proposition would be: *Output monitoring is better than input monitoring if and only if*

$$\frac{\theta_L}{\theta_H} < (3 - 2\sqrt{2}), \text{ and}$$
$$\frac{(3\theta_H - \theta_L) - \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)} < p < \frac{(3\theta_H - \theta_L) + \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}.$$

Proposition 4 can be illustrated as follows.

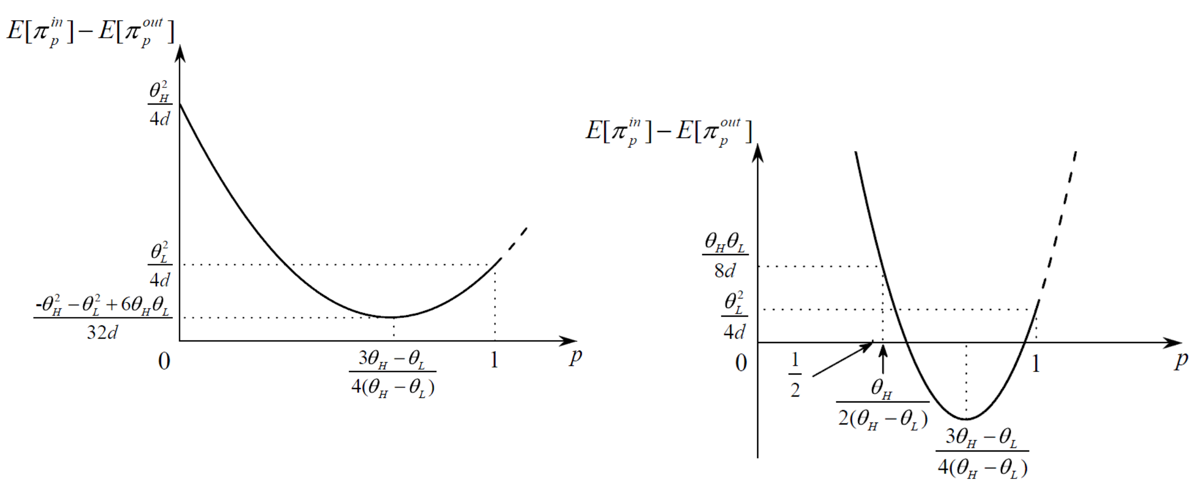

Figure 1: Difference in principal's profits under condition (a) (left panel) and (b) (right panel)

are independent, still the low-type agent enjoys a higher payoff due to the high-type agent's more generous effort (relative to the low-type agent's effort). This interpretation evokes a sense of free riding.

[17] Under some conditions input and output monitoring are equivalent for the principal. In such cases we will say that input monitoring is better.

The left panel of Fig. 1 describes the scenario when the productivity of the low-type agent is not too low or rather, the productivity of the high-type agent is not too high, indeed, bounded above by $1/(3-2\sqrt{2})\theta_L = (3+2\sqrt{2})\theta_L$. Thus, the advantage of output monitoring is too small to offset the moral hazard problem it brings about. So the principal prefers direct control over incentives through input monitoring. The only way output monitoring can benefit the principal is when the high type is substantially more productive than the low type. Only then it might make sense to let the agents self-select and respond according to their types (right panel of Fig. 1). Here we still need p to be reasonably high: for low p, principal would have set generous incentives even under input monitoring in anticipation of the agent being of high type; as p increases the principal will set flatter incentive under input monitoring, weakening its effectiveness, and output monitoring starts to dominate. Finally, if p is very close to 1, naturally the ex-ante expected gain from output monitoring disappears.

4.2 Social surplus

Now we again compare the ex-ante social surplus from input monitoring and output monitoring.

Under input monitoring, agents' payoffs are

$$\pi^{in}_{LL} = \pi^{in}_{HH} = \pi^{in}_{LH} = \pi^{in}_{HL} = \frac{[p\theta_L + (1-p)\theta_H]^2}{8d}, \tag{7}$$

and we know that the principal's expected profit is shown in equation (5). Thus, the ex-ante social surplus is

$$\begin{aligned} V^{in} &= \frac{[p\theta_L + (1-p)\theta_H]^2}{2d} + 2 \times \frac{[p\theta_L + (1-p)\theta_H]^2}{8d} \\ &= \frac{3[p\theta_L + (1-p)\theta_H]^2}{4d}. \end{aligned}$$

Under output monitoring, agents' payoffs are

$$\pi^{out}_{LL} = \frac{3\theta_L^2}{32d}, \quad \pi^{out}_{HH} = \frac{3\theta_H^2}{32d}, \quad \pi^{out}_{LH} = \frac{\theta_L^2 + 2\theta_H^2}{32d}, \quad \pi^{out}_{HL} = \frac{2\theta_L^2 + \theta_H^2}{32d}, \tag{8}$$

and we know that the principal's expected profit is shown in equation (6). Thus, the ex-ante social surplus is

$$\begin{aligned}V^{\text{out}} &= \frac{p\theta_L^2 + (1-p)\theta_H^2}{4d} + 2p^2 \times \frac{3\theta_L^2}{32d} + 2(1-p)^2 \times \frac{3\theta_H^2}{32d} + 2p(1-p)\left[\frac{\theta_L^2 + 2\theta_H^2}{32d} + \frac{2\theta_L^2 + \theta_H^2}{32d}\right] \\ &= \frac{7[p\theta_L^2 + (1-p)\theta_H^2]}{16d}.\end{aligned}$$

Therefore,

$$V^{\text{out}} - V^{\text{in}} = \frac{7(1-p)\theta_H^2 + 7p\theta_L^2 - 12[p\theta_L + (1-p)\theta_H]^2}{16d}.$$

Proposition 5. *Suppose efforts are perfect substitutes. Then the ex-ante social surplus under input monitoring is higher than that under output monitoring if and only if any one of the following conditions is satisfied:*

(a) $\frac{\theta_L}{\theta_H} \geq \frac{17-4\sqrt{15}}{7}$;

(b) $\frac{\theta_L}{\theta_H} < \frac{17-4\sqrt{15}}{7}$, *and*
either $p \leq \frac{(17\theta_H - 7\theta_L) - \sqrt{7(7\theta_H^2 - 34\theta_H\theta_L + 7\theta_L^2)}}{24(\theta_H - \theta_L)}$ *or* $p \geq \frac{(17\theta_H - 7\theta_L) + \sqrt{7(7\theta_H^2 - 34\theta_H\theta_L + 7\theta_L^2)}}{24(\theta_H - \theta_L)}$.

Another way to read the above proposition would be: *The ex-ante social surplus under output monitoring is higher than that under input monitoring if and only if*

$$\frac{\theta_L}{\theta_H} < \frac{17-4\sqrt{15}}{7}, \text{ and}$$
$$\frac{(17\theta_H - 7\theta_L) - \sqrt{7(7\theta_H^2 - 34\theta_H\theta_L + 7\theta_L^2)}}{24(\theta_H - \theta_L)} < p < \frac{(17\theta_H - 7\theta_L) + \sqrt{7(7\theta_H^2 - 34\theta_H\theta_L + 7\theta_L^2)}}{24(\theta_H - \theta_L)}.$$

As can be seen from Propositions 4 and 5, the comparison of social surplus under the two monitoring mechanisms shares a similar pattern with the comparison of principal's expected profits except for the different cutoff values of θ_L/θ_H and p. Similar pattern is due to the simple fact that social surplus is the sum of principal's profit and agents' payoffs, and these two parts are likely to often co-move: the principal does not mind conceding more rents to the agents if in the process he can improve his profits. So if input monitoring dominates output monitoring (or vice versa) in terms of social surplus, it stands to reason that the principal will have benefited under input monitoring (or output monitoring) by extracting a fraction of the higher social surplus even if he concedes part of the gains to the agents. For example, when $p = 0.5$ and $\theta_L/\theta_H = 0.3$, both the ex-ante social surplus and principal's expected profit are higher under input monitoring (Propositions 4 and 5), and the agents' combined payoffs are also higher under input monitoring (as can be verified using (7) and (8)).[18]

The θ_L/θ_H and p cutoffs in Proposition 5, however, differ from the cutoffs in Proposition 4. When θ_L/θ_H is above $(17-4\sqrt{15})/7 \approx 0.215$ or below $3-2\sqrt{2} \approx 0.172$, the agents' payoffs are more aligned with the principal's profits. When $\theta_L/\theta_H > 0.215$, the rankings of the two mechanisms are the same for comparison of either the principal's profits or the ex-ante social surplus. When $\theta_L/\theta_H < 0.172$, output monitoring dominates in terms of social surplus for a wider range of p values than for principal's profits, since $\frac{(17\theta_H-7\theta_L)-\sqrt{7(7\theta_H^2-34\theta_H\theta_L+7\theta_L^2)}}{24(\theta_H-\theta_L)} < \frac{(3\theta_H-\theta_L)-\sqrt{\theta_H^2-6\theta_H\theta_L+\theta_L^2}}{4(\theta_H-\theta_L)}$ and $\frac{(17\theta_H-7\theta_L)+\sqrt{7(7\theta_H^2-34\theta_H\theta_L+7\theta_L^2)}}{24(\theta_H-\theta_L)} > \frac{(3\theta_H-\theta_L)+\sqrt{\theta_H^2-6\theta_H\theta_L+\theta_L^2}}{4(\theta_H-\theta_L)}$. In the range $0.172 < \theta_L/\theta_H < 0.215$, the principal's attempt at maximization of ex-ante profits would have inclined him towards input monitoring, by Proposition 4, but output monitoring can dominate input monitoring in terms of social surplus (for some "intermediate" values of p) as one can see from Proposition 5. This implies that the agents' collective ex-ante payoffs will be much lower under input monitoring, and the principal is able to shift a greater share of the smaller social surplus towards himself. This is the usual efficiency loss due to the principal considering his own gains only.

4.3 Wage linear in both input and output

Similar to section 3.3, consider linear wage in both input and output: $W = \alpha_{in}e_j + \alpha_{out}y$, $\alpha_{in}, \alpha_{out} > 0$. Agent j's payoff from choosing effort e_j while agent k chooses e_k is

$$\pi_j = \alpha_{in}e_j + \alpha_{out}(\theta_j e_j + \theta_k e_k) - d \cdot \frac{(e_j)^2}{2}, \qquad j,k = 1,2, \quad k \neq j.$$

Agent j will respond by choosing an effort $e_j^{in} = \frac{\alpha_{in}+\alpha_{out}\theta_j}{d}$. Then, the principal's expected profit function can be written as:

$$\begin{aligned} E[\pi_p] = \frac{2}{d} \quad \{ & \quad [p\theta_L + (1-p)\theta_H]\alpha_{in} - 2\alpha_{in}^2 + [p\theta_L^2 + (1-p)\theta_H^2]\alpha_{out} \\ & - \quad 2[p\theta_L^2 + (1-p)\theta_H^2]\alpha_{out}^2 - 3[p\theta_L + (1-p)\theta_H]\alpha_{in}\alpha_{out}\}. \end{aligned}$$

The principal will choose α_{in} and α_{out} such that

$$\begin{aligned} \frac{\partial E[\pi_p]}{\partial \alpha_{in}} &= \frac{2}{d}\{[p\theta_L + (1-p)\theta_H](1-3\alpha_{out}) - 2\alpha_{in})\} = 0, \\ \frac{\partial E[\pi_p]}{\partial \alpha_{out}} &= \frac{2}{d}\{[p\theta_L^2 + (1-p)\theta_H^2](1-4\alpha_{out}) - 3[p\theta_L + (1-p)\theta_H]\alpha_{in}\} = 0. \end{aligned}$$

[18] Of course there is no guarantee that the agents' combined payoffs will be higher whenever social surplus is higher. It is quite possible that the higher social surplus is mainly due to the principal's gain.

Solving the above two first-order conditions we obtain:

$$\alpha_{in} = \frac{[p\theta_L + (1-p)\theta_H][p\theta_L^2 + (1-p)\theta_H^2]}{8[p\theta_L^2 + (1-p)\theta_H^2] - 9[p\theta_L + (1-p)\theta_H]^2},$$
$$\alpha_{out} = \frac{2[p\theta_L^2 + (1-p)\theta_H^2] - 3[p\theta_L + (1-p)\theta_H]^2}{8[p\theta_L^2 + (1-p)\theta_H^2] - 9[p\theta_L + (1-p)\theta_H]^2}.$$

For $\alpha_{in}, \alpha_{out} > 0$, the following condition must be satisfied:

$$\frac{p\theta_L^2 + (1-p)\theta_H^2}{[p\theta_L + (1-p)\theta_H]^2} > \frac{3}{2}. \tag{9}$$

By substituting α_{in} and α_{out} into $E[\pi_p]$, we derive principal's payoff:

$$E[\pi_p] = \frac{2(p\theta_L^2 + (1-p)\theta_H^2)[3(p\theta_L + (1-p)\theta_H)^2 - 2(\theta_L^2 + (1-p)\theta_H^2)][3(p\theta_L + (1-p)\theta_H)^2 - 4(\theta_L^2 + (1-p)\theta_H^2)]}{d\{8[p\theta_L^2 + (1-p)\theta_H^2] - 9[p\theta_L + (1-p)\theta_H]^2\}}.$$

■ **Comparison with input monitoring and output monitoring.** We know that if we compare input monitoring with output monitoring, either mechanism can dominate based on different conditions. What if we compare them with the mixed wages?

Proposition 6. *Suppose efforts are perfect substitutes. Then mixed wage dominates both input monitoring and output monitoring when* $\frac{\theta_L}{\theta_H} < 2 - \sqrt{3}$ *and* $\frac{(2\theta_H - \theta_L) - \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)} < p < \frac{(2\theta_H - \theta_L) + \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)}$.

Observe that when mixed wage dominates, there is no uniform ranking between input and output monitoring. This is easy to see from Proposition 4 where either of the two monitoring mechanisms can be the dominating one when $\frac{\theta_L}{\theta_H} < 2 - \sqrt{3}$ and $\frac{(2\theta_H - \theta_L) - \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)} < p < \frac{(2\theta_H - \theta_L) + \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)}$.[19] The main takeaway from Proposition 6 is that sometimes mixed wage can do better than relying on just one of the two measures of performance, input or output. This is not surprising given that each of the polar mechanisms has some specific advantage over the other, as discussed in the comparison before Proposition 4.

[19]Note that $2 - \sqrt{3} > 3 - 2\sqrt{2}$, $\frac{(2\theta_H - \theta_L) - \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)} < \frac{(3\theta_H - \theta_L) - \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}$ and $\frac{(2\theta_H - \theta_L) + \sqrt{\theta_H^2 - 4\theta_H\theta_L + \theta_L^2}}{3(\theta_H - \theta_L)} > \frac{(3\theta_H - \theta_L) + \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}$ when $\frac{\theta_L}{\theta_H} < 3 - 2\sqrt{2}$.

5 Other technologies

The analysis of perfect substitution and perfectly complementary technologies provides important lessons on the relative merits of different monitoring mechanisms. What has been missing, however, is a meaningful tension between agents' effort decisions in output monitoring: either the equilibrium efforts are independent (as under perfect substitution) or the efforts are perfectly coordinated (as under perfect complementarity). In this section, we expand the analysis by including other technologies in order to identify the influence of the two dominant forces in team incentives – *concerns for coordination* and *concerns for free riding*. One of these two may be more influential depending on the technology and subject to some qualifications.

5.1 Submodular technology

We first consider a non-linear production technology with linear isoquants, $y = (\theta_1 e_1 + \theta_2 e_2)^\gamma$, $0 < \gamma < 1$, as we had hinted in Section 4. The main objective is to allow agents' effort decisions under output monitoring to interact in a way to capture ***free riding in teams*** as what we understand from Holmstrom (1982). Given that the chosen technology is submodular, i.e. $\frac{\partial^2 y}{\partial e_i \partial e_j} < 0$, with linear output incentives the efforts will be ***strategic substitutes*** (as will be shown later). It might be recalled that, for the perfect substitution technology (also represented by linear isoquants) effort decisions were independent for linear incentives. We are going to argue that now with the free-rider problem, principal's ranking will swing towards input monitoring.

■ **Input monitoring.** Suppose the principal offers wages $W^{in} = \alpha_{in} e_j^{in}$, $\alpha_{in} > 0$. Agent j will choose effort $e_j^{in} = \frac{\alpha_{in}}{d}$. The principal's expected profit function is

$$\begin{aligned} E[\pi_p^{in}] &= p^2[(2\frac{\alpha_{in}\theta_L}{d})^\gamma - 2\frac{\alpha_{in}^2}{d}] + 2p(1-p)[(\frac{\alpha_{in}}{d}(\theta_L+\theta_H))^\gamma - 2\frac{\alpha_{in}^2}{d}] + (1-p)^2[(2\frac{\alpha_{in}\theta_H}{d})^\gamma - 2\frac{\alpha_{in}^2}{d}] \\ &= [2^\gamma p^2 \theta_L^\gamma + 2p(1-p)(\theta_L+\theta_H)^\gamma + 2^\gamma(1-p)^2\theta_H^\gamma]\frac{\alpha_{in}^\gamma}{d^\gamma} - 2\frac{\alpha_{in}^2}{d}. \end{aligned}$$

So he will choose α_{in} such that

$$\begin{aligned} \frac{\partial E[\pi_p^{in}]}{\partial \alpha_{in}} &= [2^\gamma p^2 \theta_L^\gamma + 2p(1-p)(\theta_L+\theta_H)^\gamma + 2^\gamma(1-p)^2\theta_H^\gamma]\gamma\frac{\alpha_{in}^{\gamma-1}}{d^\gamma} - \frac{4\alpha_{in}}{d} = 0, \\ \text{i.e., } \alpha_{in} &= \left\{\frac{[2^\gamma p^2 \theta_L^\gamma + 2p(1-p)(\theta_L+\theta_H)^\gamma + 2^\gamma(1-p)^2\theta_H^\gamma]\gamma d^{1-\gamma}}{4}\right\}^{\frac{1}{2-\gamma}}. \end{aligned}$$

Therefore, both high- and low-type agents' equilibrium effort choices are,

$$e_H^{in} = e_L^{in} = \left\{ \frac{[2^\gamma p^2 \theta_L^\gamma + 2p(1-p)(\theta_L + \theta_H)^\gamma + 2^\gamma (1-p)^2 \theta_H^\gamma]\gamma}{4d} \right\}^{\frac{1}{2-\gamma}}.$$

The expected profit for the principal is

$$E[\pi_p^{in}] = \frac{(2-\gamma)\gamma^{\frac{\gamma}{2-\gamma}} [2^\gamma p^2 \theta_L^\gamma + 2p(1-p)(\theta_L + \theta_H)^\gamma + 2^\gamma (1-p)^2 \theta_H^\gamma]^{\frac{2}{2-\gamma}}}{2^{\frac{2+\gamma}{2-\gamma}} d^{\frac{\gamma}{2-\gamma}}}. \tag{10}$$

■ **Output monitoring.** Consider output-based wages: $W^{out} = \alpha_{out} y$, $\alpha_{out} > 0$. Agent j's payoff function is $\pi_j^{out} = \alpha_{out}(\theta_j e_j^{out} + \theta_k e_k^{out})^\gamma - d \cdot \frac{(e_j^{out})^2}{2}$, $j, k = 1, 2$ and $k \neq j$. Based on the first-order condition,

$$\gamma \alpha_{out} (\theta_j e_j^{out} + \theta_k e_k^{out})^{\gamma - 1} \theta_j - d e_j^{out} = 0,$$

agent j's best-response function can be solved implicitly.

For $\frac{de_j^{out}}{de_k^{out}} < 0$, define its absolute value to be a measure of the degree of strategic substitutability.

We are not aware of any well-defined measure of strategic substitutability in the literature on submodular games. Our definition is meant to capture the responsiveness of one agent's effort to that of the other, i.e., the extent of free riding.

Lemma 1. *Suppose the production function is $y = (\theta_1 e_1 + \theta_2 e_2)^\gamma$, $0 < \gamma < 1$. Under output monitoring, agents' efforts are strategic substitutes, and the degree of substitutability decreases with γ.*

Thus the free-rider problem manifests in output monitoring, and it becomes less and less significant as γ approaches 1. In addition, an agent will lower his effort if the other agent happens to be a high rather than low type.

Solving the first-order conditions yields agent j's effort $e_j^{out} = \frac{(\gamma \alpha_{out})^{\frac{1}{2-\gamma}} \theta_j}{d^{\frac{1}{2-\gamma}} (\theta_j^2 + \theta_k^2)^{\frac{1-\gamma}{2-\gamma}}}$ and symmetrically for agent k. The principal then derives the expected profit function to determine effort incentives as follows:

$$\begin{aligned}
E[\pi_p^{out}] =& p^2 \left[\frac{2(\gamma\alpha_{out})^{\frac{1}{2-\gamma}}\theta_L^2}{d^{\frac{1}{2-\gamma}}(2\theta_L^2)^{\frac{1-\gamma}{2-\gamma}}}\right]^{\gamma}(1-2\alpha_{out}) + 2p(1-p)\left[\frac{(\gamma\alpha_{out})^{\frac{1}{2-\gamma}}}{d^{\frac{1}{2-\gamma}}(\theta_L^2+\theta_H^2)^{\frac{1-\gamma}{2-\gamma}}}(\theta_L^2+\theta_H^2)\right]^{\gamma}(1-2\alpha_{out}) \\
&+ (1-p)^2\left[\frac{2(\gamma\alpha_{out})^{\frac{1}{2-\gamma}}\theta_H^2}{d^{\frac{1}{2-\gamma}}(2\theta_H^2)^{\frac{1-\gamma}{2-\gamma}}}\right]^{\gamma}(1-2\alpha_{out}) \\
=& \left(\frac{\gamma\alpha_{out}}{d}\right)^{\frac{\gamma}{2-\gamma}}(1-2\alpha_{out})\left[2^{\frac{\gamma}{2-\gamma}}p^2\theta_L^{\frac{2\gamma}{2-\gamma}} + 2p(1-p)(\theta_L^2+\theta_H^2)^{\frac{\gamma}{2-\gamma}} + 2^{\frac{\gamma}{2-\gamma}}(1-p)^2\theta_H^{\frac{2\gamma}{2-\gamma}}\right];
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial E[\pi_p^{out}]}{\partial \alpha_{out}} \\
&= \frac{\alpha_{out}^{\frac{2\gamma-2}{2-\gamma}}}{d^{\frac{\gamma}{2-\gamma}}}\left(\frac{\gamma^{\frac{2}{2-\gamma}}(1-2\alpha_{out})}{2-\gamma} - 2\gamma^{\frac{\gamma}{2-\gamma}}\alpha_{out}\right)\left[2^{\frac{\gamma}{2-\gamma}}p^2\theta_L^{\frac{2\gamma}{2-\gamma}} + 2p(1-p)(\theta_L^2+\theta_H^2)^{\frac{\gamma}{2-\gamma}} + 2^{\frac{\gamma}{2-\gamma}}(1-p)^2\theta_H^{\frac{2\gamma}{2-\gamma}}\right] \\
&= 0,
\end{aligned}$$

i.e., $\alpha_{out} = \frac{\gamma}{4}$. (The second solution $\alpha_{out} = 0$ will fail the second-order condition.)

The agents' equilibrium efforts are solved as $e_j^{out} = \frac{\gamma^{\frac{2}{2-\gamma}}\theta_j}{(4d)^{\frac{1}{2-\gamma}}(\theta_j^2+\theta_k^2)^{\frac{1-\gamma}{2-\gamma}}}$, $j, k = 1, 2$, $j \neq k$, giving rise to principal's expected profit

$$E[\pi_p^{out}] = \left(\frac{\gamma^2}{4d}\right)^{\frac{\gamma}{2-\gamma}}\left(1-\frac{\gamma}{2}\right)\left[2^{\frac{\gamma}{2-\gamma}}p^2\theta_L^{\frac{2\gamma}{2-\gamma}} + 2p(1-p)(\theta_L^2+\theta_H^2)^{\frac{\gamma}{2-\gamma}} + 2^{\frac{\gamma}{2-\gamma}}(1-p)^2\theta_H^{\frac{2\gamma}{2-\gamma}}\right]. \quad (11)$$

■ **Comparison.** Analytical comparison of (10) and (11) does not lead to any manipulable expression, so we rely on Mathematica plots that exhibit the following pattern of dominance.

Simulation result 1. *Suppose the team production technology is* $y = (\theta_1 e_1 + \theta_2 e_2)^{\gamma}$, $0 < \gamma < 1$, *which is submodular. Then input monitoring tends to dominate output monitoring when* $\frac{\theta_L}{\theta_H}$ *is appropriately large. When* $\frac{\theta_L}{\theta_H}$ *is relatively small, input monitoring can still often dominate output monitoring when* γ *is appropriately small (equivalently, strategic substitutability is large), and output monitoring can sometimes be the dominant mechanism when* γ *is large (substitutability small).*

Figs. 2 and 3 are sample plots to indicate how the relative productivity and degree of strategic substitution matters for the monitoring choice. The only possible scenario that output monitoring could dominate input monitoring is when $\frac{\theta_L}{\theta_H}$ is small and γ is sufficiently large.

The dominance results of the two mechanisms and the underlying reasons share many similarities with the perfect substitution case. When $\frac{\theta_L}{\theta_H}$ is large, the advantage of output

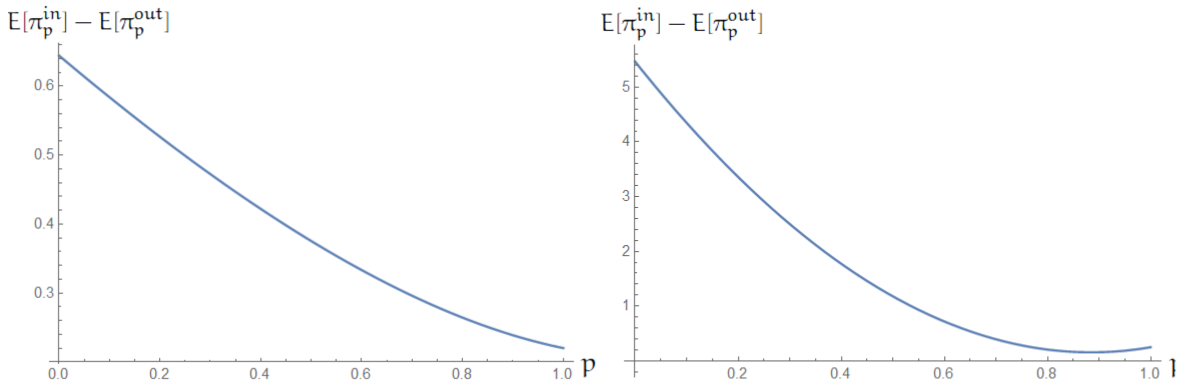

Figure 2: Profit difference, strategic substitutes – (i) $d = 1$, $\theta_L = 1, \theta_H = 5, \gamma = 0.5$ (left panel); (ii) $d = 1$, $\theta_L = 1, \theta_H = 5, \gamma = 0.98$ (right panel)

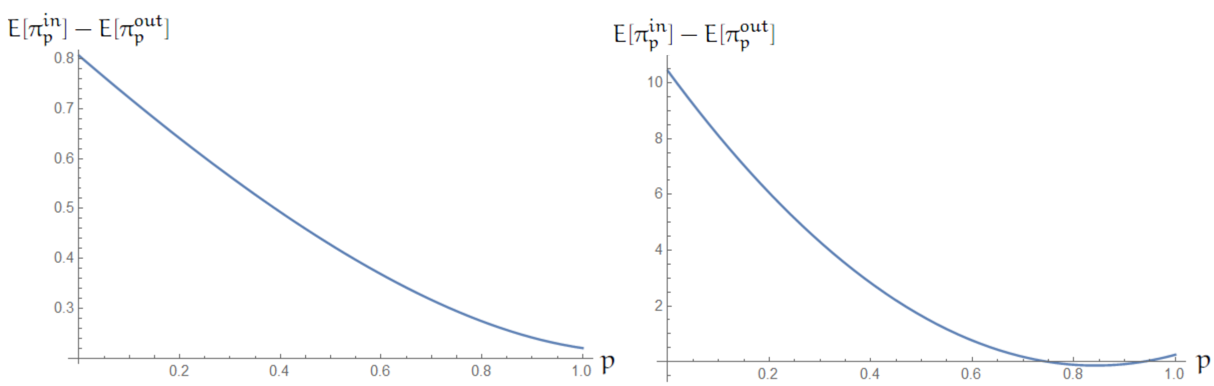

Figure 3: Profit difference, strategic substitutes – (i) $d = 1$, $\theta_L = 1, \theta_H = 7, \gamma = 0.5$ (left panel); (ii) $d = 1$, $\theta_L = 1, \theta_H = 7, \gamma = 0.98$ (right panel)

monitoring, as it induces agents to put in effort proportional to their types, are not so significant. So the team moral hazard problem as well as the free-rider problem become the major concern, making input monitoring attractive. When $\frac{\theta_L}{\theta_H}$ is small, the dominance result depends on the value of γ. When γ is 'small', the overwhelming evidence in favor of input monitoring (shown in the left panel of Fig. 3), as opposed to occasional dominance of output monitoring in Proposition 4, can be explained as follows. Now output monitoring leads to free riding that was absent in the perfect substitution case, and the free-rider problem is significant as illustrated by Lemma 1.[20] For the (low, high) team type, while the principal prefers the high type to shoulder the majority of the burden of production, the wage to be paid to the low type is as much as that for the high type, and this is a cost for the principal to bear. By switching to input monitoring, the principal can reward the agents proportional

[20]Free riding, in the form of negative interdependence of agents efforts, happens because marginal productivity of an agent's effort, and hence his marginal reward, is declining in the other agent's effort for $\gamma < 1$.

to their efforts. More importantly, input monitoring avoids the problem of **strategic substitution of efforts** that would have plagued the output monitoring mechanism leading to general under-provision of efforts. As the production function $y = (\theta_1 e_1 + \theta_2 e_2)^\gamma$ approaches perfect substitution technology with $\gamma \to 1$, the free-rider problem gradually disappears, and output monitoring starts to dominate for some values of p. The right panel of Fig. 3 thus resembles the right panel of Fig. 1 for the perfect substitution technology.

5.2 Supermodular technology

As a final variant, consider the following Cobb-Douglas production function, $y = \theta_1\theta_2(e_1 e_2)^\beta$, $\beta > 0$.[21] The technology is supermodular, i.e. $\frac{\partial^2 y}{\partial e_i \partial e_j} > 0$, which will make the agents' efforts ***strategic complements*** under linear output incentives (shown later). So, coordination should be an important concern when the principal chooses between input and output monitoring.[22]

■ **Input monitoring.** Suppose the principal offers wages $W^{in} = \alpha_{in} e_j^{in}, \alpha_{in} > 0$. Agent j will choose effort $e_j^{in} = \frac{\alpha_{in}}{d}$. The principal's expected profit function is then derived as follows:

$$\begin{aligned} E[\pi_p^{in}] &= p^2[\theta_L^2(\frac{\alpha_{in}}{d})^{2\beta} - \frac{2\alpha_{in}^2}{d}] + 2p(1-p)[\theta_L\theta_H(\frac{\alpha_{in}^2}{d^2})^\beta - \frac{2\alpha_{in}^2}{d}] + (1-p)^2[\theta_H^2(\frac{\alpha_{in}}{d})^{2\beta} - \frac{2\alpha_{in}^2}{d}] \\ &= [p^2\theta_L^2 + 2p(1-p)\theta_L\theta_H) + (1-p)^2\theta_H^2](\frac{\alpha_{in}}{d})^{2\beta} - \frac{2\alpha_{in}^2}{d}. \end{aligned}$$

The principal will choose α_{in} such that

$$\begin{aligned} \frac{\partial E[\pi_p^{in}]}{\partial \alpha_{in}} =& [p^2\theta_L^2 + 2p(1-p)\theta_L\theta_H) + (1-p)^2\theta_H^2]\frac{2\beta}{d^{2\beta}}\alpha_{in}^{2\beta-1} - \frac{4\alpha_{in}}{d} = 0, \\ \text{i.e.,}\quad \alpha_{in} =& \left(\frac{[p^2\theta_L^2 + 2p(1-p)\theta_L\theta_H) + (1-p)^2\theta_H^2]\beta}{2d^{2\beta-1}}\right)^{\frac{1}{2-2\beta}}. \end{aligned}$$

(The second solution $\alpha_{in} = 0$ will fail the second-order condition.)

Therefore, high- and low-type agents' equilibrium efforts are

$$e_H^{in} = e_L^{in} = \left(\frac{[p^2\theta_L^2 + 2p(1-p)\theta_L\theta_H) + (1-p)^2\theta_H^2]\beta}{2d}\right)^{\frac{1}{2-2\beta}},$$

[21] We use a common β in the production function rather than $y = \theta_1\theta_2(e_1^\alpha e_2^\beta)$, because of our symmetry assumption about the agents' role in production. See Section 2.

[22] Moral hazard will also play an important role.

giving rise to the following expected profit for the principal:

$$E[\pi_p^{in}] = \frac{\beta^{\frac{\beta}{1-\beta}}(1-\beta)[p^2\theta_L^2 + 2p(1-p)\theta_L\theta_H) + (1-p)^2\theta_H^2]^{\frac{1}{1-\beta}}}{(2d)^{\frac{\beta}{1-\beta}}}. \tag{12}$$

■ **Output monitoring.** Now, suppose the principal offers wages based on output: $W^{out} = \alpha_{out}y,\ \alpha_{out} > 0$. Agent j's payoff function is $\pi_j^{out} = \alpha_{out}\theta_j\theta_k(e_j^{out}\cdot e_k^{out})^\beta - d\cdot\frac{(e_j^{out})^2}{2},\ j,k = 1,2$ and $k \neq j$. The first-order condition,

$$\beta\alpha_{out}\theta_j\theta_k(e_j^{out})^{\beta-1}(e_k^{out})^\beta - de_j^{out} = 0,$$

determines agent j's best response.

Let the steepness of the slope, $\frac{de_j^{out}}{de_k^{out}} > 0$, be a measure of the *degree of strategic complementarity,* similar to the measure of strategic substitutability.

Lemma 2. *Suppose the production function is* $y = \theta_1\theta_2(e_1e_2)^\beta$, $0 < \beta < 1$. *Under output monitoring, agents' efforts are strategic complements, and the degree of complementarity increases with* β.

Thus, as β increases, agents' effort coordination becomes more significant under output monitoring. As it can be seen from the proof, efforts will be strategic complements so long as $0 < \beta < 2$, and strategic substitutes for $\beta > 2$. We restrict to the range, $0 < \beta < 1$, for much sharper Mathematica simulations.

First-order conditions yield player j and k's effort choices: $e_j^{out} = e_k^{out} = (\frac{\beta\alpha_{out}}{d}\theta_j\theta_k)^{\frac{1}{2(1-\beta)}}$. So the principal's expected profit function can be written as:

$$\begin{aligned}
E[\pi_p^{out}] =& p^2\theta_L^2(\frac{\beta\alpha_{out}}{d}\theta_L^2)^{\frac{\beta}{1-\beta}}(1-2\alpha_{out}) + 2p(1-p)\theta_L\theta_H(\frac{\beta\alpha_{out}}{d}\theta_L\theta_H)^{\frac{\beta}{1-\beta}}(1-2\alpha_{out}) \\
&+ (1-p)^2\theta_H^2(\frac{\beta\alpha_{out}}{d}\theta_H^2)^{\frac{\beta}{1-\beta}}(1-2\alpha_{out}) \\
=& (\frac{\beta\alpha_{out}}{d})^{\frac{\beta}{1-\beta}}(1-2\alpha_{out})[p^2\theta_L^{\frac{2}{1-\beta}} + 2p(1-p)(\theta_L\theta_H)^{\frac{1}{1-\beta}} + (1-p)^2\theta_H^{\frac{2}{1-\beta}}].
\end{aligned}$$

The principal will choose α_{out} such that

$$\begin{aligned}
\frac{\partial E[\pi_p^{out}]}{\partial\alpha_{out}} =& [\frac{\beta}{1-\beta}\alpha_{out}^{\frac{2\beta-1}{1-\beta}}(1-2\alpha_{out}) - 2\alpha_{out}^{\frac{\beta}{1-\beta}}][p^2\theta_L^{\frac{2}{1-\beta}} + 2p(1-p)(\theta_L\theta_H)^{\frac{1}{1-\beta}} + (1-p)^2\theta_H^{\frac{2}{1-\beta}}] = 0, \\
&\text{i.e.,}\ \ \alpha_{out} = \frac{\beta}{4}.
\end{aligned}$$

(The second solution $\alpha_{out} = 0$ will fail the second-order condition.)

Thus, the agents' equilibrium efforts, $e_j^{out} = e_k^{out} = (\frac{\beta^2}{4d}\theta_j\theta_k)^{\frac{1}{2(1-\beta)}}$, $j, k = 1, 2$, $j \neq k$, will result in the following expected profit for the principal:

$$E[\pi_p^{out}] = \left(\frac{\beta^2}{2d}\right)^{\frac{\beta}{1-\beta}} (1-\beta)[p^2\theta_L^{\frac{2}{1-\beta}} + 2p(1-p)(\theta_L\theta_H)^{\frac{1}{1-\beta}} + (1-p)^2\theta_H^{\frac{2}{1-\beta}}]. \quad (13)$$

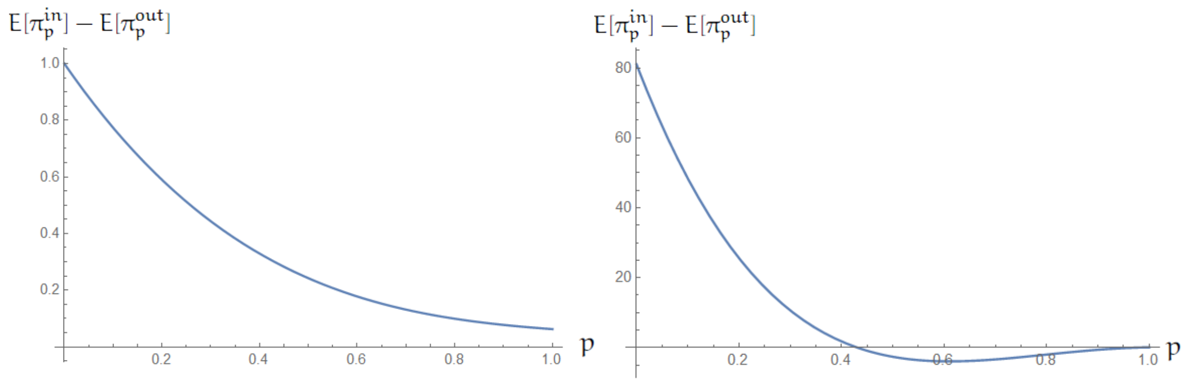

Figure 4: Profit comparisons, Cobb-Douglas, $\beta = \frac{1}{2}$ – (i) $\frac{\theta_L}{\theta_H} \geq \bar{x}$, in particular, $d = 1$, $\theta_L = 1$ and $\theta_H = 2$ (left panel); (ii) $\frac{\theta_L}{\theta_H} < \bar{x}$ ($d = 1$, $\theta_L = 1$ and $\theta_H = 6$) (right panel)

■ **Comparison.** Based on Mathematica simulation, we obtain the following pattern of dominance.

Simulation result 2. *Suppose the team production technology is* $y = \theta_1\theta_2(e_1e_2)^\beta$, $0 < \beta < 1$, *which is supermodular.*

1. *For each* β, *if* $\frac{\theta_L}{\theta_H}$ *is above a cutoff value, then input monitoring dominates output monitoring. Otherwise output monitoring may dominate.*

2. *As* β *increases, coordination starts to become important and output monitoring dominates with increasing frequency.*

We can illustrate the ambiguous nature of monitoring ranking (i.e., the first part) as follows. Left panel of Fig. 4 describes the scenario when the ratio of the productivities, i.e. $\frac{\theta_L}{\theta_H}$, is relatively high, bounded below by some number $\bar{x}$.[23] We can see that input monitoring generates a higher profit for the principal. In this case, the two agents are more alike, thus, coordination is not very important. Rather, the principal would like to reward the agents directly based on their effort in order to avoid the moral hazard problem. Right panel of Fig. 4, on the other hand, describes a scenario when the ratio of the productivities is relatively

[23] According to Mathematica simulation, $\bar{x}$ is around 0.298.

low.[24] When the value of p is either small or large, i.e., when there is a considerably high chance that the two agents are of the same type, input monitoring still dominates; when the value of p is not extreme so that the probability of high-low combination is considerably high, output monitoring dominates as it facilitates coordination.

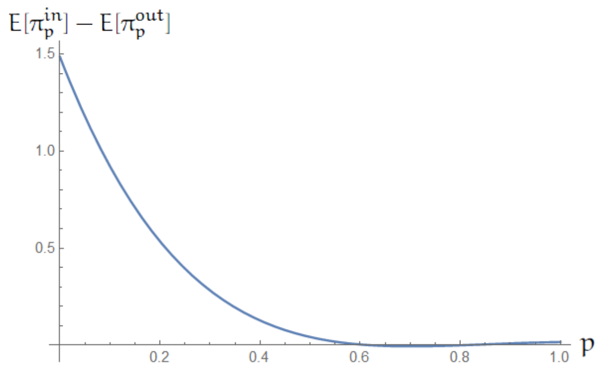

Figure 5: Profit comparisons, Cobb-Douglas, $\beta = 0.7$ – (i) $d = 1$, $\theta_L = 1$ and $\theta_H = 2$ (left panel); (ii) $d = 1$, $\theta_L = 1$ and $\theta_H = 6$ (right panel)

Fig. 5 presents a parallel scenario with only β increased to 0.7. This, together with Fig. 4, illustrates the second part of the Observation. In the left panel, the initial absolute dominance of input monitoring disappears and output monitoring may prevail sometimes. For the right panel, output monitoring dominates for a bigger range of p values. As shown in Lemma 2, higher β means stronger strategic complementarity, and thus output monitoring becomes more desirable for better coordination.

■ **Remark.** An obvious way to study the incentive issues analyzed in this paper, one might think, should be to consider the CES production technology, $y = [\theta_1 e_1^\rho + \theta_2 e_2^\rho]^{\frac{1}{\rho}}$, where the elasticity of substitution measured by $\sigma = \frac{d(e_2/e_1)}{e_2/e_1} / \frac{dTRS}{TRS}$ is constant and equal to $\frac{1}{1-\rho}$, with $\rho \in (-\infty, 1]$. Because $\rho = 1$, $\rho \to 0$ and $\rho \to -\infty$ correspond respectively to perfect substitution, Cobb-Douglas, and perfectly complementary technologies (Ch.1, Varian, 1992), one can vary ρ and possibly cover the entire span of isoquant maps. In the process we can try to understand how the curvature of isoquants, or roughly technological substitutability or complementarity, might influence the desirability of alternative monitoring mechanisms. The main limitation of this approach, however, is that the CES covers only a subset of supermodular technologies.[25] This leaves out all submodular technologies $\left(\text{i.e. } \frac{\partial^2 y}{\partial e_2 \partial e_1} < 0\right)$.[26]

[24] When p is close to 1, $E[\pi_p^{in}] - E[\pi_p^{out}] > 0$, although the graph does not show that clearly. In particular, when $p = 1$, $E[\pi_p^{in}] - E[\pi_p^{out}] = \frac{1}{16} > 0$.

[25] $\frac{\partial^2 y}{\partial e_2 \partial e_1} = (1-\rho)\theta_1\theta_2(e_1 e_2)^{\rho-1}\left(\theta_1 e_1^\rho + \theta_2 e_2^\rho\right)^{\frac{1}{\rho}-2} > 0$, for all but $\rho = 1$.

[26] As illustrated in Section 5.1, strict submodularity can give rise to free riding that plays an important

Also, as we verify in the Appendix, under output monitoring and linear incentives the agents' best-response functions are upward-sloping for the entire range $\rho \in (-\infty, 1)$, i.e., efforts are *strategic complements*. This suggests that the CES family will not be very useful to uncover any broad insight about the impact of free riding on the choice of monitoring.

6 Further discussions

We discuss two modifications of the current model.

■ **Type-dependent contracts.** An obvious alternative to our model is the one of type-dependent contract. Knowing that the agents would learn their own type as well the team member's type, which is plausible in the context of team projects, one can study a game where agents pick from a menu of contracts. Below we consider such an approach.

Consider input monitoring first. Principal can use the following type-dependent contract to achieve the first-best: The two agents are asked to report the team's type profile, i.e., own type and the other member's type. If their reports match, then the principal rewards each agent a wage just enough to cover his effort cost if the agent has exerted his part of the "first best" effort pair where first best is calculated thinking as if the type reports are truthful. Otherwise, the agent will receive *zero* wage.

Thus, each agent will have (weak) incentive to truthfully report the team's type and put in first-best effort, since any misreport of the team's type will create a mismatch of their reports,[27] which leads to zero wage, and any effort level other than the first-best effort will not be rewarded either.

The type-dependent output monitoring contract can be specified in a similar way:

The two agents are asked to report the team's type profile. If their reports match, then the principal rewards each agent a wage just enough to cover his "first-best effort" cost if the "first-best team output" is observed.[28] Otherwise, the agent receives zero wage.

Again, it is easy to see that there is a Nash equilibrium in which the agents will truthfully report the team's type and exert first-best effort.

Since both input monitoring and output monitoring can achieve the first-best outcome regardless of the technology if the principal chooses the "best" type-dependent contract (there is no point to restrict to linear contract if the principal wants to use sophisticated mechanisms), the two mechanisms are equally good.

role in the optimal monitoring choice.

[27] Here, we do not consider cooperation between the agents by jointly deviating from truthful reports. Instead, the agents play non-cooperatively.

[28] Again "first best" efforts and output are derived assuming type reports are truthful.

The above arguments rely on punishing contracts to implement the first best rather than linear contracts assumed throughout the paper. If solicitations of type reports are possible, there is no specific reason not to use such a contract because it is quite simple both in its wage specifications and in terms of implementation. There is no violation of limited liability on the equilibrium path either.

■ **Private information about types.** It is plausible to criticise that the above first-best implementation results are due to the fact that the agents know each other's types. What if the agents do not know each other's types?

First, it is not difficult to see that all the results under input monitoring should remain unaffected if we stick to our assumed class of linear contracts based on input alone and not on any additional report of own type or the team member's type. The agents are rewarded based solely on their own efforts.

When efforts are perfect substitutes, the results under output monitoring will also be the same as before since an agent's effort will be independent of his partner's type due to the linear wage structure. Thus, we can conclude that the ranking results of the two mechanisms when efforts are perfect substitutes will not be altered even if the agent does not know the other agent's type.

When the technology is of other forms, deriving optimal linear wage incentives under private information of agent types will pose a much tougher challenge. This makes comparing the monitoring mechanisms a difficult exercise.

Appendix

Proof of Proposition 1. When agents' efforts are perfect complements,

$$\begin{aligned}
&E[\pi_p^{out}] - E[\pi_p^{in}] \\
=&\frac{p(2-p)\theta_L^2 + (1-p)^2\theta_H^2}{8d} - \frac{[p(2-p)\theta_L + (1-p)^2\theta_H]^2}{8d} \\
=&\frac{1}{8d}\{p(2-p)[1-p(2-p)]\theta_L^2 + (1-p)^2[1-(1-p)^2]\theta_H^2 - 2p(2-p)(1-p)^2\theta_L\theta_H\} \\
=&\frac{1}{8d}\{p(2-p)(1-p)^2\theta_L^2 + p(2-p)(1-p)^2\theta_H^2 - 2p(2-p)(1-p)^2\theta_L\theta_H\} \\
=&\frac{1}{8d}p(2-p)(1-p)^2(\theta_L-\theta_H)^2 > 0.
\end{aligned}$$

Thus, output monitoring is better than input monitoring for the principal. **Q.E.D.**

Proof of Proposition 2. We can simplify $V^{out} - V^{in}$ into the following expression:

$$V^{out} - V^{in} = \frac{p(1-p)[3(1-k)^2p^2 - 9(1-k)^2p - (k^4 - 7k^2 + 12k - 6)]}{16d\theta_H^4},$$

where $k = \frac{\theta_L}{\theta_H}$.

Proving $V^{out} > V^{in}$ is equivalent to showing that

$$3(1-k)^2p^2 - 9(1-k)^2p - (k^4 - 7k^2 + 12k - 6) > 0.$$

The left-hand side of the inequality is quadratic in p. In order to prove the above, it is important to know whether the expression

$$3(1-k)^2p^2 - 9(1-k)^2p - (k^4 - 7k^2 + 12k - 6) = 0 \tag{14}$$

has real number roots p for which $0 < p < 1$. We first calculate that

$$\begin{aligned} \delta_c &\equiv [9(1-k)^2]^2 + 4 \times 3(1-k)^2(k^4 - 7k^2 + 12k - 6) \\ &= 3(2k^2 + 3k + 3)(1-2k)(1-k)^3. \end{aligned}$$

When $0.5 < k < 1$, $\delta_c < 0$. So there is no real roots p for equation (14). Thus, the left-hand side of equation (14) is always positive, i.e., $V^{out} > V^{in}$.

When $0 < k \leq 0.5$, $\delta_c \geq 0$, and we can write the roots as

$$p_1 = \frac{3}{2} + \frac{\sqrt{3(2k^2 + 3k + 3)(1-2k)(1-k)}}{6(1-k)} \quad \text{or} \quad p_2 = \frac{3}{2} - \frac{\sqrt{3(2k^2 + 3k + 3)(1-2k)(1-k)}}{6(1-k)}$$

It can be shown that $\frac{\sqrt{3(2k^2+3k+3)(1-2k)(1-k)}}{6(1-k)}$ is decreasing in k, so the smallest value of p approaches $\frac{3}{2} - \frac{\sqrt{9}}{6(1)} = 1$ from the right side when k approaches 0. Thus, when $0 < p < 1$, the left-hand side of equation (14) is again always positive, i.e., $V^{out} > V^{in}$. **Q.E.D.**

Proof of Proposition 4. When agents' efforts are perfect substitutes, define $D \equiv E[\pi_p^{in}] - E[\pi_p^{out}]$, thus,

$$D = \frac{1}{4d}\{(1-2p)(\theta_H - \theta_L)[p\theta_L + (1-p)\theta_H] + \theta_H\theta_L\}.$$

Note that when $p = 0$, $D = \frac{\theta_H^2}{4d} > 0$ and when $p = 1$, $D = \frac{\theta_L^2}{4d} > 0$. Now, we are looking for the minimum value of D:

$$\frac{\partial D}{\partial p} = \frac{\theta_H - \theta_L}{4d}[\theta_L - 3\theta_H - 4p(\theta_L - \theta_H)] = 0, \qquad \text{i.e., } p_{\min} = \frac{3\theta_H - \theta_L}{4(\theta_H - \theta_L)}.$$

Since $\frac{\partial D}{\partial p} < 0$ when $p < \frac{3\theta_H - \theta_L}{4(\theta_H - \theta_L)}$, as long as $p_{\min} \geq 1$, we will have $D \geq 0$. Equivalently, when $\theta_H \leq 3\theta_L$, input monitoring generates higher expected profit for the principal. If $\theta_H > 3\theta_L$, we look at the minimum value of D:

$$D_{\min} = \frac{1}{8}(-\theta_H^2 - \theta_L^2 + 6\theta_H\theta_L).$$

Thus, as long as the minimum value of D is above 0, input monitoring is better. We know $-\theta_H^2 - \theta_L^2 + 6\theta_H\theta_L \geq 0$ if and only if $(3-2\sqrt{2})\theta_L < \theta_H \leq (3+2\sqrt{2})\theta_L$. Thus, input monitoring is better when $\frac{\theta_H}{\theta_L} \leq (3 + 2\sqrt{2})$. We have proven condition (a).

If $\frac{\theta_H}{\theta_L} > (3 + 2\sqrt{2})$, then $-\theta_H^2 - \theta_L^2 + 6\theta_H\theta_L < 0$, i.e., the minimum value of D is below 0, we need to look at the horizontal intercept. When $D = 0$,

$$p = \frac{(3\theta_H - \theta_L) \pm \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}.$$

Since D is a convex function of p, $D > 0$ if and only if

$$p < \frac{(3\theta_H - \theta_L) - \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)} \quad \text{or} \quad p > \frac{(3\theta_H - \theta_L) + \sqrt{\theta_H^2 - 6\theta_H\theta_L + \theta_L^2}}{4(\theta_H - \theta_L)}.$$

We have proven condition (b). **Q.E.D.**

Proof of Proposition 5. We can simplify $V^{out} - V^{in}$ into the following expression:

$$V^{out} - V^{in} = \frac{-12(1-k)^2p^2 + (17-7k)(1-k)p - 5}{16d\theta_H^2},$$

where $k = \frac{\theta_L}{\theta_H}$.

Proving $V^{out} > V^{in}$ is equivalent to showing that

$$-12(1-k)^2p^2 + (17-7k)(1-k)p - 5 > 0.$$

The left-hand side of the inequality is quadratic in p. In order to prove the above, it is important to know whether the expression

$$-12(1-k)^2p^2 + (17-7k)(1-k)p - 5 = 0 \tag{15}$$

has real number roots p for which $0 < p < 1$. We first calculate that

$$\begin{aligned} \delta &= (17-7k)^2(1-k)^2 - 240(1-k)^2 \\ &= 7(1-k)^2(7k^2-34k+7). \end{aligned}$$

When $\frac{17-4\sqrt{15}}{7} < k < 1$, $\delta < 0$. So there is no real roots p for equation (15). Thus, the left-hand side of equation (15) is always negative, i.e., $V^{in} > V^{out}$.

When $k = \frac{17-4\sqrt{15}}{7}$, $\delta = 0$. So there is one real root for equation (15), where

$$p_0 = \frac{17-7k}{24(1-k)}.$$

Since $k = \frac{17-4\sqrt{15}}{7}$, we can find out that $0 < p_0 < 1$. Thus, $V^{in} \geq V^{out}$.

When $0 < k < \frac{17-4\sqrt{15}}{7}$, $\delta > 0$, and we can write the roots as

$$p_1 = \frac{17-7k-\sqrt{7(7-34k+7k^2)}}{24(1-k)} \quad \text{or} \quad p_2 = \frac{17-7k+\sqrt{7(7-34k+7k^2)}}{24(1-k)}$$

It can be shown that when $0 < k < \frac{17-4\sqrt{15}}{7}$, $0 < p_1, p_2 < 1$. Thus, when $p_1 < p < p_2$, $V^{out} < V^{in}$; and when $0 < p \leq p_1$ or $p_2 \leq p < 1$, $V^{in} \geq V^{out}$. **Q.E.D.**

Proof of Proposition 6. We know that mixed wage dominates both input-based wage and output-based wage as long as condition (9) is satisfied, i.e.,

$$\frac{pk^2 + (1-p)}{[pk + (1-p)]^2} > \frac{3}{2}.$$

Rearranging it, we obtain an equivalent condition as follows:

$$3(1-k)^2p^2 - 2(1-k)(2-k)p + 1 < 0. \tag{16}$$

In order to find out the condition under which the above is satisfied, it is important to know whether the expression

$$3(1-k)^2p^2 - 2(1-k)(2-k)p + 1 = 0$$

has real number roots p for which $0 < p < 1$. We first calculate that

$$\begin{aligned} \delta &= 4(1-7)^2(1-k)^2 - 12(1-k)^2 \\ &= 4(1-k)^2(k^2-4k+1). \end{aligned}$$

When $0 < k < 2 - \sqrt{3}$, $\delta > 0$. Thus, there are two real roots that can be expressed as follows:

$$p_1 = \frac{2 - k - \sqrt{1 - 4k + k^2}}{3(1-k)} \quad \text{or} \quad p_2 = \frac{2 - k + \sqrt{1 - 4k + k^2}}{3(1-k)}$$

It can be shown that when $0 < k < 2 - \sqrt{3}$, $0 < p_1, p_2 < 1$. Thus, when $0 < k < 2 - \sqrt{3}$ and $p_1 < p < p_2$, condition (16) is satisfied. **Q.E.D.**

Proof of Lemma 1. For simplicity, we will write e_j and e_k instead of e_j^{out} and e_k^{out}. The first-order condition can be rearranged as:

$$\frac{\gamma \alpha_{out} \theta_j}{d} (\theta_j e_j + \theta_k e_k)^{\gamma - 1} = e_j. \tag{17}$$

Taking total differentiation, we have

$$\frac{\gamma \alpha_{out} \theta_j}{d} (\gamma - 1)(\theta_j e_j + \theta_k e_k)^{\gamma - 2} (\theta_j \frac{de_j}{de_k} + \theta_k) = \frac{de_j}{de_k}.$$

Using (17), we get

$$(\gamma - 1)(\theta_j e_j + \theta_k e_k)^{-1} e_j (\theta_j \frac{de_j}{de_k} + \theta_k) = \frac{de_j}{de_k}.$$

By rearranging, we obtain

$$\frac{de_j}{de_k} = \frac{(\gamma - 1) e_j \theta_k}{(2 - \gamma)\theta_j e_j + \theta_k e_k}.$$

Since $0 < \gamma < 1$, $\frac{de_j}{de_k} < 0$, i.e. agents' efforts are strategic substitutes.

Now,

$$\begin{aligned} \frac{\partial (\frac{de_j}{de_k})}{\partial \gamma} &= \frac{[(2-\gamma)\theta_j e_j + \theta_k e_k] e_j \theta_k - (\gamma - 1) e_j \theta_k e_j \theta_j (-1)}{[(2-\gamma)\theta_j e_j + \theta_k e_k]^2} \\ &= \frac{\theta_j \theta_k e_j^2 + (2 - \gamma)\theta_k^2 e_j e_k}{[(2-\gamma)\theta_j e_j + \theta_k e_k]^2} > 0. \end{aligned}$$

Thus, as γ increases, $\frac{de_j}{de_k}$ will become less negative, i.e. the degree of substitutability decreases. **Q.E.D.**

Proof of Lemma 2. For simplicity, we will write e_j and e_k instead of e_j^{out} and e_k^{out}. The first-order condition can be rearranged as:

$$e_j = \left(\frac{\beta}{d} \alpha_{out} \theta_j \theta_k \right)^{\frac{1}{2-\beta}} e_k^{\frac{\beta}{2-\beta}}. \tag{18}$$

Taking total differentiation, we have

$$\frac{de_j}{de_k} = \frac{\beta}{2-\beta}\left(\frac{\beta}{d}\alpha_{out}\theta_j\theta_k\right)^{\frac{1}{2-\beta}} e_k^{\frac{\beta}{2-\beta}-1} = \frac{\beta e_j}{(2-\beta)e_k} \quad \text{(using (18))}.$$

Since $0 < \beta < 1$, $\frac{de_j}{de_k} > 0$, i.e. agents' efforts are strategic complements.

Now,

$$\frac{\partial(\frac{de_j}{de_k})}{\partial\beta} = \frac{2e_j}{e_k(2-\beta)^2} > 0.$$

Thus, as β increases, the degree of complementarity increases. **Q.E.D.**

■ **CES technology**. Consider the CES production technology satisfying constant returns to scale:

$$y = [\theta_1 e_1^\rho + \theta_2 e_2^\rho]^{\frac{1}{\rho}}.$$

We will consider only output monitoring. Let $W^{out} = \alpha_{out} y$, $\alpha_{out} > 0$. Agent j's ex-post payoff is $\pi_j^{out} = \alpha_{out}(\theta_j(e_j^{out})^\rho + \theta_k(e_k^{out})^\rho)^{\frac{1}{\rho}} - d \cdot \frac{(e_j^{out})^2}{2}$, $j, k = 1, 2$ and $k \neq j$.

In the rest of the derivation we will write e_j and e_k instead of e_j^{out} and e_k^{out}. The first-order condition yields:

$$\alpha_{out}(\theta_j e_j^\rho + \theta_k e_k^\rho)^{\frac{1-\rho}{\rho}}\theta_j e_j^{\rho-1} - d e_j = 0,$$

$$\text{i.e.} \qquad \frac{\alpha_{out}}{d} \cdot \theta_j(\theta_j e_j^\rho + \theta_k e_k^\rho)^{\frac{1}{\rho}-1} = e_j^{2-\rho}. \tag{19}$$

Taking total differentiation, we have

$$\frac{\alpha_{out}}{d} \cdot \theta_j \cdot \frac{1-\rho}{\rho}(\theta_j e_j^\rho + \theta_k e_k^\rho)^{\frac{1}{\rho}-2}(\rho\theta_j e_j^{\rho-1}\frac{de_j}{de_k} + \rho\theta_k e_k^{\rho-1}) = (2-\rho)e_j^{1-\rho}\frac{de_j}{de_k}.$$

By rearranging, we obtain

$$\frac{de_j}{de_k} = \frac{\rho\theta_k e_k^{\rho-1}}{\frac{(2-\rho)e_j^{1-\rho}}{\frac{\alpha_{out}}{d}\theta_j\frac{1-\rho}{\rho}(\theta_j e_j^\rho+\theta_k e_k^\rho)^{\frac{1}{\rho}-2}} - \rho\theta_j e_j^{\rho-1}} = \frac{\rho\theta_k e_k^{\rho-1}}{\frac{(2-\rho)e_j^{1-\rho}}{e_j^{2-\rho}\frac{1-\rho}{\rho}(\theta_j e_j^\rho+\theta_k e_k^\rho)^{-1}} - \rho\theta_j e_j^{\rho-1}} \quad \text{(using (19))}$$

$$= \frac{\rho\theta_k e_k^{\rho-1}}{\frac{(2-\rho)(\theta_j e_j^\rho+\theta_k e_k^\rho)-\rho\theta_j e_j^{\rho-1}e_j\frac{1-\rho}{\rho}}{e_j\frac{1-\rho}{\rho}}} = \frac{\rho\theta_k e_k^{\rho-1}e_j\frac{1-\rho}{\rho}}{(2-\rho)(\theta_j e_j^\rho + \theta_k e_k^\rho) - (1-\rho)\theta_j e_j^\rho} = \frac{(1-\rho)\theta_k e_k^{\rho-1}e_j}{\theta_j e_j^\rho + (2-\rho)\theta_k e_k^\rho}.$$

For $\rho \in (-\infty, 1)$, $\frac{de_j}{de_k} > 0$, whereas for $\rho = 1$, $\frac{de_j}{de_k} = 0$. We can observe the following:

For the CES technology with $\rho \in (-\infty, 1)$, under output monitoring and linear incen-

tives agents' efforts are strategic complements.

References

Bag P.K. and Pepito N. (2012), "Peer transparency in teams: Does it help or hinder incentives?" *International Economic Review* 53, 1257–1286.

Bushman R.M., Indjejikian R.J., and Smith A. (1995), "Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies," *Journal of Accounting Research* 33, 101–128.

Bergstrom, T., Blume, L. and Varian, H. (1986), "On the private provision of public goods," *Journal of Public Economics* 29, 25–49.

Drago R. and Turnbell G.K. (1988), "Individual versus group piece rates under team technologies," *Journal of the Japanese and International Economies* 2, 1–10.

Herold F. (2010), "Contractual incompleteness as a signal of trust," *Games Economic Behavior* 68, 180–191.

Holmstrom B. (1982), "Moral hazard in teams," *Bell Journal of Economics* 13, 324–340.

Itoh H. (1991), "Incentives to help in multi-agent situations," *Econometrica* 59, 611–636.

Khalil F. and Lawarrée J. (1995), "Input versus output monitoring: who is the residual claimant?" *Journal of Economic Theory* 66, 139–157.

Legros P. and Matthews S. (1993), "Efficient and nearly-efficient partnerships," *Review of Economic Studies* 60, 599–611.

Marx L. and Matthews S. (2000), "Dynamic voluntary contribution to a public project," *Review of Economic Studies* 67, 327–358.

McAfee R.P. and McMillan J. (1991), "Optimal contracts for teams," *International Economic Review* 32, 561–577.

Mohnen A., Pokorny K., and Sliwka D. (2008), "Transparency, inequity aversion, and the dynamics of peer pressure in teams: Theory and evidence," *Journal of Labor Economics* 26, 693–720.

Nandeibam, S. (2002), "Sharing rules in teams," *Journal of Economic Theory* 107, 407–420.

Rahman D. (2012), "But who will monitor the monitor?" *American Economic Review* 102, 2267–2297.

Sappington D. (1983), "Limited liability contracts between principal and agent," *Journal of Economic Theory* 29, 1–21.

Varian H.R. (1992), *Microeconomic Analysis, 3rd edition.* New York: W.W. Norton & Company.

Winter E. (2006), “Optimal incentives for sequential production processes,” *Rand Journal of Economics* 37, 376–390.