A Theory of Firm Boundaries and Knowledge Sharing*

Evgenii Fadeev
Duke University

Michael Powell
Northwestern University

Preliminary and Incomplete

June, 2025

Abstract

We develop a theory of how integration shapes knowledge sharing. Parties hold private information essential for coordination, but they can also expropriate each other's knowledge. Asset ownership grants control over expropriation, creating a trade-off: the owner freely shares knowledge, while the non-owner partially withholds information to avoid expropriation. Thus, the asset is optimally allocated to the party with the lowest temptation to expropriate. Extending the model to multiple downstream and upstream parties, we show that the optimal industry configuration may involve a large firm acting as a knowledge intermediary, aggregating information across competing firms that are reluctant to share it directly. Finally, we identify conditions on knowledge characteristics and contracting environments under which integration becomes necessary for effective knowledge transfer, as optimal outcomes cannot be achieved through contracts alone.

*Fadeev: Duke University, Fuqua School of Business, email: evgenii.fadeev@duke.edu. Powell: Kellogg School of Management, mike-powell@kellogg.northwestern.edu.

1 Introduction

Modern empirical research on firm boundaries has identified a consistent pattern: integration facilitates the transfer of productivity-enhancing knowledge from acquiring to acquired firms (Guadalupe, Kuzmina, and Thomas (2012); Atalay, Hortaçsu, and Syverson (2014); Braguinsky, Ohyama, and Syverson (2015); Eliason, Heebsh, McDevitt, and Roberts (2020); La Forgia and Bodner (2024)). This rationale for integration is frequently invoked in merger announcements and antitrust defenses.[1] However, this observation raises a fundamental question. Why must a firm acquire another firm's assets to effectively transfer knowledge? Why can the same productivity gains not be realized through contractual arrangements? In this paper, we address these questions by analyzing (*i*) the benefits and costs of integration for knowledge sharing, (*ii*) how incentives for knowledge sharing shape the configuration of industries, and (*iii*) the types of knowledge transfer that cannot be effectively achieved through contracts and thus require integration.

As an example, consider a firm that has developed proprietary knowledge for an innovative product design. This design firm considers acquiring manufacturing assets and hiring a production partner capable of manufacturing the product. Alternatively, the firm could outsource production through a licensing agreement. However, outsourcing creates a risk: the independent manufacturer might exploit the firm's valuable design knowledge to produce competing products or supply rival firms. In other words, knowledge may be expropriated: while the initial knowledge transfer may be contractible, its subsequent use is typically not. Vertical integration can mitigate this risk, allowing the design firm to safely share its proprietary knowledge with the production partner, as it now controls the assets required for expropriation. Yet integration comes at a cost: the production partner may possess valuable manufacturing know-how, which it might be reluctant to fully disclose, fearing the design firm could expropriate this knowledge—for instance, by replacing the production partner.

This example illustrates the broader mechanism developed in our model. Two parties, whether vertically or horizontally related, must coordinate over the use of a productive asset, and each holds private information that, if shared, would enhance joint performance. However, disclosure carries a risk: the asset owner may expropriate the other party's information for private gain at the sharer's expense. We adopt a mechanism-design approach to analyze how asset ownership shapes the outcomes parties can achieve. We show that ownership determines both the extent and

[1] For instance, see the Amazon-Whole Foods merger (2017), the SEC filing "Amazon and Whole Foods Market Announce Acquisition to Close This Monday, Will Work Together to Make High-Quality, Natural and Organic Food Affordable for Everyone."

direction of knowledge sharing. Specifically, the asset owner's knowledge is freely communicated since it cannot be expropriated. In contrast, the non-owner's information remains vulnerable to expropriation and is, therefore, partially withheld. The optimal allocation of ownership balances this trade-off: the asset should be owned by the party with the lower temptation to expropriate.

We extend the model to multiple parties to analyze how incentives for knowledge sharing shape industry configuration. Consider a setting with several design and manufacturing parties, where all parties can potentially share knowledge with one another. A key insight is that optimal asset ownership may involve allocating manufacturing assets to a single production party who acts as a knowledge intermediary. This intermediary aggregates information across multiple design parties, facilitating knowledge sharing among close competitors without requiring direct disclosure between them.

In general, the optimal industry configuration depends on two factors: the gains from knowledge combination and parties' proximity in the knowledge space. Proximity is defined as the amount of information parties need to engage in expropriation. We assume design parties are closer to each other than to manufacturing parties, meaning that design parties require less information to expropriate one another's knowledge. The configuration involving a knowledge intermediary is optimal when the gains from knowledge combination among design parties are sufficiently high, and the knowledge distance between design and manufacturing parties is relatively large. In contrast, if the gains from knowledge combination are low, and vertical parties are relatively close in the knowledge space, the optimal industry configuration involves multiple vertically integrated firms. Thus, allocating assets to a knowledge intermediary enables knowledge combination across horizontally related parties at the expense of partial information withholding with vertically related parties.

The model explains not only when manufacturing should be consolidated within a single party but also why this manufacturing party, acting as a knowledge intermediary, should be independently owned. If this intermediary were a subsidiary of one of the design firms, it would have to share information with its parent firm to coordinate activities. Such an arrangement, however, would discourage other design firms from sharing their knowledge with the intermediary due to the risk of expropriation by its parent firm. For instance, General Motors stated that a rationale for spinning off its parts division, Delphi Automotive Systems, was to ensure other automakers would be willing to share knowledge with Delphi.[2] Similarly, in the computer industry's transition from vertical integration toward contract manufacturing, one of the benefits of contract manufacturers was their ability to accumulate knowledge by serving multiple clients (McGee, 2025).

[2] "G.M. Plans to Spin off Parts Division," New York Times, 1998.

We conclude by discussing the conditions on knowledge and the contracting environment under which contracts become ineffective for knowledge transfer and thus necessitate integration. Specifically, we establish a Coase-theorem-like result: if expropriation actions are contractible and knowledge can be concealed during the contracting stage, then asset ownership is irrelevant, since the surplus-maximizing outcome can be achieved through contractual arrangements alone. Therefore, at least one of these conditions must fail for asset ownership to affect knowledge sharing.

If expropriation actions are noncontractible and can be controlled only through asset ownership, parties cannot credibly commit to refrain from knowledge expropriation, making integration necessary. However, even if expropriation actions are contractible, this alone does not guarantee cooperative outcomes through contracts. Since parties negotiate after observing their private information, the contract offer itself may reveal information about the proposing party's knowledge before an agreement is reached.

Whether parties can conceal their knowledge prior to signing a contract depends on the nature of the information. In economics, private information is commonly modeled as a realization of a random variable observed by some parties but not others, with its distribution being common knowledge. Under this statistical representation, information can always be concealed by appropriately specifying a contract—an insight known as the inscrutability principle in mechanism design (Myerson, 1983). In such scenarios, the contract specifies enforceable decisions and message exchanges contingent on each possible state. However, knowledge may also possess an "awareness-inducing" or "eye-opening" property (Arrow, 1962; Gabaix and Laibson, 2006; Tirole, 2009), meaning that the mere description of a state within a contract reveals it to the other party.

We construct a simple example where one party possesses "eye-opening" knowledge and all actions are contractible. Knowledge is disclosed directly through the contract offer itself. Once revealed, parties can achieve first-best outcomes, since all actions remain contractible. However, this initial disclosure changes the parties' outside options for subsequent negotiations. As a result, the informed party may prefer to refrain from making an offer, keeping its knowledge concealed, even though disclosure would result in higher joint surplus. This result contrasts with the property-rights literature initiated by Grossman and Hart (1986) and Hart and Moore (1990). In that literature, initial asset ownership is irrelevant because parties always reallocate assets to maximize joint surplus. In contrast, under "eye-opening" knowledge, parties may avoid contracting over asset ownership altogether, as contracts themselves risk knowledge revelation, placing the privately informed party in a weaker negotiating position.

Literature Review

A large empirical literature explores how firm boundaries shape the flow of knowledge among individuals and organizations. For example, Atalay, Hortaçsu, and Syverson (2014) provide evidence that integration primarily facilitates knowledge transfer rather than the transfer of physical goods. There are also several recent papers that study the effects of acquisitions and reach similar conclusions. For example, Eliason, Heebsh, McDevitt, and Roberts (2020) find that acquired firms adopt practices from their parent companies.

Much of the theoretical literature studying knowledge in organizations relies on the assumption that knowledge flows more easily within firms than between. These studies often model intra-firm knowledge transfer as frictionless, while inter-firm transfers involve communication costs. For instance, Arrow (1975) and Riordan (1991) assume that parties can observe each other's private information within a firm, but not across firm boundaries. Similarly, research applying the Garicano (2000) model of knowledge hierarchies typically think of the knowledge available to a firm being a characteristic internal to the firm and do not allow for across-firm knowledge transfer.

One critique of this assumption is that it remains silent about the precise nature of how communication costs are shaped by asset ownership. As Grossman and Hart (1986) emphasize: "Arrow (1975) has analyzed the benefits of vertical integration based on the assumption that without integration, it is more costly for one firm to communicate information to another than with integration. The incentives of people to lie may change if their incentive structure changes, but Arrow does not explain how integration changes the set of feasible incentive structures." (p. 695)

We build off Grossman and Hart's insight and model how firm boundaries influence incentives for knowledge sharing through their impact on the potential expropriation of knowledge. Practitioners and academic economists commonly argue that in-house operations offer superior knowledge protection. For instance, in Waymo v. Uber, Waymo argued that internal assembly of LiDAR sensors protected its trade secrets.[3] Similarly, Antràs and Rossi-Hansberg (2009) conjecture that Intel's ownership of operations in Costa Rica reflected concerns about technological expropriation, while Busse (2022) notes that Apple designs its own chips partly to protect proprietary knowledge. Much of the strategy and multinational-firm literature likewise assumes that knowledge is better protected within firm boundaries (e.g., Teece, 1980; Grant, 1996; Liebeskind, 1996; Ethier and Markusen, 1996; Conconi, Leone, Magerman, and Thomas, 2025).

To our knowledge, ours is the first organizational economics model of why firm boundaries offer superior knowledge protection. We formalize the idea that knowledge protection stems from asset

[3] Waymo LLC v. Uber Technologies, Inc., No. 17-2235 (Fed. Cir. 2017)

ownership granting control over expropriation actions. Using techniques from the literature on mechanism design with an informed principal (Myerson, 1983), we characterize the best achievable outcomes for the parties. This framework also allows us to identify not only the benefits of integration for knowledge transfer, but also its costs, the implications for industry configuration, and the conditions under which integration becomes necessary, as contractual arrangements alone cannot achieve optimal knowledge sharing.

Our work also speaks to the organizational economics literature on communication and delegation in organizations, particularly Dessein (2002), Dessein and Santos (2004), Alonso, Dessein, and Matouschek (2008), and Rantakari (2008), which analyze how organizational structure and communication interact when parties have private information. Galeotti, Dessein, and Santos (2015) and Matouschek, Powell, and Reich (2025) examine how communication structures affect organizational performance. All these papers ask how internal structure shapes communication. We ask how firm boundaries themselves shape parties' incentives to share information.

2 The Model

We present our baseline model of a single dyad to illustrate the key trade-offs governing how firm boundaries shape knowledge flows. A principal is privately informed about which task is valuable but has neither the knowledge about how best to approach that task nor the human capital necessary to do the task. An agent is privately informed about the best way to approach the task and has the human capital to do the task. While the parties could benefit jointly by sharing knowledge with one another, they also have the ability to expropriate each other's knowledge.

Information, Actions, and Preferences. The state is a vector $\theta = (\theta_P, \theta_A) \in \Theta$, where $\theta_P \in \Theta_P = \{-1, 1\}$ indicates which of two tasks is productive, and $\theta_A \in \Theta_A = \{-a, a\}$ indicates which of two approaches is productive, where $|a| > 1$. The variables θ_P and θ_A are independent, and all outcomes are equally likely. The principal observes θ_P, and the agent observes θ_A.

There are three actions to be taken. Action $x \in \{-1, 0, 1\}$ is the task to be undertaken. Control over this action is possessed by the agent and is inalienable. The task $x = 0$ is a safe status quo task. Action $y \in \{-a, 0, a\}$ is the approach to pursue, where $y = 0$ is a safe status quo approach, and action z is an expropriation action that we will describe in more detail below. Control over y and z is inextricably tied to a productive asset. These control rights are therefore alienable but inseparable from one another.

The preferences of party $i \in \{P, A\}$ are

$$u_i(x, y, z, t; \theta) = \underbrace{\pi_i(x, y; \theta)}_{\text{cooperation}} + \underbrace{v_i(z; \theta)}_{\text{expropriation}} + t_i,$$

where $t_i \in \mathbb{R}$ is a transfer received by i. The cooperation payoff, $u_i(\cdot)$, depends on the agent's choice of task x and the asset owner's choice of approach y. Specifically, if $x = \theta_P$ and $y = \theta_A$, cooperation succeeds, and $\pi_P = \pi_A = \pi > 0$. If either $x = -\theta_P$ or $y = -\theta_A$, cooperation fails, and $\pi_P = \pi_A = -c < 0$. Otherwise, the status quo is implemented, and both parties receive a cooperation payoff of zero, or $\pi_P = \pi_A = 0$.

The expropriation payoff, $v_i(\cdot)$, depends on the expropriation action $z \in \{-a, -1, 0, 1, a\}$. The action $z = 0$ represents no expropriation and yields an expropriation payoff of zero for both parties. All other choices represent an expropriation attempt. A choice of $z = \theta_A$ represents successful expropriation of the agent's knowledge, yielding a private benefit of b_P for the principal and a cost k for the agent. A choice of $z = \theta_P$ represents successful expropriation of the principal's knowledge, yielding a private benefit of b_A for the agent and a cost of k to the principal. Any other choice of z is an unsuccessful expropriation attempt: $z = -\theta_A$ yields a cost of c_P to the principal, and $z = -\theta_P$ yields a cost of c_A to the agent. We make the following assumptions on the payoff parameters:

Assumption. *For each $i \in \{P, A\}$, the following inequalities are satisfied:*

1. $k > 2\pi + b_i$ *and* $k > \pi + c$
2. $b_i/c_i < \pi/c < 1$.

The first assumption ensures that expropriation is sufficiently costly that it results in negative total surplus, and it rules out coercive screening contracts (an issue we discuss below). The second assumption ensures that both cooperation and expropriation are risky but that expropriation is riskier than cooperation is. That is, parties prefer to choose safe actions if they are uninformed about the other's state, and more information is needed to expropriate than to cooperate.

Contracts. We assume actions x, y, and z are noncontractible. The parties can, however, contract over monetary transfers, asset ownership (i.e., control over the decisions y and z), and the behavior of a neutral mediator with whom they can exchange messages. The contract directly determines some contractible actions and provides parties with incentives to take noncontractible actions.

It will be useful to introduce some notation to allow us to formally define a contract. Let $\mathcal{G} = \{P, A\}$ denote the set of possible control rights, and $\mathcal{T} = \mathbb{R}$ the set of transfers. Let $a_A = (x, y_A, z_A) \in \mathcal{A}_A$ be the agent's choice of noncontractible actions, where y_A and z_A are

his proposals for the actions tied to asset ownership. Similarly, let $a_P = (y_P, z_P) \in \mathcal{A}_P$ denote the principal's choice of noncontractible actions. Given control right g and proposed actions (y_A, z_A) and (y_P, z_P), the implemented actions are $(y, z) = (y_g, z_g)$. Let $\mathcal{O} = \mathcal{G} \times \mathcal{T} \times \mathcal{A}_P \times \mathcal{A}_A$, and call $o = (g, t, a_P, a_A)$ an **outcome.** A **contract** μ is a state-contingent distribution over outcomes, that is, $\mu \in \Delta(\mathcal{O} \times \Theta)$.

We can think of the contract as setting up a neutral mediator. Each party simultaneously and privately sends a message to this mediator, and then the mediator chooses an outcome $o = (g, t, a_P, a_A)$. The control rights g and transfer t are enforced, and each individual i is given a private recommendation that he choose action a_i (which he is not compelled to do). Following Myerson (1983), it is without loss of generality to focus on incentive-compatible direct mechanisms in which parties truthfully announce their private information to the mediator, and they follow the mediator's recommended action.

To define incentive compatibility, define party i's expected payoffs under contract μ:

$$U_i(\mu|\,\theta_i) = \frac{1}{2} \sum_{\theta_j \in \Theta_j} \sum_{o \in \mathcal{O}} \mu(o|\,\theta)\, u_i(o, \theta),$$

where, abusing notation, $u_i(o, \theta) = u_i(x, y, z, \theta)$, $y = y_g$, and $z = z_g$. For any possible types θ_i and θ_i' of individual i and any function $\delta_i : \mathcal{A}_i \to \mathcal{A}_i$, define the deviation payoff

$$U_i^*\left(\mu, \delta_i, \theta_i' \middle|\, \theta_i\right) = \frac{1}{2} \sum_{\theta_j \in \Theta_j} \sum_{o \in \mathcal{O}} \mu\left(o \middle|\, \theta_j, \theta_i'\right) u_i\left(\left(g, t, a_j, \delta_i(a_i)\right), \theta\right).$$

Assuming party j announces his type truthfully, this is the expected payoff party i would receive if he announced type θ_i' when his true type was θ_i, and if he chose action $\delta_i(a_i)$ following a recommendation of a_i by the mediator. The contract μ is **incentive-compatible** if for both parties i,

$$U_i(\mu|\,\theta_i) \geq U_i^*\left(\mu, \delta_i, \theta_i' \middle|\, \theta_i\right) \text{ for all } \theta_i, \theta_i', \text{ and } \delta_i.$$

Timing. Denote by g_0 the initial control right. The timing of the game is:

1. The principal and agent observe θ_P and θ_A, respectively.
2. The principal makes a take-it-or-leave-it offer of contract μ to the agent.
3. The agent accepts the offer or rejects it. If the agent rejects the offer, then he chooses x, and party g_0 chooses y and z.
4. If the agent accepts the offer, control rights are transferred to party g, and transfer t is made from the principal to the agent. Both parties then send messages $m_i \in \Theta_i$ to the mediator

who then sends private recommendations $\hat{a}_P$ and $\hat{a}_A$ to the parties. After observing these recommendations, both parties choose their actions a_P and a_A.

Solution Concept. Since the principal is privately informed about which task is productive, this is an informed-principal problem (Myerson, 1983; Maskin and Tirole, 1990, 1992). It is natural, then, to ask whether we need to consider contracts that vary with the principal's type. Following Myerson (1983), we can restrict attention without loss to **inscrutable contracts**: contracts that are the same for both principal types and do not reveal her type to the agent. If different principal types offered different contracts, the agent would infer the principal's type from the contract he received, and the contract would need to be incentive-compatible given that the agent knows the principal's type. But now suppose we construct a single, type-independent contract under which the principal truthfully reveals her type to the mediator only after the agent accepts. The mediator then implements the same contract that would have been implemented if different principal types offered different contracts. This direct mechanism is incentive-compatible and delivers the same payoffs to all parties. Hence it is without loss of generality to focus on inscrutable contracts.

Our solution concept will be **optimal contracts**, which we define to be inscrutable incentive-compatible contracts μ that maximize $\frac{1}{2}U_P(\mu|1) + \frac{1}{2}U_P(\mu|-1)$ over all incentive-compatible contracts. For informed-principal problems in general, there is not an unambiguous notion of optimal contracts since different principal types might value different contracts differently. We show in Section 6 that our notion of optimal contracts selects contracts that are, in an important sense, the only undominated contracts. In our setting, therefore, optimal contracts are selected by all of the refinements proposed by Myerson (1983).

3 Technical Preliminaries

In this section, we establish several intermediate results that enable us to describe optimal contracts as the solution to a straightforward optimization problem. In particular, it shows that we are able to restrict attention to contracts that have an invariance property that reflects the symmetry of the model. And it shows that an optimal contract is one that maximizes parties' joint surplus. This section is largely technical in its current form and the three lemmas and their proofs can be skipped by most readers.

Our setting is one in which the environment is symmetric in the sense that for any two states θ and θ', for any realized outcome in state θ, there is a payoff-equivalent outcome in state θ'. For example, in the state $(\theta_P, \theta_A) = (1, a)$, the outcome $(x, y, z) = (1, a, 0)$ yields payoffs (π, π) for the principal and agent respectively. In state $(\theta_P, \theta_A) = (-1, a)$, the outcome $(x, y, z) = (-1, a, 0)$ also

yields (π, π) for the two parties. We make use of this observation to define a class of contracts that "translate" a given contract μ in state (θ_P, θ_A) to state $\left(\theta'_P, \theta'_A\right)$ by providing parties with the same distribution of payoffs given state $\left(\theta'_P, \theta'_A\right)$ as the original contract does in state (θ_P, θ_A). To do so, note that we can write the expropriation action $z \in \{-1, -a, 0, a, 1\}$ as the sum of two components: $z = z_P + z_A$, where $z_P \in \{-1, 0, 1\}$ and $z_P \in \{-a, 0, a\}$.

Formally, given a contract μ, we say that μ' is the θ_P**-translation** of μ if for every pair of states $\theta = (\theta_P, \theta_A)$ and $\theta' = (-\theta_P, \theta_A)$ and every outcome $o = (x, y, z_P + z_A)$ and translated outcome $o' = (-x, y, -z_P + z_A)$, we have $\mu(o|\,\theta) = \mu'\left(o'|\,\theta'\right)$. Similarly, we say that μ' is the θ_A**-translation** of μ if for every pair of states $\theta = (\theta_P, \theta_A)$ and $\theta' = (\theta_P, -\theta_A)$ and every outcome $o = (x, y, z_P + z_A)$ and translated outcome $o' = (x, -y, z_P - z_A)$, we have $\mu(o|\,\theta) = \mu'\left(o'|\,\theta'\right)$. If a contract μ is a θ_P-translation of itself, we say that μ is θ_P**-invariant**. If a contract μ is a θ_A-translation of itself, we say that μ is θ_A**-invariant**. If a contract μ is both θ_P-invariant and θ_A-invariant, we say that μ is **invariant**.

We will show that we can focus, without loss of generality, on invariant contracts. To do so, let IC be the set of incentive-compatible and feasible contracts. The first relevant observation is that the set IC is a convex set and is closed under translations. The proofs are in the appendix.

Lemma 1. *If $\mu, \mu' \in IC$, then $(1-\lambda)\mu + \lambda\mu' \in IC$. And if $\mu \in IC$ and μ' is either a θ_P- or θ_A-translation of μ, then $\mu' \in IC$.*

Recall that an optimal contract maximizes $\frac{1}{2}U_P(\mu|\,1) + \frac{1}{2}U_P(\mu|\,-1)$ over all contracts that are incentive-compatible and feasible. In general, such contracts could give different principal types different payoffs and expose them to different distributions over outcomes. The next lemma shows that we can ignore those kinds of asymmetric contracts and focus on invariant contracts.

Lemma 2. *There is an optimal contract μ^* that is invariant.*

One useful property of invariant contracts is that they deliver both agent types the same utility. Since they have the same outside option payoff, this implies that in any optimal contract, the principal fully extracts the agent's surplus. The next lemma establishes this result.

Lemma 3. *If contract μ^* is an invariant optimal contract, then $U_A(\mu^*|\,\theta_A) = 0$ for all θ_A and therefore $U_P(\mu^*|\,\theta_P) = U_P(\mu^*|\,\theta_P) + U_A(\mu^*|\,\theta_A)$.*

An important implication of Lemma 3 is that we can focus on joint-surplus-maximizing contracts. This observation leads to the main result in this section. Before stating the proposition, we introduce a piece of notation. For an invariant contract μ, let $U_P(\mu) = U_P(\mu|\,\theta_P)$ for all θ_P and $U_A(\mu) = U_A(\mu|\,\theta_A)$ for all θ_A.

Proposition 1. *Suppose* μ^* *solves*

$$\max_{\mu \in IC} U_P(\mu) + U_A(\mu)$$

subject to μ *being an invariant contract. Then* μ^* *is optimal.*

Proposition 1 establishes that for the purposes of characterizing optimal contracts, it is without loss to focus on invariant incentive-compatible contracts that maximize joint surplus. In the next section, we will further simplify the problem and derive its implications for optimal contracts.

4 Control Shapes Knowledge Transmission

In this section, we characterize optimal contracts. We take as our starting point Proposition 1, which shows that we can focus on invariant incentive-compatible contracts that maximize joint surplus. We further simplify the problem by studying a relaxed problem in which we replace the incentive-compatibility condition with an obedience constraint. That is, instead of requiring that the contract provides parties incentives to truthfully reveal their private information to the mediator, we only require that the mediator's recommendations are followed.

Consider the following relaxed problem. Let μ_g be an invariant contract with asset allocation g. We can think of $\mu_g(a|\theta)$, then, as being a distribution over action recommendations to the two parties for each state θ. For each $g \in \{P, A\}$, we will characterize contracts μ_g that solve

$$\max_{\mu_g} U_P(\mu_g) + U_A(\mu_g)$$

subject to the **obedience constraint**: for all $i \in \{P, A\}$, for all action profiles a in which party i is privately recommended a_i,

$$\sum_{\theta_j} \sum_{a_j} \mu_g(a|\theta)\, u_i(a, \theta) \geq \sum_{\theta_j} \sum_{a_j} \mu_g(a|\theta)\, u_i(a_j, a_i', \theta)$$

for all alternative actions a_i'. As a reminder, under $g = A$, the agent chooses all three actions, so $a_A = (x, y, z)$. Under $g = P$, the agent chooses action x, so $a_A = x$, and the principal chooses the remaining two actions, so $a_P = (y, z)$.

To characterize the set of contracts that satisfy the obedience constraint, it suffices to consider a signal $\sigma = (\sigma_P, \sigma_A)$ that provides a distribution over action recommendations and induces a distribution over parties' beliefs in each state θ.[4] Specifically, suppose the agent privately receives

[4] This problem is an information-design-in-games problem. By formulating it the way we have, we are following the Bergemann and Morris (2016) and Taneva (2019) obedience approach to analyzing it.

signal σ_A (informative about θ_P) and the principal receives the signal σ_P (informative about θ_A) satisfying $\sigma_P \in \{-1, 1\}$ and $\sigma_A \in \{-a, a\}$ with

$$\Pr\left[\sigma_A = \theta_P | \, \theta_P\right] = \tau_A \text{ and } \Pr\left[\sigma_P = \theta_P | \, \theta_A\right] = \tau_P.$$

For simplicity, assume these signals are independent conditional on the state.[5] This signal structure is chosen so that if the principal sees the signal realization $\hat{\sigma}_P = a$, her posterior belief that $\theta_A = a$ is τ_P. And if the agent sees the signal realization $\hat{\sigma}_A = 1$, his posterior belief that $\theta_P = 1$ is τ_A.

We will argue that optimal contracts choose the accuracy of these two signals, τ_P and τ_A, to maximize the probability the parties successfully cooperate subject to the constraint that neither party chooses to expropriate the other. To this end, define the thresholds $\overline{\tau} = c/(\pi + c)$ and $\overline{\tau}^i = c_i/(b_i + c_i)$. For the party i with the control right to be willing to follow a recommendation of cooperating but not expropriating the other, the signal must have an accuracy of $\overline{\tau} \leq \tau_i \leq \overline{\tau}^i$. If it is below $\overline{\tau}$, then party i would prefer to choose the safe action. And if it is above $\overline{\tau}^i$, party i would prefer to cooperate and also expropriate.

Principal Ownership. Under $g = P$, consider the following recommendations and signals. The agent receives signal realization $\hat{\sigma}_A$ and is recommended action

$$\hat{x} = \begin{cases} 0 & \text{if} \quad 0 \leq \tau_A < \overline{\tau} \\ \hat{\sigma}_A & \text{if} \quad \overline{\tau} \leq \tau_A \leq 1. \end{cases}$$

And the principal receives signal realization $\hat{\sigma}_P$ and is recommended action

$$(\hat{y}, \hat{z}) = \begin{cases} (0,0) & \text{if} \quad 0 \leq \tau_P < \overline{\tau} \\ (\hat{\sigma}_P, 0) & \text{if} \quad \overline{\tau} \leq \tau_P \leq \overline{\tau}^P \\ (\hat{\sigma}_P, \hat{\sigma}_P) & \text{if} \quad \overline{\tau}^P < \tau_P \leq 1. \end{cases}$$

For the agent to be willing to follow the recommended action, it must be the case that

$$E_{\theta_P, \hat{y}, \hat{z}}\left[u_A\left(\hat{x}, \hat{y}, \hat{z}\right) | \, \theta_A, \hat{\sigma}_A, \hat{x}\right] \geq \max_x E_{\theta_P, \hat{y}, \hat{z}}\left[u_A\left(x, \hat{y}, \hat{z}\right) | \, \theta_A, \hat{\sigma}_A, \hat{x}\right],$$

and for the principal to be willing to follow the recommended actions, it must be the case that

$$E_{\theta_A, \hat{x}}\left[u_P\left(\hat{x}, \hat{y}, \hat{z}\right) | \, \theta_P, \hat{\sigma}_P, (\hat{y}, \hat{z})\right] \geq \max_{y,z} E_{\theta_A, \hat{x}}\left[u_P\left(\hat{x}, y, z\right) | \, \theta_P, \hat{\sigma}_P, (\hat{y}, \hat{z})\right].$$

It is straightforward to verify these constraints are satisfied given the definitions of $\overline{\tau}$, $\overline{\tau}^A$, and $\overline{\tau}^P$.

Provided $\tau_A \geq \overline{\tau}$ and $\overline{\tau} \leq \tau_P \leq \overline{\tau}^P$, cooperation is successful with probability $\tau_P \tau_A$ (i.e., when each party receives the signal realization equal to the other's state), and so total expected surplus

[5] In general, parties can benefit from making use of correlated signals, but not in this model.

is $\tau_P\tau_A 2\pi - (1 - \tau_P\tau_A)\, 2c$. The left panel of Figure 1 depicts total expected surplus as a function of the precision of the principal's signal τ_P, and choosing τ_A optimally given τ_P (i.e., for $\tau_P < \overline{\tau}$, $\tau_A = 0$, and for $\tau_P \geq \overline{\tau}$, $\tau_A = 1$). For $\tau_P < \overline{\tau}$, parties choose the safe actions, and so expected surplus is zero. When τ_P exceeds the cooperation threshold $\overline{\tau}$, the parties attempt to cooperate, and they are able to cooperate with a higher probability the higher is τ_P. When τ_P exceeds the expropriation threshold $\overline{\tau}^P$, however, the parties attempt to cooperate, but the principal also attempts to expropriate the agent, reducing total surplus.

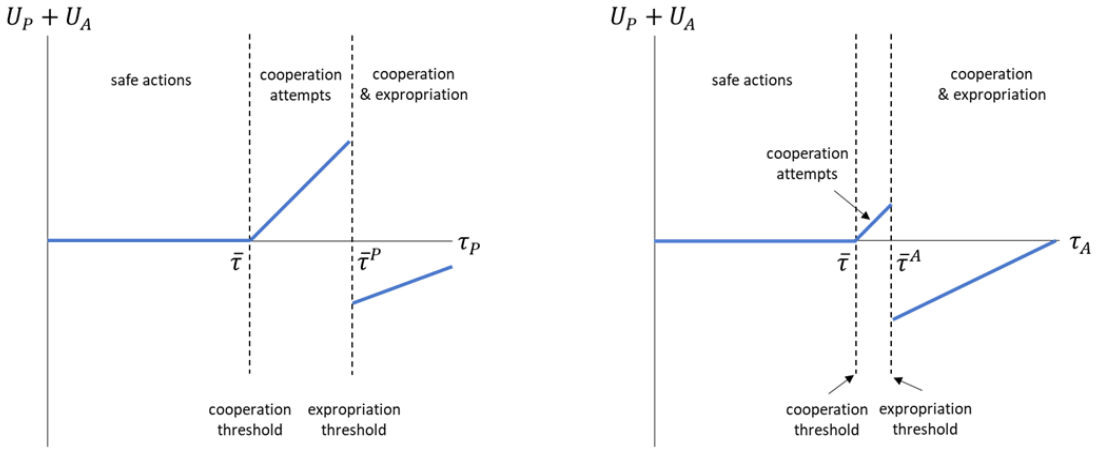


Figure 1: The left panel depicts maximized total expected surplus as a function of τ_P under principal ownership. The right panel depicts total expected surplus as a function of τ_A under agent ownership. The parameters used for both figures are $\pi = c = 2$, $b_P = 1, b_A = 2, c_P = c_A = 3$, and $k = 6$.

It is clear from this figure that total surplus is maximized in the middle region. The relaxed problem then becomes

$$\max_{\tau_A, \tau_P} \tau_P\tau_A 2\pi - (1 - \tau_P\tau_A)\, 2c$$

subject to $\tau_A \geq \overline{\tau}$ and $\overline{\tau} \leq \tau_P \leq \overline{\tau}^P$, so that under principal ownership, $\tau_A^* = 1$ and $\tau_P^* = \overline{\tau}^P$.

$$\frac{c_P}{b_P + c_P} 2\pi - \left(1 - \frac{c_P}{b_P + c_P}\right) 2c \geq 2\pi + (b_P - k)$$

Agent Ownership. Under $g = A$, consider the following recommendations and signals. The agent receives signal realization $\hat{\sigma}_A$ and is recommended actions

$$(\hat{x}, \hat{y}, \hat{z}) = \begin{cases} (0,0,0) & \text{if} \quad 0 \leq \tau_A < \overline{\tau} \\ (\hat{\sigma}_A, \theta_A, 0) & \text{if} \quad \overline{\tau} \leq \tau_A \leq \overline{\tau}^A \\ (\hat{\sigma}_A, \theta_A, \hat{\sigma}_A) & \text{if} \quad \overline{\tau}^A < \tau_A \leq 1. \end{cases}$$

And the principal receives signal realization $\hat{\sigma}_P$ but takes no actions.

For the agent to be willing to follow the recommended actions, it must be the case that

$$E_{\theta_P}\left[u_A\left(\hat{x},\hat{y},\hat{z}\right)|\,\theta_A,\hat{\sigma}_A,\left(\hat{x},\hat{y},\hat{z}\right)\right] \geq \max_{x,y,z} E_{\theta_P}\left[u_A\left(x,y,z\right)|\,\theta_A,\hat{\sigma}_A,\left(\hat{x},\hat{y},\hat{z}\right)\right],$$

which again is straightforward to verify given the definitions of $\overline{\tau}$ and $\overline{\tau}^A$.

Provided $\overline{\tau} \leq \tau_A \leq \overline{\tau}^A$, cooperation is successful with probability τ_A (i.e., when the agent receives the signal realization equal to the principal's state), and so total expected surplus is $\tau_A 2\pi - (1-\tau_A)\,2c$. The right panel of Figure 1 depicts total expected surplus as a function of the precision of the agent's signal τ_A. For $\tau_A < \overline{\tau}$, the agent chooses safe actions, and so expected surplus is zero. When τ_A exceeds the cooperation threshold $\overline{\tau}$, the agent attempts to cooperate, and they are able to cooperate with a higher probability the higher is τ_A. When τ_A exceeds the expropriation threshold $\overline{\tau}^A$, however, the parties attempt to cooperate, but the principal also attempts to expropriate the agent, reducing total surplus.

The relaxed problem then becomes

$$\max_{\tau_A} \tau_A 2\pi - (1-\tau_A)\,2c$$

subject to $\overline{\tau} \leq \tau_A \leq \overline{\tau}^A$, so that under agent ownership, $\tau_A^* = \overline{\tau}^A$, and τ_P^* can take any value.

Optimal Contracts. Under principal ownership, the principal's information is fully shared with the agent, while the agent partially withholds his information from the principal to prevent the principal from expropriating that information. The maximized total surplus is

$$\overline{\tau}^P 2\pi - \left(1-\overline{\tau}^P\right)2c.$$

Similarly, under agent ownership, the agent's information can be fully shared with the principal, but the principal partially withholds his information from the agent to prevent expropriation. The maximized total surplus is

$$\overline{\tau}^A 2\pi - \left(1-\overline{\tau}^A\right)2c.$$

Principal ownership is therefore optimal if $\overline{\tau}^P > \overline{\tau}^A$, or $b_P/c_P < b_A/c_A$. These results are summarized in the following proposition.

Proposition 2. *Under either principal or agent ownership, optimal contracts involve no expropriation. Optimal contracts under principal ownership satisfy* $\tau_A^* = 1$ *and* $\tau_P^* = \overline{\tau}^P$. *Optimal contracts under agent ownership satisfy* $\tau_A^* = \overline{\tau}^P$ *and* $\tau_P^* = 1$. *Principal ownership is optimal if* $b_P/c_P < b_A/c_A$.

Proposition 2 shows that if expropriation by the principal is riskier than is expropriation by the agent, more information will be shared under principal ownership than under agent ownership, and so principal ownership will lead to more effective cooperation. The proof of the proposition also shows how to use the solution to the relaxed problem to construct a contract that is in fact incentive-compatible and therefore optimal.

Asset ownership determines whose knowledge is fully shared and whose is partially withheld. The asset owner's knowledge is shared fully, while the non-owner's knowledge is shared only up to the expropriation threshold. Shifting ownership reallocates, but does not eliminate, the expropriation risk and therefore shapes knowledge flows. The next section shows how this logic can help us understand the role of knowledge intermediaries.

5 Knowledge Intermediaries

This section uses insights from our main model to understand the role and organization of suppliers that act as knowledge intermediaries. These are firms like contract manufacturers that work on behalf of several downstream design firms. By learning from all the downstream firms, knowledge intermediaries can combine the knowledge they gain, making them more productive in each of their relationships. Working with a knowledge intermediary opens one up to expropriation risk, however, and downstream firms will withhold some of their knowledge. We show that knowledge intermediaries arise as an optimal configuration when the knowledge combination benefit is high.

To this end, we will consider a stylized industry model that builds off a simplified version of the main model. There are two downstream parties D_1 and D_2 and two upstream parties U_1 and U_2. The state is a vector $\theta = (\theta_1, \theta_2) \in \Theta$, where $\theta_i \in \{-1, 1\}$ indicates which of two tasks is productive for D_i's production. The variables θ_1 and θ_2 are independent, and all outcomes are equally likely. D_i privately observes θ_i.

There are four types of actions. Action $x_i^j \in \{-1, 0, 1\}$ is the task that U_j chooses in their interaction with D_i. Control over actions $\left(x_1^j, x_2^j\right)$ is possessed by U_j and is inalienable. Task $x_i^j = 0$ is a safe status quo task. Action $y_i \in \{-1, 0, 1\}$ is the approach pursued on behalf of D_i, where $y_i = 0$ is a safe status quo approach. Action $z_i \in \{-1, 0, 1\}$ is an expropriation action. And action $w_i \in \{-1, 0, 1\}$ is the knowledge-combination action. Control over y_i, z_i, and w_i are inextricably linked to productive asset i. These control rights are therefore alienable but inseparable from one other. Denote $x = \left(x_1^1, x_2^1, x_1^2, x_2^2\right)$, $y = (y_1, y_2)$, $w = (w_1, w_2)$, and $z = (z_1, z_2)$.

The preferences of party $i \in \{D_1, D_2, U_1, U_2\}$ are

$$u_i(x, y, z) = \pi_i(x, y, w; \theta) + v_i(z; \theta) + t_i.$$

The cooperation payoff $u_{D_i}(\cdot)$ depends on the task chosen by the upstream party who D_i chooses to work with and the choices of y by the asset owner(s) and on whether knowledge combination was successful. Specifically, if D_i works with U_j, then if $x_i^j = y_i = 0$, the status quo is implemented, and both parties receive a cooperation payoff of zero from this interaction. If $x_i^j = y_i = \theta_i$ and $x_k^j = 0$, we say that cooperation between U_j and D_i succeeds, and both parties receive a cooperation payoff of $\pi > 0$ from this interaction. If in addition, $w_j = \theta_i$ and $w_i = \theta_j$, then we say that knowledge-combination succeeds, and the parties receive $\Pi > \pi$ rather than π. Any other configuration of actions yields a payoff of $-c < 0$ from this interaction for both parties. An upstream party working with multiple downstream parties receives the sum of the payoffs from both their interactions.

The expropriation payoffs in the interaction between D_i and U_j depend on the choice of z_i and the state θ_i. If party D_i works with U_j, and expropriation action $z_i = \theta_i$ is taken, then D_i incurs a cost of k, while U_j receives a benefit of b_U. If $z_i = 0$, then both receive an expropriation payoff of zero. And if $z_i = -\theta_i$, then U_j incurs a cost of c_U. We also allow for the possibility that one downstream party can expropriate the other. If, for example, D_1 chooses $z_2 = \theta_2$, then D_2 incurs a cost of k, while D_1 receives a benefit of b_D. If D_1 chooses $z_2 = -\theta_2$, then D_1 incurs a cost of c_D. We will assume that $b_U/c_U < \pi/c \leq b_D/c_D$, reflecting the idea that a party's downstream rival has a bigger temptation to expropriate their information than does an upstream party.

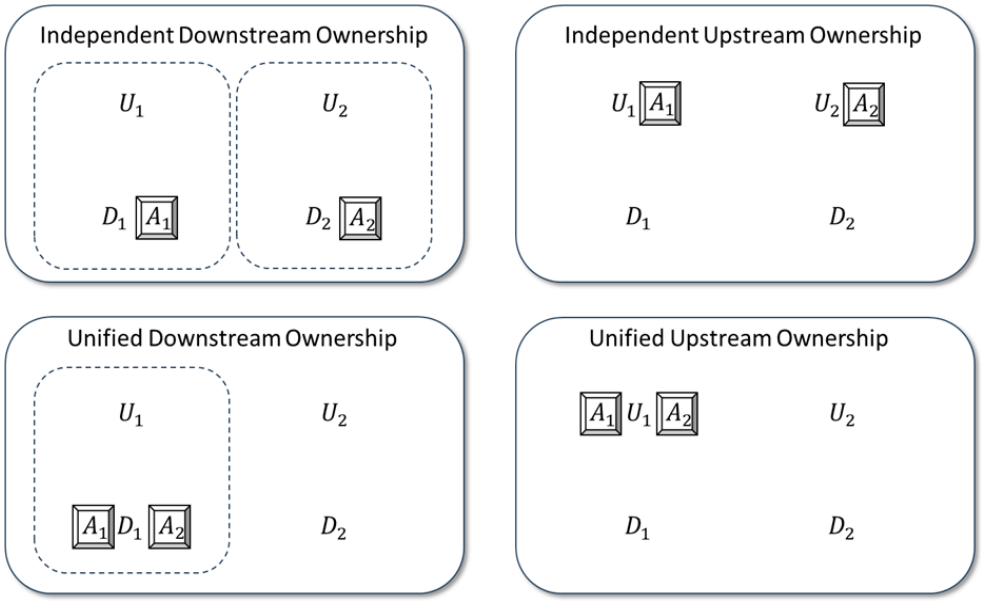


Figure 2: This figure describes the four industry configurations. Under independent downstream ownership, each downstream party owns its corresponding asset. Under independent upstream ownership, each upstream party owns its corresponding asset. Under unified downstream ownership, U_1 owns both assets, and under unified upstream ownership, U_1 owns both assets.

Denote by $g = (g_1, g_2)$ the industry configuration, where $g_\ell \in \{D_1, D_2, U_1, U_2\}$ denotes the owner of asset ℓ. We will consider four industry configurations. We refer to $g = (D_1, D_2)$ as downstream independent ownership. Each downstream party owns their own productive asset and chooses actions (y_i, w_i, z_i). We refer to $g = (U_1, U_2)$ as upstream independent ownership. Each upstream party owns the productive asset for their relationship with their corresponding downstream party, and upstream party U_j chooses actions $\left(x_1^j, x_2^j, y_j, w_j, z_j\right)$. We refer to $g = (U_1, U_2)$ as upstream unified ownership in which U_1 chooses actions $\left(x_1^1, x_2^1, y_1, w_1, z_1, w_1, y_2, w_2, z_2\right)$. Finally, we refer to $g = (D_1, D_2)$ as downstream unified ownership in which U_1 chooses actions $\left(x_1^1, x_2^1\right)$ and D_1 chooses actions $(y_1, w_1, z_1, y_2, w_2, z_2)$. These industry configurations are depicted in Figure 2.

We assume all actions are noncontractible. The parties can, however, contract over monetary transfers, the industry configuration, and the behavior of a neutral mediator with whom they can exchange messages. We assume that D_1 makes a take-it-or-leave-it offer of a contract to the remaining three parties. If all three parties accept the contract offer, play proceeds according to the offered contract. If any party rejects the offer, we assume all parties receive a payoff of zero. As in the main model, we will consider incentive-compatible contracts that are invariant and that maximize joint surplus over all four parties.

Equilibrium knowledge flows under each ownership structure. We now summarize optimal contracts under each of the four industry configurations. First, under downstream independent ownership, each downstream party fully transmits their knowledge to their corresponding upstream party, and actions $x_i^i = y_i = \theta_i$ and $z_i = 0$ are taken. The resulting total payoffs are $U_{D_1} + U_{U_1} = U_{D_2} + U_{U_2} = 2\pi$. Under upstream independent ownership, each downstream party partially withholds their knowledge to their corresponding upstream party, and actions $\left(x_i^i, y_i, w_i, z_i\right) = (\theta_i, \theta_i, 0, 0)$ are taken with probability $\overline{\tau}^U = c_U / (b_U + c_U)$ and actions $\left(x_i^i, y_i, w_i, z_i\right) = (-\theta_i, -\theta_i, 0, 0)$ are taken with probability $1 - \overline{\tau}^U$. The resulting total payoffs are $U_{D_1} + U_{U_1} = U_{D_2} + U_{U_2} = \overline{\tau}^U 2\pi - \left(1 - \overline{\tau}^U\right) 2c$.

Under downstream unified ownership, since $b_D/c_D \geq \pi/c$, if D_2 shares enough knowledge with D_1 to get D_1 to choose $y_2 = \theta_2$, then D_1 will also choose $z_2 = \theta_2$. There is therefore full knowledge sharing from D_1 to U_1 but no knowledge transmitted from D_2. The resulting total profits from the two interactions are $U_{D_1} + U_{U_1} = 2\pi$ and $U_{D_2} + U_{U_1} = 0$. Finally, under upstream unified ownership, both downstream parties partially withhold their knowledge from U_1. Actions $\left(x_1^1, x_2^1, y_1, w_1, z_1, y_2, w_2, z_2\right) = (\theta_1, \theta_2, \theta_1, \theta_2, 0, \theta_2, \theta_1, 0)$ are taken with probability $\overline{\tau}^U$ and actions $\left(x_1^1, x_2^1, y_1, w_1, z_1, y_2, w_2, z_2\right) = (-\theta_1, -\theta_2, -\theta_1, -\theta_2, 0, -\theta_2, -\theta_1, 0)$ are taken with probability $1 - \overline{\tau}^U$. Note that the optimal contract involves perfectly correlated decisions: whenever miscoordination

happens, it happens for both downstream parties. In this case, the total payoffs from the two interactions are $U_{D_1} + U_{U_1} = U_{D_2} + U_{U_1} = \overline{\tau}^U 2\Pi - \left(1 - \overline{\tau}^U\right) 2c$. These results are depicted in Figure 3.

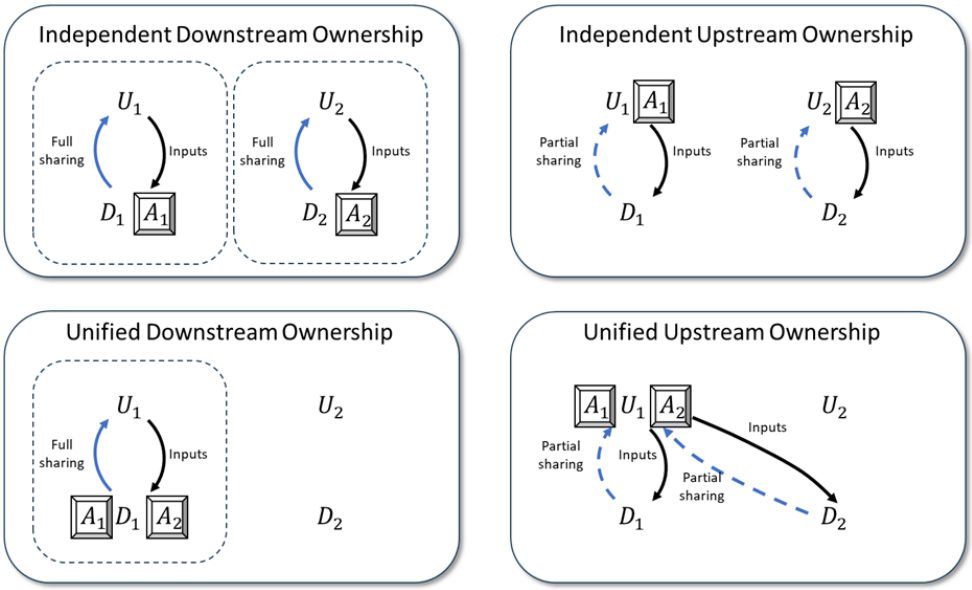


Figure 3: This figure describes knowledge flows under the four industry configurations. Under independent downstream ownership, both downstream parties fully share their information with their corresponding upstream parties. Under independent upstream ownership, both downstream parties partially share their knowledge with their corresponding upstream parties and choose $\tau = \overline{\tau}^U$. Under unified downstream ownership, D_1 fully shares their knowledge with U_1, but D_2 does not share any knowledge. Under unified upstream ownership, both downstream parties partially share their knowledge with U_1.

It is clear that upstream independent ownership performs less well than upstream unified ownership. Upstream ownership opens up parties to potential expropriation and therefore results in reduced information transmission, which in turn reduces the probability of successful cooperation. Doing so may be worthwhile, however, if the gains from knowledge combination are sufficiently high. Having an upstream knowledge intermediary (i.e., unified upstream ownership) is better than working only with one's in-house supplier (i.e., independent downstream ownership) if:

$$\Pi - \pi > \left(1 - \overline{\tau}^U\right)(\Pi + c),$$

where the left-hand-side of the inequality measures the gains from knowledge combination, and the right-hand-side measures the costs associated with miscoordination due to information withholding, which are decreasing in the riskiness of expropriation for U, which we think of as a measure of the distance between the upstream and downstream parties in the knowledge space. Knowledge intermediaries are therefore valuable when combining knowledge is essential. Figure 4 illustrates

the optimal industry configuration as a function of the gains from knowledge combination, $\Pi - \pi$, and the vertical knowledge-space distance, $c_U/(b_U + c_U)$.

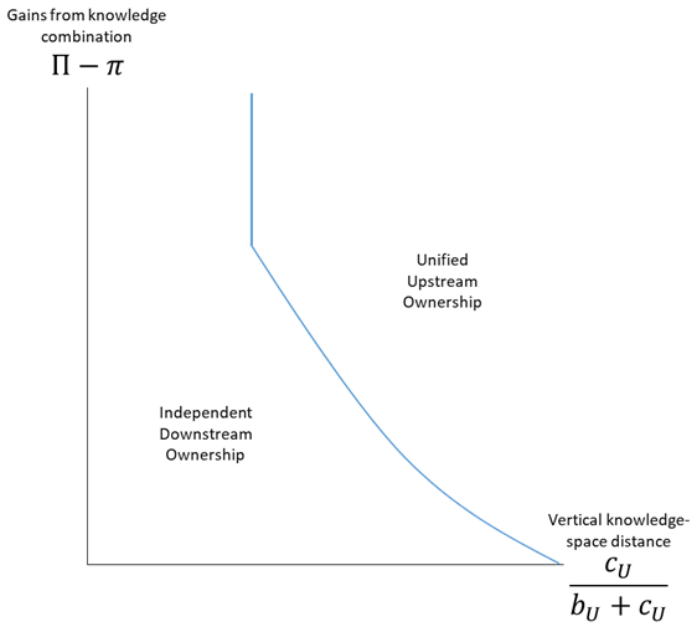


Figure 4: This figure illustrates the optimal industry configuration for different regions of the parameter space. Note that when $C_U/(C_U+B_U) \leq \bar{\tau}$, no knowledge sharing occurs under any form of upstream ownership.

In principle, one of the two downstream parties could serve as the knowledge intermediary, as under unified downstream ownership. But since it is easy for the downstream parties to expropriate one another (i.e., $b_D/c_D \geq \pi/c$), there will be no knowledge sharing between the other downstream party and the knowledge intermediary. As a result, whenever it is optimal for there to be a knowledge intermediary, that knowledge intermediary will not be one of the downstream rivals.

6 Contractibility and the Role of Property Rights

In our setting, the key assumption ensuring that the allocation of property rights matters is that the expropriation action is non-contractible. If there are no contractual safeguards preventing parties from expropriating each other's knowledge, then knowledge transmission cannot be achieved effectively through a contract alone, and other governance instruments become relevant. Assuming ex post noncontractibility also allows us to sidestep the powerful Maskin and Tirole (1999) critique of the property rights literature, namely that subgame-perfect-implementation mechanisms alone can solve hold-up problems in settings in which parties commonly observe payoff-relevant information and negotiate over decisions ex post.

In this section, we show that in settings in which expropriation actions are contractible (i.e., through the use of non-disclosure agreements and purpose-limitation covenants), whether the allocation of property rights matters for knowledge transmission depends on whether the knowledge can be concealed prior to signing a contract. Under the standard approach to modeling private information, the powerful inscrutability principle of mechanism design (Myerson, 1983) ensures that even an informed principal can offer a contract that implements efficient decisions without revealing her private information during the contract-offer stage.

For example, suppose the expropriation action is contractible, and consider an inscrutable contract that implements $z = 0$ and, after the contract has been accepted by the agent, asks the parties for their private information and implements the outcome $(x, y) = (\theta_P, \theta_A)$. Such a contract implements the efficient outcome state-by-state. At the time of the contract-acceptance decision, however, the agent learns nothing from the contract offer itself, and therefore the agent's outside option is independent of the principal's private information even if the agent initially owns the asset and can therefore choose the expropriation action following a contract rejection.

Now suppose the principal's private information is what Tirole (2009) refers to as "awareness-inducing" or, more evocatively, "eye-opening" information in which "a party's suggestion of contracting on [a decision that depends on the principal's information] gives the information away and prevents the knowledgeable party from fully benefiting from it." (Tirole, 2009, p. 266). Information of this kind cannot be withheld through the use of an inscrutable contract, and we argue, whether it is ultimately transmitted depends on the initial allocation of property rights.

There is no standard way to model eye-opening information (but see Akerlof, Holden, and Li (2025) for one approach), so we propose a particular and stark form of it to illustrate the role of property rights. Suppose the principal's state $\theta_P \in \{-1, 1\}$ tells us which of two agents, A_{-1} or A_1, has the capability to generate value for the principal, and suppose the principal only knows the identity of A_i if $\theta_P = i$, so she cannot approach both agents at the contract-offering stage. Then any offer the principal makes to agent A_i reveals that $\theta_P = i$.

Suppose productive assets are owned by the agents, so they control expropriation decisions following contract rejection. Then any contract offer to A_i reveals $\theta_P = i$, and therefore increases A_i's outside option to $\pi + b_A$. The principal therefore has to compensate A_i to avoid expropriation, and so the principal's payoff will be at most $\pi - b_A$ in any contract. If instead, the principal does not make an offer, no knowledge is transmitted, and the principal's payoff is 0. If $\pi < b_A$, then, no knowledge will be transmitted in equilibrium.

Suppose instead that the principal owns the asset. Then even though any contract offer to

A_i reveals $\theta_P = i$, A_i cannot expropriate that information, and so his outside option is π. The principal's payoff under an optimal contract will therefore be $\pi > 0$, and so she will necessarily offer a contract and transmit her information in equilibrium. Thus, if knowledge is eye-opening, then the allocation of property rights determines knowledge transmission even in environments in which the expropriation action is contractible. In contrast to the Property Rights Theory literature (Grossman and Hart, 1986; Hart and Moore, 1990), it is the initial rather than the ultimate allocation of property rights that matters. This insight is also robust to the Maskin and Tirole (1999) critique because any offer of a subgame-perfect-implementation mechanism would reveal the principal's private information to the agent.

7 Conclusion

This paper develops a mechanism-design framework for analyzing the incentives for knowledge disclosure between and within firms. When asset ownership provides control over the ability to expropriate knowledge, the allocation of that asset becomes a lever for mitigating expropriation. Assigning ownership to the party with the weaker expropriation temptation maximizes the scope of knowledge sharing while preventing expropriation.

Beyond shedding light on bilateral relationships, the analysis clarifies when industries gravitate toward an upstream "knowledge intermediary." Concentrating key assets in a single supplier can enable the benefits of combining the knowledge of many downstream clients, but it also raises a new concern: misappropriation by the intermediary itself. Our multi-buyer extension shows that integration is favored when the benefits of horizontal knowledge combination outweigh the additional withholding that clients must adopt to safeguard their knowledge. If the importance of horizontal knowledge combination has increased over time, the model can explain why the electronics industry has shifted from vertical integration of design and manufacturing toward specializing in design and working with large contract manufacturers. The model also clarifies why these contract manufacturers should be independent rather than the subsidiary of one of the design firms.

A natural next step is to study the internal organization of knowledge intermediaries and in particular how they control internal knowledge flows. Firms rely on "Chinese walls," compartmentalized IT systems, and legal firewalls to prevent leaks. These safeguards matter in settings in which leaks are possible: settings with multiple rounds of communication that lie beyond the reach of standard revelation-principal arguments (see Sugaya and Wolitzky (2021)). Extending the present framework to model endogenous information barriers promises to deepen our understanding of how organizations manage knowledge combination without sacrificing confidentiality.

Appendix

Lemma 1. *If $\mu, \mu' \in IC$, then $(1-\lambda)\mu + \lambda\mu' \in IC$. And if $\mu \in IC$ and μ' is either a θ_P- or θ_A-translation of μ, then $\mu' \in IC$.*

Proof of Lemma 1. Since μ and μ' are incentive-compatible contracts, for all $i \in \{P, A\}$, for all θ_i, θ_i' and $\delta_i : \mathcal{A}_i \to \mathcal{A}_i$,

$$\begin{aligned}
\sum_{\theta_j}\sum_{o} \mu\left(o|\,\theta\right) u_P\left(o,\theta\right) &\geq \sum_{\theta_j}\sum_{o} \mu\left(o|\,\left(\theta_i',\theta_j\right)\right) u_P\left(a_j, \delta_i\left(a_i\right), \left(\theta_i',\theta_j\right)\right) \\
\sum_{\theta_j}\sum_{o} \mu'\left(o|\,\theta\right) u_P\left(o,\theta\right) &\geq \sum_{\theta_j}\sum_{o} \mu'\left(o|\,\left(\theta_i',\theta_j\right)\right) u_P\left(a_j, \delta_i\left(a_i\right), \left(\theta_i',\theta_j\right)\right)
\end{aligned}$$

We therefore have

$$\begin{aligned}
&\sum_{\theta_j}\sum_{o} \left[(1-\lambda)\,\mu\left(o|\,\theta\right) + \lambda\mu'\left(o|\,\theta\right)\right] u_P\left(o,\theta\right) \\
\geq\ &\sum_{\theta_j}\sum_{o} \left[(1-\lambda)\,\mu\left(o|\,\theta\right) + \lambda\mu'\left(o|\,\theta\right)\right] u_P\left(a_j, \delta_i\left(a_i\right), \left(\theta_i',\theta_j\right)\right),
\end{aligned}$$

so $(1-\lambda)\mu + \lambda\mu'$ is incentive-compatible. $(1-\lambda)\mu + \lambda\mu'$ is also feasible because it is a mixture distribution for each θ and is therefore a distribution for each θ. This establishes the first part of the lemma.

For the second part of the lemma, suppose μ is incentive-compatible. Let μ' be the θ_P-translation of μ and μ'' be the θ_A-translation of μ. It is clear that μ' and μ'' are both probability distributions for each θ. It remains to show they are both incentive-compatible. Denote the collection of inequalities

$$U_i\left(\mu|\,\theta_i\right) \geq U_i\left(\mu, \theta_i', \delta_i\middle|\,\theta_i\right) \text{ for all } \delta_i : \mathcal{A}_i \to \mathcal{A}_i$$

by $IC_i\left(\theta_i, \theta_i'\middle|\,\mu\right)$. Since μ is incentive-compatible, $IC_i\left(\theta_i, \theta_i'\middle|\,\mu\right)$ holds for all i, θ_i, and θ_i'.

By construction, μ' satisfies, for all o and all θ_A,

$$\begin{aligned}
u_P\left(o, \left(\theta_P, \theta_A\right)\right) &= u_P\left(o', \left(-\theta_P, \theta_A\right)\right) \\
u_A\left(o, \left(\theta_P, \theta_A\right)\right) &= u_A\left(o', \left(-\theta_P, \theta_A\right)\right) \\
\mu\left(o|\,\left(\theta_P, \theta_A\right)\right) &= \mu'\left(o'\middle|\,\left(-\theta_P, \theta_A\right)\right).
\end{aligned}$$

We therefore have

$$\begin{aligned}
IC_P\left(\theta_P, \theta_P|\,\mu'\right) &\iff IC_P\left(-\theta_P, -\theta_P|\,\mu\right) \\
IC_P\left(\theta_P, -\theta_P|\,\mu'\right) &\iff IC_P\left(-\theta_P, \theta_P|\,\mu\right) \\
IC_A\left(\theta_A, \theta_A'|\,\mu'\right) &\iff IC_A\left(\theta_A, \theta_A'|\,\mu\right),
\end{aligned}$$

and so μ' is incentive-compatible.

Next, by construction, μ'' satisfies, for all o and all θ_P,

$$\begin{aligned} u_P(o,(\theta_P,\theta_A)) &= u_P\left(o'',(\theta_P,-\theta_A)\right) \\ u_A(o,(\theta_P,\theta_A)) &= u_A\left(o'',(\theta_P,-\theta_A)\right) \\ \mu(o|(\theta_P,\theta_A)) &= \mu''\left(o''\middle|(\theta_P,-\theta_A)\right). \end{aligned}$$

We therefore have

$$\begin{aligned} IC_P\left(\theta_P,\theta_P'|\mu''\right) &\iff IC_P\left(\theta_P,\theta_P'|\mu\right) \\ IC_A\left(\theta_A,\theta_A|\mu''\right) &\iff IC_A\left(-\theta_A,-\theta_A|\mu\right) \\ IC_A\left(\theta_A,-\theta_A|\mu''\right) &\iff IC_A\left(-\theta_A,\theta_A|\mu\right), \end{aligned}$$

and so μ'' is incentive-compatible.■

Lemma 2. *There is an optimal contract μ^* that is invariant.*

Proof of Lemma 2. To establish this lemma, we first show that there is an optimal contract μ^* for which $U_P(\mu^*|1) = U_P(\mu^*|-1)$. Then, we show that if μ satisfies $U_P(\mu|1) = U_P(\mu|-1)$ and $\mu \in IC$, then there is an invariant contract $\overline{\mu} \in IC$ that satisfies $U_P(\overline{\mu}|\theta_P) = U_P(\mu|\theta_P)$ for all θ_P. Putting these two results together, we can construct an invariant contract $\overline{\mu}^*$ that is optimal.
Step 1. There exists an optimal contract μ^* for which $U_P(\mu^*|1) = U_P(\mu^*|-1)$.

Proof of Step 1. Suppose μ_1 is optimal. That is, μ_1 solves

$$\max_{\mu} \frac{1}{2}U_P(\mu|1) + \frac{1}{2}U_P(\mu|-1)$$

subject to μ_1 feasible and μ_1 incentive compatible. Suppose without loss that $U_P(\mu_1|1) \geq U_P(\mu_1|-1)$. Let contract μ_{-1} be the θ_P-translation of μ_1. By symmetry, since μ_1 is incentive-compatible, μ_{-1} is also incentive compatible. Moreover,

$$U_P(\mu_1|1) = U_P\left(\mu_{-1}\middle|1\right) \text{ and } U_P(\mu_1|-1) = U_P\left(\mu_{-1}\middle|1\right).$$

Let $\mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_{-1}$. By Lemma 1, since μ_1 and μ_{-1} are incentive-compatible, μ is incentive-compatible. And

$$\begin{aligned} U_P(\mu|1) &= \frac{1}{2}U_P(\mu_1|1) + \frac{1}{2}U_P\left(\mu_{-1}\middle|1\right) \\ &= \frac{1}{2}U_P\left(\mu_{-1}\middle|-1\right) + \frac{1}{2}U_P(\mu_1|-1) = U_P(\mu|-1), \end{aligned}$$

which establishes the first part of the claim.

Step 2. Suppose μ satisfies $U_P(\mu|1) = U_P(\mu|-1)$ and is incentive-compatible. Then there is an invariant contract $\overline{\mu}$ that is incentive compatible and satisfies $U_P(\overline{\mu}|\theta_P) = U_P(\mu|\theta_P)$ for all θ_P.

Proof of Step 2. Suppose μ satisfies $U_P(\mu|1) = U_P(\mu|-1)$ and incentive compatibility. Let μ' be the θ_P-translation of μ, μ'' be the θ_A-translation of μ, and μ''' be the θ_P-translation of μ''. By

Lemma 1, μ', μ'', and μ''' are incentive compatible, and by construction,

$$\begin{aligned} U_P\left(\mu'\middle|\,\theta_P\right) &= U_P\left(\mu\middle|\,-\theta_P\right) = U_P\left(\mu\middle|\,\theta_P\right) \\ U_P\left(\mu''\middle|\,\theta_P\right) &= U_P\left(\mu\middle|\,\theta_P\right) \\ U_P\left(\mu'''\middle|\,\theta_P\right) &= U_P\left(\mu\middle|\,-\theta_P\right) = U_P\left(\mu\middle|\,\theta_P\right). \end{aligned}$$

Let $\overline{\mu} = \frac{1}{4}\left(\mu + \mu' + \mu'' + \mu'''\right)$. Then by Lemma 1, $\overline{\mu}$ is incentive compatible, and by construction,

$$U_P\left(\overline{\mu}\middle|\,\theta_P\right) = \frac{1}{4}\left(U_P\left(\mu\middle|\,\theta_P\right) + U_P\left(\mu'\middle|\,\theta_P\right) + U_P\left(\mu''\middle|\,\theta_P\right) + U_P\left(\mu'''\middle|\,\theta_P\right)\right) = U_P\left(\mu\middle|\,\theta_P\right),$$

and since

$$\begin{aligned} \overline{\mu} &= \frac{1}{4}\left(\mu + \mu' + \mu'' + \mu'''\right) = \frac{1}{4}\left(\mu' + \mu + \mu''' + \mu''\right) \\ &= \frac{1}{4}\left(\mu'' + \mu''' + \mu + \mu'\right), \end{aligned}$$

$\overline{\mu}$ is both θ_P-invariant and θ_A-invariant. It is therefore invariant.■

Lemma 3. *If contract μ^* is an invariant optimal contract, then $U_A\left(\mu^*\middle|\,\theta_A\right) = 0$ for all θ_A and therefore $U_P\left(\mu^*\middle|\,\theta_P\right) = U_P\left(\mu^*\middle|\,\theta_P\right) + U_A\left(\mu^*\middle|\,\theta_A\right)$.*

Proof of Lemma 3. By Lemma 2, we can focus without loss of generality on incentive-compatible invariant contracts μ. Any invariant contract μ has $U_A\left(\mu\middle|\,a\right) = U_A\left(\mu\middle|\,-a\right)$. Suppose μ is an invariant optimal contract and satisfies $U_A\left(\mu\middle|\,a\right) > 0$. Then consider an alternative invariant contract μ' that has the same distribution over outcomes as μ but reduces t_A and increases t_P. The contract μ' is incentive compatible, satisfies participation, and has $U_P\left(\mu'\middle|\,\theta_P\right) > U_P\left(\mu\middle|\,\theta_P\right)$, contradicting the claim that μ was optimal.■

Proposition 1. *Suppose μ^* solves*

$$\max_{\mu \in IC} U_P\left(\mu\right) + U_A\left(\mu\right)$$

subject to μ being an invariant contract. Then μ^* is optimal.

Proof of Proposition 1. By Lemma 2, we can focus on invariant contracts. By Lemma 3, we can focus on contracts that maximize expected joint surplus. These two observations establish the result.■

Proposition 2. *Under either principal or agent ownership, optimal contracts involve no expropriation. Optimal contracts under principal ownership satisfy $\tau_A^* = 1$ and $\tau_P^* = \overline{\tau}^P$. Optimal contracts under agent ownership satisfy $\tau_A^* = \overline{\tau}^P$ and $\tau_P^* = 1$. Principal ownership is optimal if $b_P/c_P < b_A/c_A$.*

Proof of Proposition 2. By the preceding arguments, the statement of the proposition holds for contracts that solve the relaxed problem. It remains to show that the resulting contracts are also incentive compatible. That is, they satisfy truth-telling conditions, and neither party can benefit from a double deviation in which they announce the wrong state and disobey the resulting recommendations.

Both these claims follow because these contracts induce no expropriation. Sending message $m_i \neq \theta_i$ reduces the probability of cooperation without any benefit to party i, so these contracts

satisfy truth-telling conditions. Moreover, sending message $m_i \neq \theta_i$ and then disobeying the resulting recommendation involves an even lower probability of cooperation and possibly a costly expropriation attempt. The contracts that solve the relaxed problem are incentive-compatible and therefore solve the full problem.■

References

[1] Akerlof, Robert, Richard Holden, and Hongyi Li. 2025. Getting the Picture. *Working Paper*.

[2] Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. When Does Coordination Require Centralization? *American Economic Review*. 98(1): 145–179.

[3] Antràs, Pol and Esteban Rossi-Hansberg. 2009. Organizations and Trade. *Annual Review of Economics*. 1: 43–64.

[4] Atalay, Enghin, Ali Hortaçsu, and Chad Syverson. 2014. Vertical Integration and Input Flows. *American Economic Review*. 104(4): 1120–1148.

[5] Arrow, Kenneth. 1962. Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Incentive Activity: Economic and Social Factors*, ed. Richard Nelson, 609–626. Princeton: Princeton University Press.

[6] Arrow, Kenneth. 1975. Vertical Integration and Communication. *Bell Journal of Economics*. 6(1): 173–183.

[7] Bergemann, Dirk and Stephen Morris. 2016. Bayes Correlated Equilibrium and the Comparison of Information Structures in Games. *Theoretical Economics*. 11(2): 487–522.

[8] Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson. 2015. Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry. *American Economic Review*. 105(7): 2086–2119.

[9] Busse, Meghan. 2022. Apple's Custom Chips: A Genius Decision? Kellogg School of Management Case.

[10] Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. The Anatomy of French Production Hierarchies. *Journal of Political Economy*. 123(4): 809–852.

[11] Conconi, Paola, Fabrizio Leone, Glenn Magerman, and Catherine Thomas. 2025. Multinational Networks and Trade Participation. *Working Paper*.

[12] Dessein, Wouter. 2002. Authority and Communication in Organizations. *Review of Economic Studies*. 69(4): 811–838.

[13] Dessein, Wouter and Tano Santos. 2006. Adaptive Organizations. *Journal of Political Economy*. 114(5): 956–995.

[14] Eliason, Paul, Benjamin Heebsh, Ryan McDevitt, and James Roberts. 2020. How Acquisitions Affect Firm Behavior and Performance: Evidence from the Dialysis Industry. *Quarterly Journal of Economics*. 135(1): 221–267.

[15] Ethier, Wilfred and James Markusen. 1996. Multinational Firms, Technology Diffusion and Trade. *Journal of International Economics*. 41(1–2): 1–28.

[16] Gabaix, Xavier and David Laibson. 2006. Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *Quarterly Journal of Economics*. 121(2): 505–540.

[17] Galeotti, Andrea, Wouter Dessein, and Tano Santos. 2016. Rational Inattention and Organizational Focus. *American Economic Review*. 106(6): 1522–1536.

[18] Garicano, Luis. 2000. Hierarchies and the Organiation of Knowledge in Production. *Journal of Political Economy*. 108(5): 874–904.

[19] Grant, Robert. 1996. Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal*. 17(S2): 109–122.

[20] Grossman, Sanford and Oliver Hart. 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*. 94(4): 691–719.

[21] Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas. 2012. Innovation and Foreign Ownership. *American Economic Review*. 102(7): 3594–3627.

[22] Hart, Oliver and John Moore. 1990. Property Rights and the Nature of the Firm. *Journal of Political Economy*. 98(6): 1119–1158.

[23] La Forgia, Ambar and Julia Bodner. 2024. Getting Down to Business: Chain Ownership and Fertility Clinic Performance. *Management Science*. 71(6): 5022–5044.

[24] Liebeskind, Julia. 1996. Knowledge, Strategy, and the Theory of the Firm. *Strategic Management Journal*. 17: 93–107.

[25] Maskin, Eric and Jean Tirole. 1990. The Principal-Agent Relationship with an Informed Principal: The Case of Private Values. *Econometrica*. 58(2): 379–409.

[26] Maskin, Eric and Jean Tirole. 1992. The Principal-agent Relationship with an Informed Principal, II: Common Values. *Econometrica*. 60(1): 1–42.

[27] Maskin, Eric and Jean Tirole. 1999. Unforeseen Contingencies and Incomplete Contracts. *Review of Economic Studies*. 66(1): 83–114.

[28] Matouschek, Niko, Michael Powell, and Bryony Reich. 2025. Organizing Modular Production. *Journal of Political Economy*. 133(3): 986–1046.

[29] McGee, Patrick. 2025. Apple in China: The Capture of the World's Greatest Company. Simon and Schuster, 2025.

[30] Myerson, Roger. 1983. Mechanism Design by an Informed Principal. *Econometrica*. 51(6): 1767–1797.

[31] Rantakari, Heikki. 2008. Governing Adaptation. *Review of Economic Studies*. 75(4): 1257–1285.

[32] Riordan, Michael. 1991. Ownership without Control: Towards a Theory of Backward Integration. *Journal of Japanese and International Economies*. 5(2): 101–119.

[33] Sugaya, Takuo and Alexander Wolitzky. 2021. The Revelation Principle in Multistage Games. *Review of Economic Studies*. 88(3): 1503–1540.

[34] Taneva, Ina. 2019. Information Design. *American Economic Journal: Microeconomics*. 11(4): 151–185.

[35] Teece, David. 1980. Economies of Scope, and the Scope of the Enterprise. *Journal of Economic Behavior and Organization*. 1(3): 223–247.

[36] Tirole, Jean. 2009. Cognition and Incomplete Contracts. *American Economic Review*. 99(1): 265–294.