Organizing for Cooperation: Fostering Relational Contracts to Support Interunit Knowledge Sharing

Nadav Levy*

June 21, 2022

Abstract

I study how a multiunit firm can promote knowledge sharing, creating conditions to foster the development of a relational contract that would support a mutual exchange of know-how between its units. Powered profit incentives are shown to be effective as they increase the future value of access to peer units' know-how and strengthens the incentives to cooperate with peers today. Linking the units' compensation to the performance of their peers improves the conditions for cooperation and further lower the cost of incentives. Relative performance evaluation is generally superior to joint performance evaluation as it increase the cost to a non-cooperative unit of being excluded from knowledge sharing between the other units in the future.

PRELIMINARY AND INCOMPLETE. COMMENTS WELCOME.

*Tiomkin School of Economics, Reichman Univeristy, Israel. E-mail: nadavl@idc.ac.il

1 Introduction

The transfer of knowledge within firms has long been recognized as a key to their success and competitive advantage. According to a recent Harvard Business Review article, Fortune 500 firms lose a combined \$31.5 billion per year from employees failing to share knowledge effectively.[1] A vast literature in the areas of Management, Organizational Behavior and Economics has studied the mechanisms of knowledge transfer and identified factors which support and impede the successful sharing of knowledge and innovations between units (e.g, Cool, Dierickx, and Szulanski, 1997; Argote and Ingram, 2000).[2]

Historically, organizations tended to rely on centralized management to coordinate and supervise the transfer of know-how between their units.[3] Over the last several decades, firms have aimed to improve initiative, response times, and performance by disaggregating, eliminating layers of middle managers and central functional staff, and moving an increasing amount of decision-making authority to the business units. These steps are often accompanied by strong monetary incentives, based on various measures of units' performance. However, because interdependencies between the units abound, it is crucial that the decentralization of the firm would not hamper internal cooperation and, in particular, the sharing of technical and managerial know-how and best-practice between its units. Roberts (2004) describes the organizational restructuring in BP Explorations. The firm has disaggregated, creating empowered business units with clear performance responsibilities. These measures "*had the intended effect of increasing initiative,... But the changes also created a great need for the business units to cooperate in sharing best practice and in supporting one another in solving technical and commercial problems – activities that were previously handled by the center, but that it now lacked resources to undertake.*"

[1] See, "Is Your Company Encouraging Employees to Share What They Know?," Christopher G. Myers, Harvard Business Review, November 2015.

[2] Argote (2013) notes that one reason for the interest in knowledge transfers is the growing importance of multinational firms. According to Argote, "effective operation of global firms requires that distributed expertise be coordinated—that knowledge be transferred from one expert to another and from one site to another." In addition, the growth of the multiunit, multimarket organizational form, which includes franchises and chains in the services, hospitality and health care sectors as another trend that contributed to interest in knowledge transfers. Such multiunit organizations produce the same product or deliver the same service in different markets and thus knowledge acquired at one establishment can be highly relevant to others.

[3] Argote, Beckman, and Epple (1990) discuss the production of the Liberty ship during World War II. Production was carried out by 13 shipyards that went into production in different points in time. While shipyards were independent, "a central agency was responsible for purchasing, designing each yard's layout, approving the technology it used, and supervising its construction. The central agency also stationed engineers, inspectors, and auditors at each site to share information about "best practices"."

Because flatter firms have fewer dedicated personnel to collect knowledge from the units and disseminate it within the firm, the burden of seeking their peers' knowledge and communicating their own knowledge to their peers now falls directly on the units themselves. A potential concern is that units would fail to share their knowledge, especially if they are given powered performance-based incentives. Profit-based incentives may hinder cooperation and knowledge sharing between units, as that they increase the opportunity cost for the units of the time and resources spend helping their peers. A potential solution is to foster a cooperative relational contract between the units, an informal agreement to share their knowledge.[4] The repeated interaction between the units can create implicit incentives for cooperation by leveraging the units' ability to monitor their peers and sanction noncooperative behavior.[5]

In this paper I study how multiunit firms can foster the development of a relational contract between their units in order to support interunit knowledge sharing. In the model, units constantly come up with process innovations, improved production methods, and new work practices. These innovations are valuable also to other units within the firm and can improve their productivity■ and the firm thus has an interest in disseminating them as broadly as possible. Knowledge sharing requires the units' managers to exert effort communicating their knowledge to their peers. The firm's center does not observe communication and thus it cannot secure such effort by fiat, nor compensate the units directly for exerting it. However, because communication is observable by the units themselves, a relational contact between the units could generate *implicit* incentives for cooperation that can substitute for the missing *explicit* incentives. Under the relational contract, all units are expected to share their know-how today in order to secure the cooperation of their peers in the future. A non-cooperative unit would be excluded from future cooperation with its peers and would suffer a lower productivity as a result.

I show that firms can use various features of the organizational design to strengthen their units' ability to self-discipline noncooperative peers. I first consider how firms can use their compensation system to foster cooperation. Both the power of incentives and

[4] Baker, Gibbons, and Murphy (2002) define a relational contracts as "informal agreements, sustained by the value of future relationships." Informal quid pro quos between coworkers which are the subject of this paper, are one of the examples which they mention.

[5] According to Roberts, BP Explorations was able to overcome the challenges arising from the organizational restructuring and nurture a strong culture of cooperation between its units. As he writes, *"The solution was to link collections of business units facing similar technical and commercial challenges into peer groups. The members of these peer groups developed norms of sharing information and helping one another resolve problems. It became a strong norm to ask for help in trouble and to respond wholeheartedly if asked for help. This cooperation was crucial to BP's success."*

the link between the managers' compensation and the performance of their peer have a positive impact on the stability of the relational contract between the units. When compensation is based exclusively on the unit's performance, or *independent performance evaluation* (IPE), powered monetary incentives increase the value of future technological cooperation with peer units and strengthen the implicit incentives to cooperate generated by the relational contract. This stands in contrast to a common intuition that powered incentives hamper cooperation. The power of incentives is tied to the stability of the relational contract. When units are more patient, the firm relies more heavily on the relational contract and awards the units' managers a smaller share of their units' profits.

Next, I consider alternatives forms of compensation in which the units' compensation is tied to the performance of their peers: *joint performance evaluation* (JPE), where compensation increases with peers' performance and *relative performance evaluation* (RPE), where compensation decreases with peers' performance. It has been argued group incentives can promote a climate of trust and cooperation which is conductive to organizational knowledge sharing (Collins and Smith, 2006). In my model, the firm prefers JPE to IPE since it lowers the monetary cost of the incentives that are required to sustain the relational contract. Under JPE, the immediate gains from deviation are lower than under IPE since the deviator's compensation decreases due to the adverse effect of holding on to its knowledge on the performance of its peers. On the other hand, concerns have been raised that the use of RPE can have on the undesired effect on knowledge sharing.[6] While RPE is detrimental to knowledge sharing in the static setup, its no longer the case under the relational contract. With three units or above RPE is superior to IPE. Excluding only the deviator from future knowledge sharing, while maintaining the cooperation among the other units, the deviator's punishment is harsher under RPE, since its performance falls below that of his peers. Moreover, the cost of incentives to the firm under RPE is also lower than under JPE. The firm is better off using RPE, except for when the units are too impatient so that the full sharing equilibrium cannot be sustained with RPE but is feasible with JPE.

In addition to the compensation scheme, firms can leverage other aspects of the organizational design to support knowledge sharing between their units. I show that allowing the units to compete over clients strengthen the incentives for knowledge sharing between the units. Similarly to the effect of RPE, the introduction of interunit competition over

[6] For example, Sandvik, Saouma, Seegert, and Stanton (2020) consider knowledge flows between sales agents. They note that in many sales firms a portion of compensation depends on performance relative to other employees, potentially increasing opportunity costs of helping others.

clients, severely disadvantages units in the case that they are excluded from the relational contract, as they have to compete with higher-productivity peers.

Early contributions to the theory of incentives have considered short-term relationships between employees. Holmstrom (1982), Lazear and Rosen (1981) and others have shown how relative performance evaluation and rank-order tournaments between agents can be used to filter away common uncertainties that affect the performance of all agents. Lazear (1989) shows that the value of relative performance evaluation is diminished when agents engage both in individual and other team-related tasks. Other papers have demonstrated the value of group incentives when explicit side trading between the agents is feasible (Holmstrom and Milgrom, 1990) and when the principal wants to motivate agents to help each other (Itoh, 1991).

Closer to this paper, later work have considered repeated long-term interactions between agents. Che and Yoo (2001) study how the explicit incentives chosen by the firm interact with the implicit incentives that arise from the repeated interaction between its employees. In their setup, relative performance evaluation is optimal in the static setup as it allows to filter to a common shock to the productivity of the agents but with repeated interaction it cannot generate implicit incentives for high effort, whereas group performance evaluation can generate such incentives. Che and Yoo show that group performance evaluation is optimal when the agents are sufficiently patient, whereas independent performance evaluation or relative performance evaluation are optimal if the agents are not patient enough. In contrast, my framework considers effort exerted in order to share knowledge with peers. Relative performance evaluation is superior to other compensation forms in the repeated interaction as long as there are 3 or more agents and the agents are sufficiently patient. Another strand of papers (Rob and Zemsky, 2002; Spagnolo, 1999) use dynamic models to study how the firm can affect the accumulation of "social capital" between its employees and how it can use such capital to promote them to help their peers.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 considers the effect of the compensation scheme on knowledge sharing. Section 4 discusses additional organizational variables that can support knowledge sharing.

2 The model

A firm has a set of units $\mathcal{N} = \{1, ..., n\}$ where $n \geq 2$. In each period $t = 1, 2, ..., \infty$, the units manufacture a new generation of products. At the beginning of each period each unit comes up with a single cost-reducing innovation. This innovation is also valuable to the unit's peers if shared with them. Denote by $\pi(k)$ the profit of a unit when it has access to k of the n current-generation innovations, where $\pi'(k) \geq 0$. Denote the marginal value of the k'th innovation by $\Delta(k) \equiv \pi(k) - \pi(k-1)$ and assume that $\Delta'(k) \leq 0$. Units' profits are observable by all and verifiable, and thus can be contracted upon.

Each unit is headed by a manager who bears a private disutility from the effort required to share the innovation to its peers.[7] Denote by $\Psi(m)$ the disutility of communicating to m peers, where $\Psi'(m) \geq 0$ and $\Psi''(m) \leq 0$. Manager i's utility in period t is $u_{it} = w_{it} - \Psi_{it}$, where w_{it} is manager i's wage and Ψ_{it} his disutility from communication. The manager has an outside option worth u_0 per period, where $u_0 < \pi(1)$, and discounts future utility streams with a discount factor δ.

The next assumption implies that the communication of all innovations by all n units is efficient.

A1 $(n-1)\Delta(n) > \Psi(n-1)$.

The total benefit of unit i sharing its innovation with all of its peers if all other units share their innovations, $(n-1)\Delta(n)$, is higher than the disutility from communication, $\Psi(n-1)$. The concavity of $\pi(k)$ implies that the benefits of sharing the innovation are yet higher if any of i's peers do not share their innovations.

The firm compensates the managers using either a fixed wage or with an incentive contract that depends on unit i's profits π_i and on the profile of its peers' profits, π_{-i}. Managers receive identical wage contracts that are invariant over time. Each period's wage depends only on the current profits. Thus, $w_{it} = w(\pi_{it}, \pi_{-it})$ for all $i = 1, ..., n$ and $t = 1, 2, ..., \infty,$. We consider three alternative incentive schemes:

i. Independent performance evaluation (IPE) - In addition to a base wage b, each unit manager is paid a share α_I of his unit's profit. The manager's compensation is independent of its peers' performance. Thus, $w(\pi_i, \pi_{-i}) = b + \alpha_I \pi_i$.

ii. Joint performance evaluation (JPE) - Each unit manager is paid a share α_J of the average profit of all n units, $w(\pi_i, \pi_{-i}) = b + \alpha_J \overline{\pi}$ where $\overline{\pi} = \left(\sum_{k=1}^{n} \pi_k\right) \backslash n$. Under JPE, the manager's wage increases in the performance of its peers.

[7] Communication effort may be exerted in various ways. The manager may oversee the codification of new knowledge so it can be accessible by other units. Alternatively, he may spend time in various forms of informal communication such as internal presentations, etc.

iii. Relative performance evaluation (RPE) - each unit manager is paid $w(\pi_i,\pi_{-i}) = b + \alpha_R [\pi_i + z(\pi_i - \overline{\pi}_{-i})]$ where $z > 0$ and $\overline{\pi}_{-i} = \frac{1}{n-1} \sum_{j \neq i} \pi_j$. A share α_R is paid of the unit's profit that is adjusted by $z(\pi_i - \overline{\pi}_{-i})$ to reflect the unit's relative performance compared to the average performance of its peers. Under RPE, the manager's pay decreases in the performance of its peers.

Consider first the one-period stage game. In the static setup it is never unilaterally beneficial for a manager to exert effort under either IPE or RPE. While communication is privately costly to the manager, it has no effect on the unit's compensation under IPE, and a negative effect on compensation under RPE. Thus, in the unique stage game equilibrium under these two compensation schemes, no unit communicates and the managers wage is $w_{it} = u_0$ in each period.

In contrast, under JPE, it can be beneficial for a unit to share its innovations with their peers. Suppose each one of unit i's peers has access to $m-1$ other innovations, where $1 \leq m \leq n-1$, whereas unit i itself has access to k innovations. Then, manager i finds it optimal to share its knowledge with his peers if

$$\alpha_J \cdot \frac{\pi(k) + (n-1)\pi(m)}{n} - \Psi(n-1) \geq \alpha_J \cdot \frac{\pi(k) + (n-1)\pi(m-1)}{n}$$

or if

$$\alpha_J \cdot \frac{n-1}{n} \Delta(m) \geq \Psi(n-1).$$

For clarity of exposition, and to focus on the role of dynamic incentives, we rule out this possibility. Since $\Delta(m)$ is decreasing in m, a sufficient condition (though stronger than needed) is

A2 $\Delta(2) < \Psi(n-1)$ for all $n \geq 3$.

Assumption A2 implies that for the average cost of communication $\Psi(n-1)/(n-1)$ per peer is higher than an equal share of the benefits, $\Delta(n)/n$. Since under JPE the units share the benefits from communication equally, the disutility from communication is too high for it to be unilaterally beneficial. Hence under Assumption A2 no static equilibria with communication exists under JPE as well.

3 Knowledge sharing and the pay scheme

I now turn to consider a symmetric subgame-perfect equilibria of the repeated game in which on the equilibrium path, all units share their innovations with all of their peers in all periods (a *full sharing* equilibrium). Adherence to the equilibrium is supported by an implicit threat to revert to a non-cooperative (punishment) equilibrium in case one of the units would deviate. I show below that under our assumptions, the full sharing equilibrium maximizes the managers' payoffs given the compensation schemes set the firm. The exact equilibrium strategies are presented in the sections below.

Denote by u^C the per-period utility of a manager on the equilibrium path, by u^D the instantaneous utility to a manager that deviates from the equilibrium play and by u^P the deviator average payoff in the ensuing periods (the punishment subgame). The manager's incentive constraint can be written as follows:

$$u^C \geq (1-\delta)\, u^D + \delta u^P. \tag{1}$$

Abreu (1988) shows that every perfect equilibria path is the outcome of equilibrium strategies which specify the same player-specific maximal punishment after any deviation by that player. We characterize the full sharing equilibria below following Abreu's insight.

We begin by considering the equilibrium under Independent performance evaluation (IPE). We then consider in turn Joint performance evaluation (JPE) and Relative performance evaluation (RPE) and compare each alternative with IPE. Finally we consider the optimal pay scheme.

3.1 Independent performance evaluation

Under IPE, a manager who is paid a share α_I of its unit's profit can unilaterally guarantee himself a utility $\alpha_I \pi\,(1)$ in every period when not sharing its innovation with its peers. Thus, $\alpha_I \pi\,(1)$ is a lower bound on the manager's payoff in any perfect equilibrium. The maximal punishment equilibrium is therefore one in which all units revert to play the Nash equilibrium of the stage game in all periods. Since the units do not share their innovations, this equilibrium attains the lower bound, and hence $u^P = b + \alpha_I \pi\,(1)$. The manager's per-period utility in the full sharing equilibrium, is $u^C = b + \alpha_I \pi\,(n) - \Psi\,(n-1)$. Finally, the manager's instantaneous utility if deviating from the equilibrium and not sharing its innovation, but receiving the innovations of all of its peers is $u^D = b + \alpha_I \pi\,(n)$.

Substituting these terms into (1) and rearranging, the incentive constraint becomes

$$\delta \geq \frac{\Psi (n-1)}{\alpha_I [\pi (n) - \pi (1)]}.$$

Since the right-hand side is decreasing in α_I,the incentive constraint is more likely to hold the higher is the profit share, Thus, **powered incentives helps knowledge sharing between the units**. Intuitively, when α_I is too small, knowledge sharing cannot be sustained. As α_I increases, the net immediate gain from deviation, $u^D - u^C$, is unaffected, since it equals the private cost of communication, $\Psi (n-1)$. By contrast, the future loss per period rises, as the difference between the cooperative and punishment payoffs to the manager, $u^C - u^P$, increases with α_I.

Next, rewrite the incentive constraint to express the minimal profit share $\underline{\alpha}_I^{ic}$ for which it is satisfied, given the discount factor δ. The full sharing equilibrium incentive constraint is satisfied if and only if $\alpha_I \geq \underline{\alpha}_I^{ic}$, where

$$\underline{\alpha}_I^{ic} \equiv \frac{\Psi (n-1)}{\delta [\pi (n) - \pi (1)]}. \tag{2}$$

Note that $\underline{\alpha}_I^{ic} < 1$ if and only if $\delta > \delta_1$, where $\delta_1 \equiv \Psi (n-1) / [\pi (n) - \pi (1)]$, and where $\delta_1 < 1$, as $\pi (n) - \pi (1) = \sum_{k=2}^{n} \Delta (k) > (n-1) \Delta (n) > \Psi (n-1)$, where the first inequality follows as $\Delta (k)$ decreases in k, and the last inequality follows from Assumption A1.

In addition, the manager's participation constraint is satisfied provided $u^C \geq u_0$. For the special case in which $u_0 = 0$, the participation constraint is implied by the incentive constraint since $u^C \geq (1-\delta) u^D + \delta u^P$ and as $u^D \geq 0$ and $u^P \geq 0$.

To induce the units to share their innovations at the lowest cost, the firm chooses a base wage b_I and a profit share α_I that minimizes the managers' compensation under the participation and incentive constraints, i.e.,

$$\min_{\alpha, b} b + \alpha \pi (n)$$

$$s.t.$$

$$\begin{aligned} \alpha \ & \geq \underline{\alpha}_I^{ic} \\ b + \alpha \pi (n) - \Psi (n-1) & \geq u_0 \end{aligned}$$

While there many possible solutions to the minimization problem above, I focus without loss of generality on the one in which the incentive constraint binds, i.e., $\alpha_I = \underline{\alpha}_I^{ic}$ and

$b_I = max\{u_0 + \Psi(n-1) - \underline{\alpha}_I^{ic}\pi(n), 0\}$. If $\underline{\alpha}_I^{ic}\pi(n) > u_0 + \Psi(n-1)$, the participation constraint is slack and $b_I = 0$. Otherwise, the participation constraint binds and $b_I > 0$.

Finally, one has to check the conditions under which providing the managers with incentive to share their knowledge is optimal. If not incentivizing knowledge sharing, the firm's gross profit per unit is $\pi(1)$ and it pays the manager his reservation wage u_0. Incentivizing knowledge sharing is desirable if

$$\pi(n) - [b_I + \alpha_I \pi(n)] \geq \pi(1) - u_0.$$

Clearly, incentivizing knowledge sharing is optimal when the participation constraint binds, as then the wage bill is the same in both cases. Thus, we only need to check when the condition hold when the participation constraint is slack and $b_I = 0$. In that case, the inequality above implies $\underline{\alpha}_I^{ic}\pi(n) - u_0 \leq \pi(n) - \pi(1)$, which is satisfied if and only if $\delta \geq \underline{\delta}_I$, where

$$\underline{\delta}_I \equiv \delta_1 \cdot \frac{\pi(n)}{\pi(n) + u_0 - \pi(1)},$$

and where $\underline{\delta}_I > \delta_1$ since $u_0 < \pi(1)$. At $\underline{\delta}_I$, the total wage paid to the manager $\underline{\alpha}_I^{ic}(\underline{\delta}_I)\pi(n) = u_0 + \pi(n) - \pi(1) > u_0 + \Psi(n-1)$, exceeds the manager's reservation value.

The next lemma summarizes the results under IPE.

Lemma 1 *When the firm uses independent performance evaluation (IPE) and when $\delta < \underline{\delta}_I$, it pays units' managers his reservation wage u_0 and does not incentivize knowledge sharing. When $\delta > \underline{\delta}_I$, the firm incentivizes units' managers to share their knowledge. At $\underline{\delta}_I$, the manager's utility exceeds u_0. The managers' profit share, $\underline{\alpha}_I^{ic}$ and the total wage are decreasing in δ.*

Figure 1 below plots the equilibrium wage w as a function of the discount factor δ

[[Figure 1 Here]]

In our framework, communication effort is non-verifiable and cannot be compensated directly by the firm and in the stage game, knowledge sharing is not feasible. However, in the repeated game, a sufficiently high profit share can induce a relational contract between the units with implicit incentives to support knowledge sharing. While the managers' effort is not compensated directly, if they choose not to share their innovations today they would forego the monetary value of the productivity gains from future knowledge sharing. Explicit and implicit incentives are thus complements in the sense that the firm awards the unit a share of their unit's profits only in when the units can sustain a relational contract between themselves. However, under the relational contract, the two are substitutes. As unit managers become more patient, the incentive constraint is satisfied with an increasingly smaller profit share. The firm takes advantage of the implicit incentives generated by the relational contract to reduce the managers' explicit pay.

A key insight from the literature on multitasking in static environments (Holmstrom and Milgrom, 1991) is that low-powered incentives can be optimal when agents multitask and preform concurrently both easy to measure and hard to measure tasks (such as helping peers and knowledge sharing). Multitasking considerations aside, the result here can be viewed as complementary, suggesting that in repeated setups high-powered incentives can encourage agents to take hard to measure tasks such as knowledge sharing.

3.2 Joint performance evaluation

Under JPE, managers compensation is based on the average profit of all units, $w = b_J + \alpha_J \overline{\pi}$. Although under JPE each manager gains directly from an improvement in its peers' performance, we have seen in Section 2 that, under our assumptions, the cost of communication is such that in the unique Nash equilibrium the units do not share their innovations. Thus, as in the case of IPE, the maximal punishment for a deviation is attained by a reversion to an infinite play of the Nash equilibrium of the stage game, and hence $u^P = b_J + \alpha_J \pi(1)$. Similarly, in the full sharing equilibrium, all units share their innovations and have identical profits, and thus the average profit is $\overline{\pi} = \pi(n)$. Hence $u^C = b_J + \alpha_J \pi(n) - \Psi(n-1)$. Finally, the manager's one-period payoff from deviation is

$$u^D = b_J + \alpha_J \left[\frac{\pi(n) + (n-1)\pi(n-1)}{n} \right].$$

Note that by Assumption A2, $u^D - u^C = \Psi(n-1) - \alpha_J \frac{n-1}{n} \Delta(n) > 0$. Still, u^D is lower under JPE than under IPE, as the deviator's compensation takes a hit as a result of the decrease in its peers' profitability when it does not share its innovation.

Substituting u^C, u^D and u^P into (1) and rearranging, the minimal profit share that satisfies the incentive constraint under JPE is

$$\underline{\alpha}_J^{ic} \equiv \frac{\Psi(n-1)}{\delta[\pi(n) - \pi(1)] + (1-\delta)\frac{n-1}{n}\Delta(n)}.$$

$\underline{\alpha}_J^{ic}$ is decreasing in δ. As $\underline{\alpha}_J^{ic} < \frac{\Psi(n-1)}{\delta[\pi(n)-\pi(1)]} = \underline{\alpha}_I^{ic}$, the incentive constraint under JPE can be sustained for a lower profit share than under IPE, for all levels of the discount factor

δ.

Finally, to incentivize knowledge sharing, the firm chooses a wage (α_J, b_J) to solve

$$\min_{\alpha,b} b + \alpha\pi(n)$$

$$s.t.$$

$$\begin{aligned}\alpha &\geq \underline{\alpha}_J^{ic} \\ b + \alpha\pi(n) - \Psi(n-1) &\geq u_0\end{aligned}$$

As above, we focus on the solution in which the incentive constraint binds, i.e., $\alpha_J = \underline{\alpha}_J^{ic}$ and $b_J = max\{u_0 + \Psi(n-1) - \underline{\alpha}_J^{ic}\pi(n), 0\}$. Incentivizing knowledge sharing is optimal if $\pi(n) - [b_J + \underline{\alpha}_J^{ic}\pi(n)] \geq \pi(1) - u_0$. Again, we only need to check the case in which the participation constraint is slack and $b_J = 0$, implying $\underline{\alpha}_J^{ic}\pi(n) \leq u_0 + \pi(n) - \pi(1)$. Since $\underline{\alpha}_J^{ic} < \underline{\alpha}_I^{ic}$ for all δ and as $\underline{\alpha}_I^{ic}$ and $\underline{\alpha}_J^{ic}$ are decreasing in δ, this condition is satisfied if $\delta > \underline{\delta}_J$, where $\underline{\delta}_J < \underline{\delta}_I$. Finally, for future use, denote by $\overline{\delta}_J$ a threshold value such that $\underline{\alpha}_J^{ic}\pi(n) = u_0 + \Psi(n-1)$ if and only if $\delta > \overline{\delta}_J$.

The next lemma follows immediately from the results above.

Lemma 2 *For the firm, JPE weakly dominates IPE. When $\delta \in [\underline{\delta}_J, \underline{\delta}_I]$, incentivizing knowledge sharing is optimal with JPE but not with IPE. The cost of incentives is strictly lower under JPE than under IPE, whenever $\underline{\alpha}_I^{ic}\pi(n) > u_0 + \Psi(n-1)$.*

3.3 Relative performance evaluation

Recall that under RPE, all unit managers are paid $w = b_R + \alpha_R[\pi_i + z(\pi_i - \overline{\pi}_{-i})]$ where $z > 0$ and $\overline{\pi}_{-i} = \frac{1}{n-1}\sum_{j\neq i}\pi_j$. In the full sharing equilibrium, $\pi_i = \pi(n)$ for all $i \in N$ and thus $u^C = b_R + \alpha_R\pi(n) - \Psi(n-1)$. Compared with IPE, the gain to deviation under RPE is higher. A deviating manager who does not share its unit innovation both avoids the communication disutility, and gains a one-period productivity advantage over its peers, which under RPE boosts his compensation. As for the ability to punish deviations under RPE, it depends on the number of units. With only two units, the punishment stage payoffs are identical to the case of IPE and thus it is clear that RPE would not be used when $n = 2$. With $n \geq 3$ units however, a deviation can be punished more harshly under RPE than under IPE. To lower a deviator's payoff as much as possible, all other units should discontinue sharing their innovations with the deviator, while continuing to cooperate among themselves, thus maximizing the productivity gap between themselves and the deviator. The deviator's payoff in this case is

$$\underline{u} = b_R + \alpha_R \left[\pi(1) + z(\pi(1) - \pi(n-1))\right].$$

Hence, $\underline{u}$ is a lower bound to a deviator payoff in any perfect equilibrium. In Appendix A.2, I characterize a maximal punishment subgame perfect equilibrium σ^k in which the average per period payoff of unit k equals the lower bound $\underline{u}$. Under σ^k, unit k is expected to cooperate with its own punishment, and continue to provide its know-how to its peers' throughout the T periods of punishment, during which it is excluded from their know-how. Unit k's cooperation is achieved using a "carrot and stick" strategy: After T periods unit k is forgiven and brought back into cooperation with its peers, But the punishment would be restarted from period 1 if k fails to cooperate in any of the first T periods. k's discounted payoff under σ^k is exactly $\underline{u}$. While simpler trigger strategies that cease all cooperation with k forever following a deviation would also attain the lower bound, the payoff to all other units in $\mathcal{N} \setminus \{k\}$ under σ^k is higher, a fact which makes it easier to incentivize them to adhere to the punishment strategies.

Next consider the full equilibrium. On the equilibrium path all units are expected to share their innovations with all of their peers in all periods. A deviation by unit k leads all units to play the punishment equilibrium σ^k in all future periods. A unit deviating and not sharing its innovation with any of its peers gets an immediate payoff

$$u^D = b_R + \alpha_R \left[\pi(n) + z\Delta(n)\right].$$

in the period of deviation. Substituting u^D above and $u^P = \underline{u}$ into the incentive constraint (1) yields

$$u^C \geq (1-\delta)\,\alpha_R \left[\pi(n) + z\Delta(n)\right] + \delta \underline{u}. \tag{3}$$

Compared with IPE, and for $\alpha_R = \alpha_I$, RPE increases the deviation payoff u^D by $\alpha_R z \Delta(n)$ but lowers the punishment payoff $\underline{u}$ by a larger sum, $\alpha_R z \left[\pi(n-1) - \pi(1)\right]$. Intuitively, if δ is sufficiently large, the lower punishment payoff $\underline{u}$ looms larger than the higher deviation payoff u^D and the overall effect on the right-hand side of the incentive constraint is negative. To see this formally, note that z appears only on the right-hand side of (3), which decreases with z when $\Phi(\delta) \equiv (1-\delta)\,\Delta(n) + \delta\left[\pi(1) - \pi(n-1)\right] < 0$. Since $\Phi(\delta)$ is continuous, $\Phi(0) > 0$ and $\Phi(1) < 0$ and since $\Phi'(\delta) < 0$, it follows that

$\Phi(\delta) < 0$ if and only if $\delta > \underline{\delta}_R$, where

$$\underline{\delta}_R \equiv \frac{\Delta(n)}{\pi(n) - \pi(1)}.$$

and where $\underline{\delta}_R < 1$. RPE thus strengthen incentives for cooperation if and only if $\delta > \underline{\delta}_R$.

From the incentive constraint we can also derive the minimal profit share that satisfies the constraint $\underline{\alpha}_R^{ic}(z)$, as a function of z. This equals

$$\underline{\alpha}_R^{ic}(z) \equiv \frac{\Psi(n-1)}{\delta[\pi(n) - \pi(1)] - \Phi(\delta) z}.$$

Note that $\underline{\alpha}_R^{ic}(0) = \underline{\alpha}_I^{ic}$ and that $\underline{\alpha}_R^{ic}(z) < \underline{\alpha}_I^{ic}$ and is decreasing in z if and only if $\Phi(\delta) < 0$.

Finally, to incentivize the full sharing equilibrium, the firm chooses a compensation scheme $(b_{R,}\alpha_R, z_R)$ to solve

$$(b_R, \alpha_R, z_R) = \min_{b,\alpha,z} b + \alpha\pi(n)$$

$$s.t.$$

$$\begin{aligned} &\alpha \ \geq \underline{\alpha}_R^{ic}(z) \\ &b + \alpha\pi(n) - \Psi(n-1) \geq u_0 \end{aligned}$$

As in the previous sections, and without loss of generality, let $\alpha_R = \underline{\alpha}_R^{ic}(z)$ and $b_R = max\{u_0 + \Psi(n-1) - \underline{\alpha}_R^{ic}(z)\pi(n), 0\}$.

The next lemma shows that RPE both expands the scope of the full sharing equilibrium compared with IPE and lowers the cost of incentives to the units' managers. Incentivizing knowledge sharing is always optimal whenever the use of RPE supports the relational contract.

Lemma 3 *Under RPE and when $\delta \geq \underline{\delta}_R$, the center chooses $b_R = 0$, $\alpha_R = [u_0 + \Psi(n-1)]/\pi(n)$ and z_R such that $\alpha_R = \underline{\alpha}_R^{ic}(z_R)$. When $\delta \in [\underline{\delta}_R, \underline{\delta}_I]$, incentivizing knowledge sharing is optimal with RPE but not with IPE. The cost of incentivizing knowledge sharing with RPE is strictly lower than with IPE when $\underline{\alpha}_I^{ic}\pi(n) > u_0 + \Psi(n-1)$. When $\delta < \underline{\delta}_R$, the center does not incentivize the managers to share their innovations under either RPE or IPE.*

Proof. Note first that by Assumption A1 it follows that $\underline{\delta}_R < \underline{\delta}_I$. When $\delta \geq \underline{\delta}_R$ and as $\underline{\alpha}_I^{ic}\pi(n) > u_0 + \Psi(n-1)$, then by setting $\alpha_R = [u_0 + \Psi(n-1)]/\pi(n)$, and z_R sufficiently high so that $\underline{\alpha}_R^{ic}(z_R) = \alpha_R$ the firm reduces the managers' compensation under RPE to

$u_0 + \Psi(n-1)$, below its level under IPE. Providing incentives for knowledge sharing is clearly optimal in this case. By contrast, when $\delta \leq \underline{\delta}_R$, a relative evaluation component $z > 0$ increases the cost of incentives and should be avoided. Since $\underline{\delta}_R < \underline{\delta}_I$, then by Lemma 1, the firm prefers not to incentivize knowledge sharing in that case. ■

For $\underline{\delta}_R \leq \delta < \underline{\delta}_I$, the center incentivizes the full sharing equilibrium with RPE but not with IPE. Whenever $\delta \geq \underline{\delta}_I$, full sharing is optimal with both compensation schemes, but the managers' compensation is lower with RPE, as long as $\underline{\alpha}_I^{ic}\pi(n) - \Psi(n-1) > u_0$. Only when δ is sufficiently high for the participation constraint to bind under IPE is the cost of incentives the same under RPE and IPE. The level of the relative-performance component z_R does not change the managers' compensation directly, since on the equilibrium path the units' profits do not diverge. However, the use of RPE enables the firm to lower the managers' compensation as it increases the (off-equilibrium) cost to a deviator of being excluded from the innovations of other units in the future.

To finish this section, some comments are in order. First, the (off-equilibrium) wage of a deviating manager in the punishment stage is negative when $\pi(1) + z(\pi(1) - \pi(n)) < 0$ or when $z > \overline{z} \equiv \frac{\pi(1)}{\pi(n) - \pi(1)}$. If the manager cannot pay the firm ex post in case his performance falls behind, the manager's limited liability would constrain the power of the relative performance component that can be used to $z \leq \overline{z}$. This would not change the results of Lemma 3 as long as $\underline{\alpha}_R^{ic}(\overline{z}) \leq [u_0 + \Psi(n-1)] / \pi(n)$.

Second, the effectiveness of using RPE can be further enhanced if the compensation scheme would penalize poor performance harsher than it rewards good performance. A "neutral" way to implement this the current context is to have the relative component of the wage (either a reward or a fine) kick in only if the difference between a unit's performance and that of its peers exceed a threshold $\widetilde{\Delta}$, where $\Delta(n) < \widetilde{\Delta} < [\pi(n) - \pi(1)] / (n-1)$. A deviating unit would then not receive an additional compensation as a result holding on to its innovation (as its productivity advantage $\Delta(n)$ would be below the threshold $\widetilde{\Delta}$) but would be fined in the punishment equilibrium when the difference in performance compared to its peers' is larger. Under such scheme, the scope of the full sharing equilibrium under RPE would be further extended and the cost of incentives would be diminished.

Third, the firm could achieve an incentive structure similar to RPE if it would set performance goals for the units based on the outcomes of the high performers in the previous period. In the punishment stage, a deviator would be unable to reach the performance standard without access to its peers' know-how and accordingly would be fined.

3.4 Optimal pay schemes

The next proposition combines the results of the previous sections to characterize the optimal pay scheme for the unit managers.

Proposition 1 *With* 2 *units, JPE is optimal. With* 3 *or more units,*

i. If $\delta > \overline{\delta}_J$ either RPE or JPE are optimal. The managers' utility is u_0.

ii. If $\delta \in \left(\underline{\delta}_R, \overline{\delta}_J\right)$ RPE is optimal. Managers' utility is u_0 and the cost of incentives to the firm is strictly lower than with JPE.

iii. If $\underline{\delta}_J < \underline{\delta}_R$, then for $\delta \in [\underline{\delta}_J, \underline{\delta}_R]$, JPE is optimal. Managers' utility exceeds u_0.

iv. If $\delta < \min\{\underline{\delta}_J, \underline{\delta}_R\}$, the managers are paid a fixed wage u_0 and do not share knowledge.

Proof. As by Lemma 3, JPE weakly dominates IPE, it is optimal for $n = 2$. For $n \geq 3$ we need to consider the choice between RPE and JPE.

i. If $\delta > \overline{\delta}_J$ the total wage bill is $u_0 + \Psi(n-1)$ under either JPE or RPE . Thus, either one is optimal.

ii. If $\delta \in \left(\underline{\delta}_R, \overline{\delta}_J\right)$, the total wage bill under RPE is $u_0 + \Psi(n-1)$ and higher under JPE. Thus, RPE is optimal for the firm.

iii. If $\underline{\delta}_J < \underline{\delta}_R$, then for $\delta \in [\underline{\delta}_J, \underline{\delta}_R]$, RPE cannot support the full sharing equilibrium and hence JPE is optimal. Since $\delta < \overline{\delta}_J$ the manager's total wage exceeds $u_0 + \Psi(n-1)$.

iv. If $\delta < \min\{\underline{\delta}_J, \underline{\delta}_R\}$, a full sharing is not feasible under RPE and not beneficial under JPE. The firm then pays the manager a fixed wage u_0 and does not incentivize knowledge sharing.

■

Linking the unit managers' pay to the performance of their peers is beneficial either because it limits the deviation profits (under JPE), or lowers the punishment equilibrium payoffs (under RPE). While the cost of incentives is lower under RPE, it can only be sustained when the units are patient enough for the lower (future) punishment profits to outweigh the higher (immediate) deviation profits. Under RPE the firm lowers the managers' profit share and increases the relative-pay component. While in equilibrium all units profits are the same and no unit is actually paid a bonus (fined) for exceptional (weak) relative performance, the use RPE gives rise to a relational contract in which the punishment for deviation is more severe. In contrast, when the units are relatively impatient, RPE hinders knowledge sharing and therefore should not be used. In some cases, the use of JPE may allow the firm to maintain cooperation in this case.

4 Additional aspects of organizational design that support knowledge sharing

4.1 Interunit competition over clients

According to Tsai (2002) "W*hile multiunit organizations encourage interunit knowledge sharing to realize economies of scope, they also allow interunit competition to achieve efficiency. Different organizational units not only collaborate with each other to share knowledge, but also compete with each other to maximize their own benefit. Internally, they vie for limited resources within the organization. Externally, they try to outperform other units that offer similar products or services in the marketplace."*

While it may seem that interunit competition can discourage knowledge sharing between the units, I argue here that interunit competition on clients can be part of an organizational design intended to support knowledge sharing between the units. Granting units exclusivity on their clients shields the units from their peers, even if their performance lags behind. In contrast, allowing the units to compete internally over clients would lower the "market share" of under-performing units. Under a knowledge sharing relational contract between the units, an internal competition over clients increases the vulnerability of a unit to being excluded from sharing knowledge with its peers. This in turn helps sustaining the full sharing equilibrium. Internal competition thus acts similarly to the use of relative performance evaluation (RPE) in the previous section.

To see how this idea can be modeled, assume that the firm has a fixed client base, normalized for the simplicity to a measure n, as the number of units. Each unit is assigned an equal number of clients (measure 1), but the firm allows competition over a share $\theta \in [0, 1]$ of them ("competitive clients"), while the unit retains exclusivity over the rest, $1 - \theta$. To compete on its peer's competitive clients, a unit must pay a small positive fixed cost. This assumption implies that a unit would only compete on its peer's clients when it has a strict cost advantage. Absent competition, a unit charges the client a price p that equals its willingness to pay. If two units or more compete on a client, the lowest cost unit wins at a price that equals the cost of the second lowest cost unit. Denote a unit's cost when having access to m innovations by c_m where $c_1 > c_2 ... > c_n$. For simplicity assume that $u_0 = 0$ and that the firm employs IPE and sets a profit share α_I.

Consider a full sharing equilibrium, in which all future knowledge sharing with a deviating unit ceases following a deviation. Under the above assumptions

$$\begin{aligned} u^C &= \alpha_I (p - c_n) - \Psi (n-1), \\ u^P &= \alpha_I (1-\theta)(p - c_1), \\ u^D &= \alpha_I [(p - c_n) + (n-1) \theta (c_{n-1} - c_n)]. \end{aligned}$$

Note that the cooperative profits u^C are unaffected by the share of competitive consumers. Since in the full sharing equilibrium all units have the same costs, they do not find it profitable to compete with their peers. Since $\partial u^P / \partial \theta = -\alpha_I (p - c_1)$, an increase in the share of competitive clients disadvantages a potential deviator in all periods after the deviation. Because a deviator is excluded from all future knowledge sharing, it has higher cost than its peers and would lose his competitive clients to them. However, since $\partial u^D / \partial \theta = \alpha_I (n-1)(c_{n-1} - c_n)$, the deviator wins a larger share of competitive clients in the deviation period itself.

As in Section 3.1, to incentivize knowledge sharing, the firm chooses the minimal profit share that satisfies the relational contract incentive constraint, i.e.,

$$\underline{\alpha}_I^{ic} \equiv \frac{\Psi (n-1)}{p - c_n - (1-\delta)[(p - c_n) + (n-1)\theta (c_{n-1} - c_n)] - \delta (p - c_1)(1-\theta)}.$$

Since $\underline{\alpha}_I^{ic} (p - c_n) - \Psi (n-1) > 0 = u_0$, the manager's participation constraint is satisfied. Taking a derivative of the denominator with respect to θ, it can be seen that $\underline{\alpha}_I^{ic}$ decreases in θ if and only if

$$-(1-\delta)(n-1)(c_{n-1} - c_n) + \delta (p - c_1) > 0. \tag{4}$$

The condition holds for $\delta \geq \widehat{\delta}$ for some $\widehat{\delta} \in [0, 1)$,where $\widehat{\delta}$ decreases in the clients' willingness to pay p and increases in n. In this range, increasing the share of competitive clients θ lowers the cost of incentives to the units' managers under the full sharing equilibrium (the equilibrium is supported with similar strategies to those described in

Section 3.3).[8] If increasing the share of competitive clients also entails a cost, the firm's choice of θ would balance this cost against the lower costs of incentives. For example, the firm may be required to publish competitive internal bids for the "competitive" clients. Alternatively, internal competition may erode the profits that are extracted from the competitive clients also on the equilibrium path. This would be the case for example, if in contrast to the maintained assumption, no fixed cost is required to compete over a peers' competitive clients.

4.2 Decentralization of resources

In this section, I explore how the control of the firm's resources affects knowledge sharing. Under "centralization" the firm's center controls all resources and allocates them to the units based on their productivity at each period. Under "decentralization" the firm allocates the resources equally to the units at the beginning of the game.

To simplify we consider the case of $n = 2$. We modify the setup of Section 2 and assume that in each period $t = 1, 2, ..., \infty$ exactly one of the two units innovates with equal probability. Denote by ω_{it} the productivity of unit i in period t. Prior to any knowledge sharing, $\omega_{it} = \underline{a}$ for the noninnovating unit and $\omega_{it} = \overline{a}$ for the innovating unit, where $\underline{a} < \overline{a}$. If the innovator shares its knowledge with it peer then $\omega_{1t} = \omega_{2t} = \overline{a}$.

The firm has a single unit of resources which can be allocated between the units. Denote by x_{it} the amount of resources allocated to i, and by $\pi_{it} = \omega_{it} f(x_{it})$ the profit, where f is increasing and strictly concave. Under decentralization, each unit is assigned exactly half of the resources at $t = 0$, whereas under centralization the firm allocates its resources optimally in every period based on the units' productivities, i.e., the firm solves

$$\max_{0 \leq x_{1t} \leq 1} \omega_{1t} f(x_{1t}) + \omega_{2t} f(1 - x_{1t}). \tag{5}$$

Denote the optimal resource allocation by $x(\omega_1, \omega_2)$, where $x(\omega_1, \omega_2) = 1 - x(\omega_2, \omega_1)$. Since f is strictly concave it immediately follows that $x(\omega, \omega) = 0.5$ and $x(\omega_1, \omega_2) > 0.5$ if and only if $\omega_1 > \omega_2$.

Suppose the firm employs IPE and sets a profit share α_I to the units' managers. Consider a full sharing equilibrium, in which the along the equilibrium path, the innovating

[8] Specifically, in the punishment equilibrium σ^k, a deviating unit k is required to share its innovation with its peers for T periods, after which cooperation resumes. Units $j \in \mathcal{N} \setminus \{k\}$, take turn in competing with k over its competitive clients and thus in each period only one firm pays the fixed cost required to compete. If unit j deviates from σ^k,all units switch to play σ^j in the next period.

unit shares its innovation with its peer at a cost Ψ. If a unit fails to share its knowledge, all future knowledge sharing between the units ceases. On the equilibrium path, $\omega_{1t} = \omega_{2t} = \overline{a}$ and therefore $x_{1t} = x_{2t} = 0.5$ under both centralization and decentralization. In contrast, upon deviation, under decentralization both units continue to have a half unit of resources, whereas under centralization the resources are allocated optimally, i.e., $\overline{x} \equiv x\left(\overline{a}, \underline{a}\right) > 0.5$ for the innovator and $\underline{x} \equiv x\left(\underline{a}, \overline{a}\right)$ for the non-innovator and where $\overline{x} = 1 - \underline{x}$.

To gain intuition why decentralization can help, note that whereas the full sharing equilibrium are the same identical under both centralization and decentralization, the deviation payoff is higher under centralization. Because the resource allocation under centralization is optimal, upon deviation an innovating unit is "rewarded" with additional $\overline{x}-$ 0.5 resources in the deviation period. In addition, the optimal allocation of resources leads to higher profits under centralization than under decentralization in the punishment equilibrium. Thus, sustaining the full sharing equilibrium is made easier with the decentralization of resources, which "commits" the firm to an inefficient allocation of resources in case knowledge sharing breaks down. The next lemma summarizes the implications, and shows that decentralization lowers the cost of incentives to the firm.

Lemma 4 *Denote by $\underline{\delta}^C$ and $\underline{\delta}^D$ the minimal discount rates for which the firm opts to incentivizes the managers to share knowledge under centralization and decentralization, respectively. Then $\underline{\delta}^D < \underline{\delta}^C$ and for $\delta \in \left[\underline{\delta}^D, \underline{\delta}^C\right]$, the firm incentivizes knowledge sharing only under decentralization. For $\delta > \underline{\delta}^C$, the firm incentivizes knowledge sharing under both decentralization and centralization, but the managers' profit share is weakly lower with decentralization.*

5 Conclusions

Recent research has demonstrated that many significant management practices employed by successful firms are governed by relational contracts (for an extensive survey, see Gibbons and Henderson, 2013). This paper contributes to this literature, describing how relational contracts can help to sustain a consistent exchange of know-how between units in a multiunit firm. The analysis offers two main insights: First, powered incentives can help, rather than impede, knowledge sharing in firms, as they increase the monetary value of future cooperation with peer units. Second, while group incentives can help the units to internalize the value of their know-how to their peers, the firm can often do better by employing relative performance based incentives and committing to other competitive elements in its organizational design, such as allowing its units to compete over clients. The competitive environment helps the units to discipline non-cooperative peers, since when a unit is sanctioned and excluded from the cooperation between the units in response for non-cooperative behavior, it suffers a cost disadvantage which leads to a substantial drop in its future compensation. Thus, contrary to a common intuition, interunit competition can contribute to a strong cohesion between the units and lead to a higher productivity for the firm.

References

Dilip Abreu. On the theory of infinitely repeated games with discounting. *Econometrica: Journal of the Econometric Society*, 56(6):383–396, 1988.

Linda Argote. *Organizational learning: Creating, retaining and transferring knowledge.* Springer Science & Business Media, 2013.

Linda Argote, Sara L Beckman, and Dennis Epple. The persistence and transfer of learning in industrial settings. *Management science*, 36(2):140–154, 1990.

Linda Argote and Paul Ingram. Knowledge transfer: A basis for competitive advantage in firms. *Organizational behavior and human decision processes*, 82(1):150–169, 2000.

George Baker, Robert Gibbons, and Kevin J Murphy. Relational contracts and the theory of the firm. *The Quarterly Journal of Economics*, 117(1):39–84, 2002.

Yeon-Koo Che and Seung-Weon Yoo. Optimal incentives for teams. *American Economic Review*, 91(3):525–541, 2001.

Christopher J Collins and Ken G Smith. Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of management journal*, 49(3):544–560, 2006.

Karel O Cool, Ingemar Dierickx, and Gabriel Szulanski. Diffusion of innovations within organizations: Electronic switching in the bell system, 1971–1982. *Organization science*, 8(5):543–559, 1997.

Robert Gibbons and Rebecca Henderson. What do managers do? exploring persistent differences among seemingly similar enterprises. In Robert Gibbons and John Roberts, editors, *The Handbook of Organizational Economics*, chapter 17, pages 680–731. 2013.

Bengt Holmstrom. Moral hazard in teams. *The Bell Journal of Economics*, pages 324–340, 1982.

Bengt Holmstrom and Paul Milgrom. Regulating trade among agents. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 85–105, 1990.

Bengt Holmstrom and Paul Milgrom. Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics, and Organization*, 7:24–52, 1991.

Hideshi Itoh. Incentives to help in multi-agent situations. *Econometrica: Journal of the Econometric Society*, pages 611–636, 1991.

Edward P Lazear. Pay equality and industrial politics. *Journal of political economy*, 97 (3):561–580, 1989.

Edward P Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5):841–864, 1981.

Rafael Rob and Peter Zemsky. Social capital, corporate culture, and incentive intensity. *RAND Journal of Economics*, pages 243–257, 2002.

J. Roberts. *The Modern Firm.* Oxford University Press, 2004.

Jason J Sandvik, Richard E Saouma, Nathan T Seegert, and Christopher T Stanton. Workplace knowledge flows. *The Quarterly Journal of Economics*, 135(3):1635–1680, 2020.

Giancarlo Spagnolo. Social relations and cooperation in organizations. *Journal of Economic Behavior & Organization*, 38(1):1–25, 1999.

Wenpin Tsai. Social structure of coopetition within a multiunit organization: Coordination, competition, and intraorganizational knowledge sharing. *Organization science*, 13(2):179–190, 2002.

A Appendix

A.1 Proofs

Proof of Lemma 4. Under decentralization, the full sharing equilibrium incentive constraint for the unit that innovated is

$$\alpha_I \, \overline{a} f\left(0.5\right) - \Psi + \frac{\delta}{1-\delta}\left[\alpha_I \, \overline{a} f\left(0.5\right) - 0.5\Psi\right] \geq \alpha_I \, \overline{a} f\left(0.5\right) + \frac{\delta}{1-\delta}\alpha_I \, \frac{\left(\overline{a} + \underline{a}\right) f\left(0.5\right)}{2}. \quad (6)$$

Under centralization, the incentive constraint for the innovator is

$$\alpha_I \, \overline{a} f\left(0.5\right) - \Psi + \frac{\delta}{1-\delta}\left[\alpha_I \, \overline{a} f\left(0.5\right) - 0.5\Psi\right] \geq \alpha_I \, \overline{a} f\left(\overline{x}\right) + \frac{\delta}{1-\delta}\alpha_I \, \left[\frac{\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right)}{2}\right] \quad (7)$$

By definition, $\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right) \geq \overline{a} f\left(x\right) + \underline{a} f\left(1-x\right)$ for all $x \in [0.5, 1]$. Moreover, since $\left(\overline{a} - \underline{a}\right) f'\left(0.5\right) > 0$, the objective of the resource allocation problem is strictly increasing at $x_{1t} = 0.5$ and thus $\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right) > \left(\overline{a} + \underline{a}\right) f\left(0.5\right)$.

It follows from (6) that the full sharing equilibrium can be sustained under decentralization if and only if $\frac{\delta}{1-\delta} \geq d\left(\alpha_I\right)$, where

$$d\left(\alpha_I\right) = \frac{\Psi}{\alpha_I \, \frac{\left(\overline{a} - \underline{a}\right) f\left(0.5\right)}{2} - 0.5\Psi},$$

and is decreasing in α_I. Similarly, the full sharing equilibrium can be sustained under centralization if and only if $\frac{\delta}{1-\delta} \geq c\left(\alpha_I\right)$, where

$$c\left(\alpha_I\right) = \frac{\Psi + \alpha_I \, \overline{a}\left[f\left(\overline{x}\right) - f\left(0.5\right)\right]}{\alpha_I \, \left[\overline{a} f\left(0.5\right) - \frac{\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right)}{2}\right] - 0.5\Psi}$$

where $c\left(\alpha_I\right)$ is decreasing in α_I as well. To see this last point, denote $A = \overline{a}\left[f\left(\overline{x}\right) - f\left(0.5\right)\right] > 0$ and $B = \overline{a} f\left(0.5\right) - \frac{\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right)}{2}$, where $B > 0$ by concavity. Then, $c'\left(\alpha_I\right) \propto A\left[\alpha_I B - 0.5\Psi\right] - B\left[\Psi + \alpha_I A\right] = \left[-0.5A - B\right]\Psi < 0$.

Next we show that $c\left(\alpha_I\right) > d\left(\alpha_I\right)$ for all α_I. As the numerator of $c\left(\alpha_I\right)$ is higher than of $d\left(\alpha_I\right)$, a sufficient condition is that the denominator is smaller, or

$$\overline{a} f\left(0.5\right) - \frac{\overline{a} f\left(\overline{x}\right) + \underline{a} f\left(\underline{x}\right)}{2} < \frac{\left(\overline{a} - \underline{a}\right) f\left(0.5\right)}{2}.$$

Straightforward derivations show that this condition is equivalent to

$$\frac{\overline{a}}{\underline{a}} > \frac{\frac{f(0.5)-f(\underline{x})}{0.5-\underline{x}}}{\frac{f(\overline{x})-f(0.5)}{\overline{x}-0.5}}.$$

To show that the sufficient condition holds, write the first-order condition of the resource allocation problem, (5), as follows:

$$\frac{\overline{a}}{\underline{a}} = \frac{f'(\underline{x})}{f'(\overline{x})}.$$

As f is concave,

$$\begin{aligned} \frac{f(0.5) - f(\underline{x})}{0.5 - \underline{x}} &< f'(\underline{x}) \\ \frac{f(\overline{x}) - f(0.5)}{\overline{x} - 0.5} &> f'(\overline{x}) \end{aligned}$$

and thus

$$\frac{\overline{a}}{\underline{a}} = \frac{f'(\underline{x})}{f'(\overline{x})} > \frac{\frac{f(0.5)-f(\underline{x})}{0.5-\underline{x}}}{\frac{f(\overline{x})-f(0.5)}{\overline{x}-0.5}}.$$

Next, define the minimal profit shares that satisfy the incentive constraint, under centralization and decentralization, by $\underline{\alpha}_I^C(\delta)$ and $\underline{\alpha}_I^D(\delta)$. These minimal profit shares are implicitly defined by $c\left(\underline{\alpha}_I^C\right) = \frac{\delta}{1-\delta} = d\left(\underline{\alpha}_I^D\right)$. As $c(\alpha_I) > d(\alpha_I)$ and as both functions are strictly decreasing, $\underline{\alpha}_I^C(\delta) > \underline{\alpha}_I^D(\delta)$ for all δ.

Last, the firm incentivizes knowledge sharing under decentralization if

$$\underline{\alpha}_I^D(\delta) \leq 1 - \frac{(\underline{a} + \overline{a}) f(0.5) - 2u_0}{2\overline{a} f(0.5)}$$

and thus as $\underline{\alpha}_I^D(\delta)$ is decreasing in δ, there exists $\underline{\delta}^D$ such that the condition holds for $\delta \geq \underline{\delta}^D$. Similarly, under centralization the firm incentivizes knowledge sharing if

$$\underline{\alpha}_I^C(\delta) \leq 1 - \frac{\overline{a} f(\overline{x}) + \underline{a} f(\underline{x}) - 2u_0}{2\overline{a} f(0.5)},$$

which is satisfied for $\delta \geq \underline{\delta}^C$. Note that the expression in the right-hand side is larger under decentralization, as $(\underline{a} + \overline{a}) f(0.5) < \overline{a} f(\overline{x}) + \underline{a} f(\underline{x})$. Thus, as $\underline{\alpha}_I^D(\delta) \leq \underline{\alpha}_I^C(\delta)$ for all δ, it follows that $\underline{\delta}^D < \underline{\delta}^C$. ■

A.2 Optimal punishment strategy under RPE

Consider the following strategy profile, $\sigma^k = \left\{\sigma_i^k\right\}_{i=1}^n$ designed to punish unit $k \in \mathcal{N}$. σ^k has two substages: (i) In the first $T \geq 1$ periods, each unit $j \in \mathcal{N} \backslash \{k\}$ shares its innovation with all of its peers, except for unit k, whereas unit k shares its innovation with all of its peers. If unit k fails to share its innovation in any of the first T periods, the play of σ^k is restarted from period 1. (ii) After T punishment periods, full sharing between all units is resumed. At any stage, if a unit $j \in \mathcal{N} \backslash \{k\}$ fails to share its innovation, all units switch to play the strategy profile σ^j which is structured similarly.

In each of the first T periods of σ^k, unit $j \in \mathcal{N} \backslash \{k\}$'s payoff is

$$\overline{u} = \alpha_R \left[\pi(n) + z \cdot \frac{\pi(n) - \pi(1)}{n-1}\right] - \Psi(n-2),$$

while unit k's payoff is

$$\underline{\underline{u}} = \alpha_R \left(\pi(1) - z\left[\pi(n) - \pi(1)\right]\right) - \Psi(n-1).$$

Since $\underline{\underline{u}} < \underline{u}$ (defined in the text), unit k is expected to cooperate with its own punishment in the first T periods. Its cooperation is achieved using a "carrot and stick" strategy. The carrot offered to k is that after T punishment periods it would be forgiven and brought back into cooperation with its peers. The stick is that the punishment would be restarted from period 1 if k fails to act its part in any of the first T periods.

The discounted payoff of unit k under σ^k, as a function of $T \geq 0$, is

$$v_k(T) \equiv \sum_{t=0}^{T-1} \delta^t \cdot \underline{\underline{u}} + \sum_{t=T}^{\infty} \delta^t u^C.$$

Since $\underline{\underline{u}} < \underline{u} < u^C$, there exists $T_0 \geq 0$ such that $v_k(T_0+1) \leq \sum_{t=0}^{\infty} \delta^t \underline{u} = \frac{\underline{u}}{1-\delta} \leq v_k(T_0)$ and thus $A \times v_k(T_0+1) + (1-A) \times v_k(T_0) = \frac{\underline{u}}{1-\delta}$ for some $A \in (0,1)$. Upon deviation, the length of the punishment period is set to $T = T_0 + 1$ in probability A and to $T = T_0$ in probability $1 - A$.[9] If k is punished for T periods, its average payoff per period under σ^k is thus $\underline{u}$ and equals the lower bound.

While a ("trigger") strategy by which following a deviation by k all peers would cease cooperation with k forever would also attain the lower bound $\underline{u}$ for the punished unit,

[9] To implement, we may assume the existence of a public coordinating device, a random variable θ that is uniformly distributed over $[0,1]$. Then $T = T_0 + 1$ if $\theta \leq A$ and $T = T_0$ otherwise.

the payoff to all other units in $\mathcal{N}\setminus\{k\}$ ("punishers") under σ^k is higher. To see why this is the case observe that the sum of units payoffs under the "carrot-and-stick" strategy is higher than the trigger strategy. To see why note that under RPE

$$\sum_{i=1}^{n} u_i = \alpha_R \sum_{i=1}^{n} \pi_i - \alpha_R z \sum_{i=1}^{n} (\pi_i - \overline{\pi}_{-i}) - \sum_{i=1}^{n} \Psi_i = \alpha_R \sum_{i=1}^{n} \pi_i - \sum_{i=1}^{n} \Psi_i, \tag{8}$$

where the second equality follows as $\sum_{i=1}^{n} (\pi_i - \overline{\pi}_{-i}) = 0$. Since under the "carrot-and-stick" strategy, units have access to more innovations in all periods than under the trigger strategy the sum of profits net of communication costs must be higher. Finally, since k's payoff is the same under both strategies, the profits for all other units unit $j \in \mathcal{N}\setminus\{k\}$ must be higher. This makes it easier to incentivize them to adhere to the punishment strategies.

For σ^k to be a subgame perfect equilibrium, the punished unit k and all other units $j \in \mathcal{N}\setminus\{k\}$ must adhere to it. First, if k would choose not to cooperate with its punishment in any period in the first stage of σ^k and not share its innovations in one of the first T periods, it would get an immediate payoff $\underline{u}$, following which σ^k would be restarted from the first period. The payoff to such deviation is thus

$$\underline{u} + \delta v_k(T) = \underline{u} + \delta \frac{u}{1-\delta} = \frac{u}{1-\delta} = v_k(T),$$

i.e., the same as under σ^k. Next, if a unit $j \in \mathcal{N}\setminus\{k\}$ were to deviate during the first T periods of σ^k and withhold its innovation from any of its peers in $\mathcal{N}\setminus\{k\}$, all units would switch to play σ^j, lowering j's payoff to $\underline{u}$ per period in all future periods. Deviating from σ^k is not profitable for $j \in \mathcal{N}\setminus\{k\}$ if

$$\overline{u} \geq (1-\delta)\,\alpha_R \left[\pi(n) + z[\pi(n) - \frac{(n-2)\,\pi(n-1) + \pi(1)}{n-1}]\right] + \delta \underline{u}. \tag{9}$$

Last we show that whenever the equilibrium incentive constraint (3) is satisfied, then the incentive constraint (9) for a unit $j \in \mathcal{N}\setminus\{k\}$ in the punishment equilibrium σ^k are satisfied as well, and thus we only need to verify that the equilibrium incentive constraint holds. Since, by construction, a deviation by unit j leads in both cases to a switch to σ^j in all future periods, then if (3) is satisfied (9) is satisfied as well provided that the difference between the the discounted collusive payoffs in both cases , is larger than the difference in the profits in the period of deviation. That is, if

$$\overline{u}-u^{C}\geq\left(1-\delta\right)\alpha_{R}\left[\pi\left(n\right)+z[\pi\left(n\right)-\frac{\left(n-2\right)\pi\left(n-1\right)+\pi\left(1\right)}{n-1}]-\pi\left(n\right)-z\Delta\left(n\right)\right].$$

Rearranging and omitting identical terms from both sides, we get

$$\overline{u}-u^{C}\geq\left(1-\delta\right)\alpha_{R}z\frac{\pi\left(n-1\right)-\pi\left(1\right)}{n-1}.$$

This condition is satisfied since,

$$\overline{u}-u^{C}=\alpha_{R}z\cdot\frac{\pi\left(n\right)-\pi\left(1\right)}{n-1}+\Psi(n-1)-\Psi\left(n-2\right)>\alpha_{R}z\cdot\frac{\pi\left(n\right)-\pi\left(1\right)}{n-1}>\alpha_{R}z\cdot\frac{\pi\left(n-1\right)-\pi\left(1\right)}{n-1}.$$