A Theory of Firm Boundaries and Knowledge Sharing

Evgenii Fadeev* Michael Powell†

February 9, 2026

Abstract

We study optimal asset ownership in collaborations that require sharing private knowledge. Knowledge sharing facilitates production but creates a risk of expropriation. Optimal disclosure balances these forces by restricting information flows to parties with the incentive and ability to expropriate. Asset ownership shapes these flows through two channels: it confers control over production (determining who must receive information) and control over expropriation (determining who has the means to misuse the information). Ownership can protect the owner's knowledge — making the owner more willing to disclose — while discouraging others from disclosing their knowledge to an owner who can expropriate it. We apply the theory to explain the role of third-party intermediaries that facilitate knowledge sharing among competitors, and to determine when a vertically integrated firm must spin off a subsidiary to encourage knowledge sharing from rivals, versus when internal "Chinese Walls" suffice.

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

1 Introduction

How do firm boundaries shape the transfer and use of knowledge? Recent empirical work in organizational economics documents a regularity: when one firm buys another, acquired firms often come to resemble their acquirers along a range of organizational and operational dimensions (Atalay, Hortaçsu, and Syverson, 2014; Braguinsky et al., 2015; Eliason et al., 2020; Demirer and Karaduman, 2025). These findings are typically interpreted as evidence that integration facilitates the transfer of knowledge or other intangible inputs.

One view, going back to Arrow (1975), is that it is less costly to communicate knowledge within firms than across firm boundaries. As an explanation, however, this view shifts the question rather than answering it, leaving unexplained the mechanism through which integration reduces these communication costs.[1] Instead of taking communication costs as primitive, we focus on incentives to share knowledge. We assume that knowledge can be freely shared regardless of firm boundaries, but that its subsequent use is difficult to control once it is revealed (Arrow, 1962). In particular, sharing knowledge with another party can enable non-contractible follow-on use that benefits the recipient at the expense of the original knowledge holder. We refer to this follow-on use as knowledge expropriation. The core question we study is how firm boundaries shape incentives to share knowledge and, in turn, determine its use in production.

We argue that firm boundaries matter in two related but distinct ways. To see both in a stark setting, consider the following stylized example. There are two master chefs. Each has a secret recipe for a component of a dish. If both recipes are used together, they can produce a breakthrough dish. Recipes can be disclosed freely, but disclosure creates a risk of expropriation: once one chef learns the other's recipe, the recipient can use it outside the collaboration, for example, to launch a competing dish. Following the property-rights tradition (Grossman and Hart, 1986; Hart and Moore, 1990; Hart, 1995), we model a firm boundary choice as allocating ownership — and hence control — over a key productive asset. In this example, suppose the asset is proprietary equipment required to produce the dish.

First, ownership may confer *control over production*. Suppose the dish must be produced by a single chef who owns and operates the equipment. Both chefs can expropriate each other's knowledge, but the owner can deploy their own recipe directly without disclosing it. The dish nevertheless requires both component recipes, so the operating chef must learn the other chef's recipe. Anticipating expropriation risk, the non-owner may be unwilling to fully disclose the

[1] See also Grossman and Hart (1986): "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)

secret. Swapping ownership changes which chef can act on their knowledge without disclosure and which chef must reveal information and thereby become exposed to expropriation.

Second, ownership may confer *control over expropriation*. In contrast to the previous scenario, suppose production requires mutual disclosure: both chefs must know both recipes to succeed. Crucially, however, assume that expropriation requires the proprietary equipment. In this case, the owner is willing to disclose their recipe, as the non-owner cannot commercialize the knowledge without the asset. The non-owner, by contrast, remains unwilling to fully disclose their secret, as the owner possesses the means to expropriate it. Swapping ownership therefore determines which party can safely disclose their knowledge.

This example illustrates how ownership shapes knowledge use by allocating protection against expropriation. This protection takes two forms. First, ownership can confer *control over production*: the owner's knowledge is protected because it can be applied directly without being disclosed. Second, ownership can confer *control over expropriation*: the owner's knowledge is protected because the non-owner lacks the means to misuse it. These two forms of protection have distinct implications for interpreting empirical patterns. When an acquired firm begins to resemble its acquirer, this pattern is conventionally interpreted as evidence of *knowledge transfer*. However, it may instead reflect *knowledge application*: the acquirer gains control over key decisions and applies its knowledge directly, without necessarily revealing it to the acquired firm.

This example is stylized in one important respect. In practice, ownership decisions are often shaped not only by bilateral incentives but by broader industry configurations. For instance, competing firms often share knowledge indirectly through common partners such as manufacturing suppliers (Küpper et al., 2020). Similarly, General Motors argued that its spin-off of Delphi Automotive was intended to encourage other automakers to work with Delphi — competitors that had previously been reluctant to share sensitive knowledge with a supplier owned by a rival.[2] At the same time, firms sometimes claim that internal informational "Chinese Walls" between a parent and a subsidiary can achieve the same goal by assuring competitors that their knowledge will not reach the parent (Martin and Mickle, 2017; Gawer and Cusumano, 2002; Steinhilber, 2008).

We now turn to a model that lets us study these issues in a unified way. In our model, there are multiple parties who collaborate on projects that require actions to be tailored to underlying information. Actions are complementary in the O-ring sense (Kremer, 1993): a project succeeds only if all required actions are chosen correctly. Sharing information improves cooperation, but

[2]See CNN Money (1999): "Delphi... said the spinoff will allow the company to pursue supply contracts with other automakers, which previously were reluctant to use Delphi because it was a unit of GM". Bradsher (1998) notes that rivals avoided Delphi "for fear that it would share parts designs and other competitive information with G.M."

it creates an incentive problem because disclosed information can be misused through non-contractible expropriation actions. Parties can contract over transfers, asset ownership, and an information-sharing protocol that governs what each party learns prior to choosing actions. We focus on the case in which expropriation is sufficiently costly that optimal contracts deter it.

For a fixed ownership structure, the problem reduces to designing the information structure. Following Bergemann and Morris (2019), we represent this choice as a distribution over private action recommendations. These recommendations must satisfy an obedience constraint: upon receiving a recommendation, each party must find it optimal to follow it given the information the recommendation conveys. In our setting, obedience requires both choosing the recommended productive action and refraining from expropriation.

A convenient way to summarize each party's no-expropriation constraint is through their safety threshold. Consider any party that (i) must be informed for the project to succeed and (ii) is capable of expropriation. For such a party, the no-expropriation constraint places an upper bound on the precision of the signal it can receive. We define the safety threshold as the maximum signal precision that can be disclosed while ensuring the party refrains from expropriation. This threshold is higher when expropriation is privately less attractive and lower when it is more attractive.

Our first main result is that, for any fixed ownership structure, the optimal information structure takes a nested form. Rank the required parties by their safety thresholds. The optimal contract coordinates recommendations so that whenever a party with a lower threshold receives the correct recommendation, all parties with higher thresholds also receive the correct recommendation. Nesting is efficient because actions are complementary: accuracy for a high-threshold party is wasted in states where a low-threshold party makes a mistake. The optimal contract therefore concentrates accuracy in the same states. As a result, project success is constrained by its information bottleneck: the required party with the lowest safety threshold.

We use this characterization to explain the role of ownership. Ownership operates through two channels: (i) it determines who must be informed (by allocating control over productive actions), and (ii) it determines who can expropriate (by allocating the asset required for misuse). Shifting ownership alters the set of parties who are both required for success and capable of expropriation, thereby changing the identity of the bottleneck. Specifically, ownership can relax the bottleneck constraint in two ways: by disarming a low-threshold party (removing their ability to expropriate) or by reorganizing production (removing the need to inform them). Our second main result characterizes the optimal ownership structure by how it maximizes the safety threshold of the induced information bottleneck.

We apply this characterization to several prominent industry configurations discussed above. First, we show how common suppliers and other third parties can serve as knowledge

hubs. When direct sharing among competitors would make a low-threshold rival the bottleneck, it is efficient to route information through a supplier that is safer to inform. Second, we rationalize vertical spin-offs as a way to encourage rivals to share knowledge. When an upstream unit is owned by a downstream competitor, outside customers may withhold valuable information for fear it will reach the parent (as illustrated by GM's spin-off of Delphi). Separation shifts the bottleneck to a party with weaker incentives to expropriate, unlocking cooperation. Third, we characterize when internal Chinese Walls can substitute for organizational separation. We find that informational firewalls are feasible only when the rival-facing project can be implemented without requiring the parent's operational involvement — and thus exposure to the rival's information — in a way that recreates the bottleneck. This provides a sharp sense in which Chinese Walls can be credible for some technologies but not others.

We conclude with three extensions that clarify our theory. The first two extensions discuss the key frictions preventing the first-best outcome. One friction is the *non-contractibility* of expropriation actions: we show that if parties could sign a binding no-expropriation agreement before disclosing knowledge, they could achieve the first-best regardless of ownership. By contrast, contracting over productive actions or the division of collaborative profits does not resolve the underlying incentive problem. The non-contractibility of knowledge expropriation is what makes ownership matter. We view it as a fundamental property of knowledge: once disclosed, it cannot be unlearned, and it is generally difficult to verify how it is used outside the collaboration.

The second friction is the *inseparability* of control rights associated with asset ownership. In our baseline model, ownership bundles multiple decision rights, creating unavoidable trade-offs: reallocating ownership can change which party becomes the bottleneck, but it typically cannot remove bottlenecks altogether. We show that unbundling control rights across multiple assets can achieve the first best through two distinct channels. First, when all key productive actions are alienable, unbundling productive control allows assigning each decision to the party with the relevant information, eliminating the need for disclosure. Second, when expropriation is asset-dependent, unbundling expropriation control allows assigning the means of expropriation to the potential victim, neutralizing the threat. We view inseparability of control rights as a fundamental limitation of integration as a mechanism for achieving the efficient use of knowledge: ownership reallocations typically transfer a coarse bundle of rights, rather than the targeted control needed to sustain efficient knowledge sharing.

In our third extension, we discuss the role of expropriation costs. In the baseline model, expropriation is sufficiently costly that optimal contracts deter it. When expropriation is costly but not prohibitive, it can be optimal to tolerate some expropriation in order to disclose more

information and increase the probability of successful cooperation.

Literature. We build on the incomplete-contracts approach to firm boundaries in which ownership is modeled as residual control rights (Grossman and Hart, 1986; Hart and Moore, 1990; Hart, 1995). In contrast to the classic property-rights models, inefficiencies are ex post and arise from ex post non-contractible actions (Hart and Holmstrom, 2010; Legros and Newman, 2013). Control over these actions matters in our model because it constrains who needs to know what and who is willing to share their knowledge with whom.

Within the firm-boundaries literature, our paper is closest methodologically to work that studies boundary choices through a mechanism-design approach (Matouschek, 2004; Baliga and Sjöström, 2018). Whereas these papers emphasize how ownership affects information rents when ex post actions are contractible, we study optimal ownership when ex post actions are non-contractible. The central issue is therefore about obedience rather than truth-telling. To characterize the optimal information-sharing and ownership arrangement in this environment, we use modern information-design tools (Rayo and Segal, 2010; Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019; Mathevet, Perego, and Taneva, 2020).

We also connect to the literature on communication and delegation in organizations (Dessein, 2002; Alonso, Dessein, and Matouschek, 2008; Rantakari, 2008), which studies how the allocation of control shapes communication. Because ours is a paper on firm boundaries rather than internal organization, we emphasize two features of the environment that delegation models typically abstract from: parties have access to rich contracting instruments (i.e., they can contract jointly on decision rights, information-sharing protocols, and transfers), and ownership allocates a bundle of control rights.

Finally, our paper relates to the literature on selling information when knowledge is partially excludable. A central theme in this work is how an informed seller can extract value from knowledge despite appropriation hazards through, for example, leveraging competition and credible threats of wider disclosure (Anton and Yao, 1994, 2002), or through gradual revelation that relaxes enforcement or informational constraints (Hörner and Skrzypacz, 2016; Garicano and Rayo, 2017; Fudenberg and Rayo, 2019). In our model, when ownership gives control over expropriation actions, it allows one party to share knowledge without fear of expropriation.

2 Model

There are n parties, $i \in \mathcal{N} = \{1, \ldots, n\}$, who collaborate on a collection of projects. Each project succeeds only if a set of required actions are correctly tailored to private information held by the parties. Sharing information improves cooperation but creates a risk: parties may use what they learn to expropriate others' knowledge for their private benefit.

Information. The state is $\theta = (\theta_1, \ldots, \theta_n)$, where $\theta_i \in \{-1, 1\}$ is privately observed by party i. The components of θ are independent, and each realization is equally likely, $p(\theta) = \left(\frac{1}{2}\right)^n$.

Actions and Asset Ownership. Each party $i \in \mathcal{N}$ chooses a vector of non-contractible actions. We distinguish actions along two dimensions: (i) whether they are *cooperative* or *expropriation* actions and (ii) whether they are *alienable* or *inalienable*. Cooperative actions are potentially productive. When correctly matched to the relevant component of the state, they contribute to project success. Expropriation actions are privately beneficial but socially destructive. If correctly matched to others' information, they generate a private gain for the acting party and impose a cost on others. Alienable actions require access to a productive asset. Inalienable actions reflect parties' specific human capital and thus do not require access to a productive asset.

There is a single productive asset owned by party $g \in \mathcal{N}$. We model control over alienable actions in the following way. Each party $i \in \mathcal{N}$ selects a plan for *alienable cooperative actions*, denoted $x_{ij} \in \{-1, 1\}$ for all $j \in \mathcal{N}$, and for *alienable expropriation actions*, denoted $z_{ij} \in \{-1, 0, 1\}$ for all $j \neq i$. While every party specifies a plan, only the choices of the asset owner g are payoff-relevant.

Each party $i \in \mathcal{N}$ also chooses *inalienable cooperation actions*, denoted by $\bar{x}_{ij} \in \{-1, 1\}$ for all $j \in \mathcal{N}$, and *inalienable expropriation actions*, denoted by $\bar{z}_{ij} \in \{-1, 0, 1\}$ for $j \neq i$.

For each party $i \in \mathcal{N}$, let $a_i = \left((x_{ij})_{j\in\mathcal{N}}, (\bar{x}_{ij})_{j\in\mathcal{N}}, (z_{ij})_{j\neq i}, (\bar{z}_{ij})_{j\neq i}\right)$ denote party i's *action plan*. We denote by $\mathcal{A}_i$ the set of possible action plans by party i and write $\mathcal{A} = \prod_{i\in\mathcal{N}} \mathcal{A}_i$ and $a = (a_i)_{i\in\mathcal{N}} \in \mathcal{A}$. Equivalently, we may write $a = (x, \bar{x}, z, \bar{z})$.

Some parties can expropriate only if they own the asset, while others can expropriate even if they do not. Let $\mathcal{E} \subseteq \mathcal{N}$ denote the set of parties whose expropriation opportunities are tied to the asset. We define party i's effective expropriation action against party j as ζ_{ij}. For those parties who can intrinsically expropriate, $i \notin \mathcal{E}$, $\zeta_{ij} = \bar{z}_{ij}$. For those parties in $i \in \mathcal{E}$, $\zeta_{ij} = z_{ij}$ for the asset owner $(g = i)$ and $\zeta_{ij} = 0$ for all others $(g \neq i)$.

Projects. Production is organized into a set of projects $\mathcal{K}$. Each project $k \in \mathcal{K}$ has value $V_k > 0$ if successfully implemented and value 0 otherwise. Project success requires that a set of cooperative actions be correctly matched to the relevant components of the state.

Formally, each project k has two sets of requirements. First, $\mathcal{R}_k \subseteq \mathcal{N}$ is the set of alienable requirements: for each $j \in \mathcal{R}_k$, the asset owner g must set $x_{gj} = \theta_j$. Second, $\bar{\mathcal{R}}_k \subseteq \mathcal{N} \times \mathcal{N}$ is the set of inalienable requirements: for each $(i, j) \in \bar{\mathcal{R}}_k$, party i must choose $\bar{x}_{ij} = \theta_j$. Given asset owner g, define the project-success indicator $w_k \in \{0, 1\}$ by

$$w_k\left(x, \bar{x}; g, \theta\right) = \left(\prod_{j\in\mathcal{R}_k} \mathbf{1}\left\{x_{gj} = \theta_j\right\}\right)\left(\prod_{(i,j)\in\bar{\mathcal{R}}_k} \mathbf{1}\left\{\bar{x}_{ij} = \theta_j\right\}\right).$$

Project k succeeds if and only if all of its alienable and inalienable requirements are met.

Payoffs. Party i's payoffs are quasilinear and consist of three components:

$$u_i(a,t;g,\theta) = u_i^c(x,\bar{x};g,\theta) + u_i^e(z,\bar{z};g,\theta) + t_i,$$

where u_i^c is the *cooperative payoff* from project success, u_i^e is the *expropriation payoff*, and t_i is a monetary transfer. Transfers are budget balanced, so $\sum_{i\in\mathcal{N}} t_i = 0$. We next describe the cooperative and expropriation payoff terms.

The cooperative payoff depends only on which projects succeed. We assume party i receives an exogenous share $\alpha_{ik} \geqslant 0$ of project k's value, with $\sum_{i\in\mathcal{N}} \alpha_{ik} = 1$. Its cooperative payoff is

$$u_i^c(x,\bar{x};g,\theta) = \sum_{k\in\mathcal{K}} \alpha_{ik} V_k w_k(x,\bar{x};g,\theta).$$

Thus, cooperative payoffs depend on actions only through project success and do not depend directly on asset ownership beyond its effect on which actions are payoff-relevant.

Expropriation payoffs depend on the effective expropriation actions ζ_{ij}. If party i attempts to expropriate party j's knowledge, and the attempt is successful (i.e., $\zeta_{ij} = \theta_j$), then party i obtains a private benefit $b_{ij} > 0$, while party j incurs a loss $\lambda_{ij} > 0$. If the attempt is unsuccessful (i.e., $\zeta_{ij} = -\theta_j$), party i incurs a cost $c_{ij} > 0$. If $\zeta_{ij} = 0$, both parties receive 0. Party i's expropriation payoff is therefore

$$u_i^e(z,\bar{z};g,\theta) = \sum_{j\neq i} \left[b_{ij} 1\{\zeta_{ij} = \theta_j\} - c_{ij} 1\{\zeta_{ij} = -\theta_j\} \right] - \sum_{j\neq i} \lambda_{ji} 1\{\zeta_{ji} = \theta_i\}.$$

We define total surplus as $TS(a;g,\theta) = \sum_{i\in\mathcal{N}} u_i(a,t;g,\theta)$. Because transfers are budget-balanced, they do not directly enter this expression.

We impose the following two assumptions on expropriation payoff:

Assumption 1. *Let $V_{total} = \sum_{k\in\mathcal{K}} V_k$ be the total potential value of all projects. For each party $i \in \mathcal{N}$, (i) $\lambda_{ij} - b_{ij} > V_{total}$ and (ii) $b_{ij} < c_{ij}$.*

The first part ensures any successful expropriation destroys more total surplus than can be generated by all projects. The second part ensures expropriation is risky and thus attractive only when a party has sufficiently precise information about the target's state.

Solution Concept. Consider an *omniscient* mediator who decides how information is disclosed to parties. Specifically, before the state $\theta \in \Theta$ is realized, the mediator commits to a *decision rule*

$$\sigma(\cdot \mid \theta) \in \Delta(\mathcal{A}) \qquad \text{for each } \theta \in \Theta,$$

which maps each realized state θ into a distribution over action profiles $a = (a_i)_{i\in\mathcal{N}}$. The mediator observes θ, draws an action profile a according to $\sigma(\cdot \mid \theta)$, and privately recommends the component a_i to each party i.

Recommendations are not enforceable. Instead, a decision rule must be designed so that parties optimally choose to follow their recommendations given their information. Accordingly, we restrict attention to *obedient decision rules*, under which no party wishes to deviate from the recommended action after observing its private recommendation and private state. We denote the set of obedient recommendation rules by Σ. Formally, $\sigma \in \Sigma$ if

$$\sum_{\theta_{-i}\in\Theta_{-i}} \sum_{a_{-i}\in\mathcal{A}_{-i}} p(\theta_{-i}|\theta_i)\sigma(a_i, a_{-i} \mid \theta_i, \theta_{-i}) \Big[u_i\big((a_i, a_{-i}), g; \theta_i, \theta_{-i}\big) - u_i\big((a_i', a_{-i}), g; \theta_i, \theta_{-i}\big)\Big] \geqslant 0 \tag{2.1}$$

This formulation follows the information-design-in-games approach of Bergemann and Morris (2019). An obedient decision rule σ induces a joint distribution over states and actions, $\psi(a, \theta) = p(\theta)\sigma(a \mid \theta)$. This decision rule is Bayes-plausible, $\sum_a \psi(\theta, a) = p(\theta)$, and obedience corresponds to the requirement that the induced distribution constitutes a Bayes correlated equilibrium, as captured by (2.1).

We study *optimal decision rules* that maximize total surplus subject to obedience. Let

$$TS(g) \equiv \max_{\sigma\in\Sigma} \mathbb{E}\left[TS(a, g; \theta)\right], \tag{2.2}$$

where the expectation is taken with respect to the distribution $\psi(a, \theta) = p(\theta)\sigma(a \mid \theta)$. *Optimal asset ownership* is g^* that solves $\max_{g\in\mathcal{N}} TS(g)$.

Discussion of Modeling Assumptions. The model relies on several assumptions that we relax and discuss in the extensions and appendix.

First, our solution concept assumes an omniscient mediator who observes the full state and issues private action recommendations that parties have incentives to follow (obedience). We justify this formulation as a reduced form of a general contracting game. Suppose parties can contract over (i) (possibly random) asset ownership, (ii) monetary transfers, and (iii) the behavior of a neutral mediator. A contract is a *coordination mechanism* in the sense of Myerson (1982, 1983): parties send messages with the mediator, the mediator enforces the contractible components (ownership and transfers), sends recommendations, and parties then choose non-contractible actions. A feasible mechanism must satisfy several incentive constraints. In addition to obedience described above, it must satisfy truth-telling (parties report their states truthfully), rule out profitable double deviations that combine misreporting with disobedience, and satisfy participation constraints ensuring that each party prefers acceptance to the disagreement payoff.

In Appendix A.1, we show that the solution to the full mechanism-design problem coincides with the solution concept defined in (2.2). Viewed this way, the omniscient mediator formulation provides a convenient characterization of the outcomes induced under optimal mechanisms.

Second, we assume that parties receive exogenous shares of project values, $\{\alpha_{ik}\}_{i,k}$. Section 5.1 allows these shares (and cooperative actions) to be contractible and shows that doing so does not increase the maximum attainable surplus.

Third, the baseline model assumes a single productive asset that bundles control over all alienable actions. Section 5.2 allows for multiple assets and for cooperative and expropriation control rights to be allocated separately. This extension clarifies which results are driven by bundling control in a single asset and which reflect more general properties of information sharing under expropriation risk.

Finally, we assume that successful expropriation destroys total surplus. Assumption 1 implies that expropriation is never optimal and therefore does not occur on the equilibrium path, even though the threat of expropriation shapes information sharing and ownership. This assumption is stronger than necessary for our main results; Appendix A.2 derives a weaker condition in the proof of Proposition 1. Section 5.3 also discusses the case in which expropriation is costly but not prohibitively costly, $0 < \lambda_{ij} - b_{ij} < V_{\text{total}}$.

3 Analysis

The fundamental constraint on total surplus is the threat of expropriation. Party i can profitably attempt to expropriate party j's knowledge only if it has both the *means* to do so (i.e., access to an effective expropriation action) and sufficiently precise *information*. As we show below, it is optimal to ensure that no party simultaneously has the means and sufficiently precise information to expropriate. We refer to this requirement as the *disarmament principle*.

The disarmament principle can be achieved either by limiting information disclosure or by reallocating asset ownership. In Section 3.1, we derive the optimal obedient decision rule for a fixed ownership structure and show that information disclosed to parties capable of expropriation must be capped. In Section 3.2, we analyze optimal ownership and show how it shapes both which parties must be informed for project success and which parties possess asset-dependent means of expropriation. Thus, a party can be disarmed either by removing it from the production process or by denying it access to the means of expropriation.

3.1 Optimal Information Disclosure

Fix an owner $g \in \mathcal{N}$ and an obedient decision rule $\sigma(\cdot \mid \theta) \in \Delta(\mathcal{A})$. After observing private state θ_i and recommendation a_i, party i forms posterior beliefs about others' states. For $j \neq i$, let

$$\tau_{ij}(\theta_i, a_i) \equiv \mathbb{P}_\psi\big(\theta_j = 1 \,|\, \theta_i, a_i\big),$$

where the posterior is defined under the joint distribution $\psi(\theta, a) = p(\theta)\sigma(a|\theta)$. When the dependence on (θ_i, a_i) is clear from context, we write τ_{ij} for $\tau_{ij}(\theta_i, a_i)$.

If party i can engage in expropriation, it will choose $\zeta_{ij} = 1$ if $\tau_{ij} b_{ij} - (1 - \tau_{ij}) c_{ij} > 0$ and $\zeta_{ij} = -1$ if $(1 - \tau_{ij}) b_{ij} - \tau_{ij} c_{ij} > 0$. Combining these inequalities, expropriation occurs if and only if the belief is sufficiently precise:

$$\max\{\tau_{ij}, 1 - \tau_{ij}\} > \tau^\star_{ij} \equiv \frac{c_{ij}}{b_{ij} + c_{ij}} > \frac{1}{2},$$

where we assume that in the case of indifference, parties do not expropriate.

Before stating the general results on the structure of the optimal decision rule, it is useful to build intuition with two examples.

Example 1: Optimal Decision Rule For One Party. Suppose there are two parties, $\mathcal{N} = \{1, 2\}$, and one project, $\mathcal{K} = \{1\}$. The project does not require an asset, $\mathcal{R}_1 = \varnothing$, but requires party 1 to match the state of party 2 via its inalienable action, $\bar{\mathcal{R}}_1 = \{(1, 2)\}$. That is, the project succeeds if and only if $\bar{x}_{12} = \theta_2$. Parties do not require the asset to engage in expropriation, $\mathcal{E} = \varnothing$. In this example, the asset ownership is irrelevant. Suppose both parties receive half of the project's value, $\alpha_{11} = \alpha_{12} = 1/2$.

The optimal decision rule can be found using the techniques from Kamenica and Gentzkow (2011). Specifically, denote by τ the posterior belief of party 1 that $\theta_2 = 1$ after receiving the recommendation from the mediator. Figure 1 illustrates the expected total surplus as a function τ. It is optimal for party 1 to select $\bar{x}_{12} = 1$ if $\tau > 1/2$ and $\bar{x}_{12} = -1$ if $\tau < 1/2$. If party 1's information about θ_2 is sufficiently precise, it also optimally engages in expropriation. Specifically, If $\tau > \tau^\star_{12}$, party 1 selects $\zeta_{12} = \bar{z}_{12} = 1$, and $\tau < \tau^\star_{12}$, party 1 selects $\zeta_{12} = \bar{z}_{12} = -1$.

Under expropriation, the total expected surplus is negative. Thus, the optimal decision rule is to randomize between posteriors $\tau^\star_{12}$ and $1 - \tau^\star_{12}$. It can be constructed as follows. Suppose there is a random variable $\xi \sim U[0, 1]$ independent of θ. The recommendations to party 1 are

$$x_{12} = \bar{x}_{12} = \begin{cases} \theta_2 & \text{if } \xi \leqslant \tau^\star_{12} \\ -\theta_2 & \text{if } \xi > \tau^\star_{12} \end{cases} \quad \text{and} \quad z_{12} = \bar{z}_{12} = 0$$

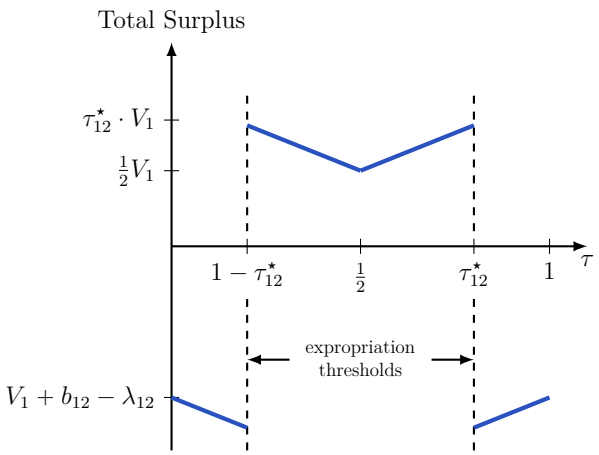


Figure 1: Total Surplus As a Function of Beliefs (Example 1)

This figure illustrates the relationship between total expected surplus and party 1's belief that $\theta_2 = 1$ for Example 1.

Example 2: Nested Decision Rule For Multiple Parties. Suppose there are three parties, $\mathcal{N} = \{1, 2, 3\}$, and two projects, $\mathcal{K} = \{1, 2\}$. Neither project has alienable requirements, so $\mathcal{R}_1 = \mathcal{R}_2 = \varnothing$. Project 1 requires two inalienable actions from parties 1 and 2 to match the state of party 3, so $\bar{\mathcal{R}}_1 = \{(1, 3), (2, 3)\}$. Project 2 requires only an inalienable action from party 2 to match the state of party 3, so $\bar{\mathcal{R}}_2 = \{(2, 3)\}$. All parties possess inalienable means of expropriation ($\mathcal{E} = \varnothing$), so asset ownership plays no role in this example.

Suppose expropriation parameters are $b_{12} = 2$, $c_{13} = 3$, $b_{23} = 1$, and $c_{23} = 9$. Then

$$\tau_{13}^\star = \frac{3}{2+3} = 0.6 \qquad \text{and} \qquad \tau_{23}^\star = \frac{9}{1+9} = 0.9.$$

To deter expropriation, party i's recommendation for the action $\bar{x}_{i3}$ must be such that its posterior belief τ_{i3} about θ_3 lies in the safety region $[1 - \tau_{i3}^\star, \tau_{i3}^\star]$. Following Example 1, a natural obedient rule is to recommend the correct match to party i with probability $\tau_{i3}^\star$ and the incorrect match with probability $1-\tau_{i3}^\star$. However, with two parties required for Project 1, the key question is whether these recommendations should be generated independently.

If parties 1 and 2 receive independent recommendations satisfying their respective safety thresholds, then Project 1 succeeds with probability $\tau_{13}^\star \cdot \tau_{23}^\star$, while Project 2 succeeds with probability $\tau_{23}^\star$. Hence the expected total surplus is

$$V_1 \cdot \tau_{13}^\star \cdot \tau_{23}^\star + V_2 \cdot \tau_{23}^\star = V_1 \cdot 0.48 + V_2 \cdot 0.9$$

The mediator can do strictly better by *positively correlating* recommendations. Intuitively, party 1 is the bottleneck in Project 1, so whenever party 1 is recommended the correct action,

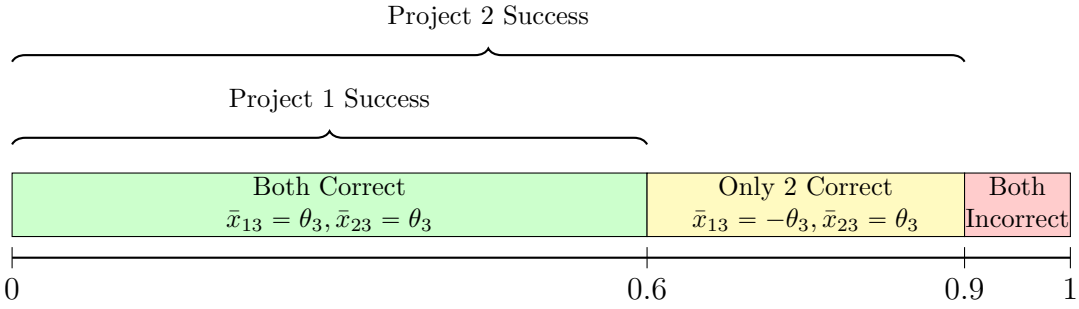


Figure 2: Nested Decision Rule For Example 3

The figure illustrates the realization of recommendations based on the random variable $\xi \sim U[0,1]$. Party 1 receives the correct recommendation only when $\xi \leqslant 0.6$. Party 2 receives the correct recommendation when $\xi \leqslant 0.9$. This decision rule ensures that Party 2 is always informed whenever Party 1 is informed.

party 2 should also be recommended the correct action.

Let $\xi \sim U[0,1]$ be a *common* randomization device. Consider the nested rule in which, for each $i \in \{1,2\}$, the mediator recommends $\bar{x}_{i3} = \theta_3$ if $\xi \leqslant \tau^\star_{i3}$ and $\bar{x}_{i3} = -\theta_3$ otherwise. This rule is illustrated in Figure 2. Under it, Project 1 succeeds with probability $\min\{\tau^\star_{13}, \tau^\star_{23}\} = \tau^\star_{13}$, while Project 2 succeeds with probability $\tau^\star_{23}$. Thus, expected total surplus is

$$V_1 \cdot \min\{\tau^\star_{13}, \tau^\star_{23}\} + V_2 \cdot \tau^\star_{23} = V_1 \cdot 0.6 + V_2 \cdot 0.9,$$

which strictly exceeds the independent benchmark. Proposition 1 below shows that such nested decision rules — implemented via a common randomization device — are optimal more generally.

General Results. Fix an owner $g \in \mathcal{N}$. Let $\mathrm{Arm}(g)$ denote the set of "armed" parties, that is parties, who can engage in expropriation under owner $g \in \mathcal{N}$:

$$\mathrm{Arm}(g) = \{i \in \mathcal{N} : i = g \text{ or } i \notin \mathcal{E}\}.$$

This set includes the owner and all parties with inalienable means for expropriation ($i \notin \mathcal{E}$).

Disarmed parties can be fully informed about others because they cannot expropriate. For armed parties, by contrast, information must be capped at the expropriation threshold. Let $\tau^{\mathrm{safe}}_{ij}(g)$ denote the maximum precision of party i's belief about θ_j that is consistent with no expropriation:

$$\tau^{\mathrm{safe}}_{ij}(g) = \begin{cases} \tau^\star_{ij} = \frac{c_{ij}}{b_{ij}+c_{ij}} & \text{if } i \in \mathrm{Arm}(g) \\ 1 & \text{if } i \notin \mathrm{Arm}(g) \end{cases}.$$

For convenience, we set $\tau^{\mathrm{safe}}_{ii}(g) \equiv 1$ for all $i \in \mathcal{N}$ and $g \in \mathcal{N}$.

For project k, it is useful to represent the production requirements as *matching tasks* of the

form "party i must match party j's state." Let

$$\mathcal{Q}_k(g) \equiv \{(g, j) : j \in \mathcal{R}_k\} \ \cup \ \bar{\mathcal{R}}_k \ \subseteq \ \mathcal{N} \times \mathcal{N}$$

denote the set of such tasks that must be performed for project k to succeed under ownership g. In Examples 1 and 2, these sets are $\mathcal{Q}_k(g) = \{(1, 2)\}$ and $\mathcal{Q}_k(g) = \{(1, 3), (2, 3)\}$, respectively. Since these examples feature only inalienable requirements (i.e., $\mathcal{R}_k = \varnothing$), the set of matching tasks is independent of the asset owner.

The next proposition establishes that for each project $k \in \mathcal{K}$ the maximum attainable expected surplus is determined by the *bottleneck task*: the lowest safety threshold $\tau_{ij}^{\text{safe}}(g)$ among tasks in $(i, j) \in \mathcal{Q}_k(g)$.

Proposition 1. *Fix an asset owner g. The maximum attainable expected total surplus is*

$$TS(g) = \sum_{k \in \mathcal{K}} V_k \cdot \left(\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{safe}(g) \right). \tag{3.1}$$

Moreover, there exists an optimal obedient decision rule with a nested structure: whenever a party with a lower safety threshold receives accurate recommendations, all parties with higher safety thresholds receive accurate recommendations as well. The mediator draws a single common random variable $\xi \sim U[0, 1]$ and, conditional on (θ, ξ), privately recommends:

(i) No expropriation: each party i is recommended to choose $z_{ij} = \bar{z}_{ij} = 0$ for all $j \neq i$.

(ii) Cooperation: party i is recommended the correct match ($x_{ij} = \bar{x}_{ij} = \theta_j$) if and only if $\xi \leqslant \tau_{ij}^{\text{safe}}(g)$, and is recommended the opposite match otherwise ($x_{ij} = \bar{x}_{ij} = -\theta_j$).

Proof. See Section A.2 in the appendix. □

The intuition behind Proposition 1 is as follows. The mediator prefers to avoid expropriation because it destroys total surplus. To do so, information must be partially withheld from *armed* parties — those who have the means to expropriate. While withholding information reduces the probability of successful cooperation, it is necessary to prevent expropriation. Consequently, the planner discloses information to each party i about state θ_j only up to its safety threshold $\tau_{ij}^{\text{safe}}(g)$.

Since cooperative actions are complementary, the probability of successful cooperation in a project is bounded the bottleneck task: the party-target pair with the lowest safety threshold, captured by the term $\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g)$. The nested decision rule described in the proposition achieves this bound by correlating agents' information through a common randomization device. This decision rule creates a strict hierarchy of information: if a less trustworthy agent is correctly

informed about a particular state, then all more trustworthy agents are necessarily informed about this state as well.

3.2 Optimal Asset Ownership

Proposition 1 characterizes the maximum total surplus as a function of ownership structure $g \in \mathcal{N}$. Optimal asset ownership solves the following problem

$$g^* \in \arg\max_{g \in \mathcal{N}} TS(g) = \arg\max_{g \in \mathcal{N}} \left\{ \sum_{k \in \mathcal{K}} V_k \cdot \left(\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) \right) \right\}.$$

Ownership plays two roles in shaping total surplus. First, control over production operates through the task set $\mathcal{Q}_k(g)$ by determining who selects alienable cooperative actions. Second, control over expropriation operates through the safety thresholds $\tau_{ij}^{\text{safe}}(g)$ by determining who has the means to expropriate. We build intuition for these roles using two examples.

Example 3: Ownership as Control Over Production. Suppose there are two parties, $\mathcal{N} = \{1, 2\}$, and a single project, $\mathcal{K} = \{1\}$. The project requires the asset owner to match the states of both parties to succeed, so $\mathcal{R}_1 = \{1, 2\}$. No inalienable actions are required ($\bar{\mathcal{R}}_1 = \varnothing$), and all parties possess the inalienable means to expropriate ($\mathcal{E} = \varnothing$).

Given an owner $g \in \{1, 2\}$, the project succeeds if and only if $x_{g1} = \theta_1$ and $x_{g2} = \theta_2$. The owner can always match its own state, $x_{gg} = \theta_g$, without disclosing it. By contrast, the non-owner's state must be disclosed to the owner for the project to succeed. Since the owner is armed, the information disclosed about the non-owner's state must be capped to deter expropriation.

Formally, ownership affects the set of matching tasks, $\mathcal{Q}_1(g) = \{(g, 1), (g, 2)\}$. Because expropriation is inalienable ($\mathcal{E} = \varnothing$), ownership does not affect safety thresholds: $\tau_{ij}^{\text{safe}}(g) = \tau_{ij}^{\star} = \frac{c_{ij}}{b_{ij} + c_{ij}}$. Hence the asset should be owned by the party with the weaker temptation to expropriate (equivalently, the higher safety threshold):

$$g^* \in \arg\max_{g \in \mathcal{N}} \left\{ \tau_{gj}^{\star} \cdot V_1 \right\} \quad \Leftrightarrow \quad g^* \in \arg\min_{g \in \mathcal{N}} \left\{ \frac{b_{gj}}{c_{gj}} \right\}, \tag{3.2}$$

where $j \neq g$.

Example 4: Ownership as Control Over Expropriation. Suppose again there are two parties, $\mathcal{N} = \{1, 2\}$, and a single project, $\mathcal{K} = \{1\}$. The project requires the asset owner to match the states of both parties ($\mathcal{R}_1 = \{1, 2\}$). Unlike the previous case, cooperation now requires complementary inalienable actions: each party must use their human capital to match the other party's state ($\bar{\mathcal{R}}_1 = \{(1, 2), (2, 1)\}$). Furthermore, we assume that both parties require the asset to engage in expropriation ($\mathcal{E} = \{1, 2\}$).

There are two key differences from Example 3. First, cooperation now requires inalienable actions by both parties, and ownership does not affect matching tasks, $\mathcal{Q}_1(g) = \{(1,2),(2,1)\}$ for all $g \in \mathcal{N}$. Second, expropriation is alienable and controlled by the owner: the non-owner is disarmed and has no effective expropriation action. Therefore, ownership affects safety thresholds, $\tau_{ij}^{\text{safe}}(g) = \tau_{ij}^{\star}$ if $i = g$ and $\tau_{ij}^{\text{safe}}(g) = 1$ if $i \neq g$.

As in the previous example, the optimal asset owner is the party with the lowest temptation to expropriate (see (3.2)). The logic, however, differs across the two settings. In Example 1, ownership determines who participates in production. In Example 4, ownership determines who lacks the means to expropriate.

Additional Results. We characterize the optimal ownership structure for several special cases. Define the set of active participants in project $k \in \mathcal{K}$ under ownership $g \in \mathcal{N}$ as

$$\mathcal{P}_k(g) = \left\{ i \in \mathcal{N} : \exists j \neq i \text{ such that } (i,j) \in \mathcal{Q}_k(g) \right\}$$

That is, $\mathcal{P}_k(g)$ consists of parties who must match the state of some other party for project k to succeed. Furthermore, for all results below we assume that each party i has symmetric preferences regarding expropriating other parties' knowledge:

$$\tau_{ij}^{\star} \equiv \tau_i^{\star} \quad \text{for all } j \neq i \quad \text{and} \quad \tau_i^{\text{safe}}(g) \equiv \tau_{ij}^{\text{safe}}(g) \tag{3.3}$$

Under this assumption, the formula for the maximum expected surplus under ownership $g \in \mathcal{N}$ simplifies to

$$TS(g) = \sum_{k \in \mathcal{K}} V_k \cdot \min_{i \in \mathcal{P}_k(g)} \tau_i^{\text{safe}}(g)$$

Ownership and Temptation to Expropriate. First, we generalize the insight from Examples 3 and 4 regarding the relationship between ownership and the temptation to expropriate. Suppose that all active non-owners are disarmed. Formally, assume that for every ownership assignment $g \in \mathcal{N}$ and project $k \in \mathcal{K}$, any active participant other than the owner belongs to the set of agents with alienable expropriation opportunities, $\mathcal{E}$:

$$\mathcal{P}_k(g) \backslash \{g\} \subseteq \mathcal{E} \text{ for all } g \in \mathcal{N}, k \in \mathcal{K}. \tag{3.4}$$

Under this assumption, any project in which the owner is *not* an active participant can be implemented with certainty: all of its active participants are disarmed and can safely be fully informed. The only projects that are "throttled" by an information cap are those in which the

owner must act. Thus, the maximum attainable surplus simplifies to

$$TS(g) = \tau_g^\star \sum_{k\in\mathcal{K}} V_k \mathbf{1}\left\{g \in \mathcal{P}_k(g)\right\} + \sum_{k\in\mathcal{K}} V_k \mathbf{1}\left\{g \notin \mathcal{P}_k(g)\right\} = V_{\text{total}} - (1-\tau_g^\star)T_g, \tag{3.5}$$

where $T_g \equiv \sum_{k\in\mathcal{K}} V_k \mathbf{1}\left\{g \in \mathcal{P}_k(g)\right\}$ denotes the total value of projects in which owner g is an active participant.

Corollary 1 (Ownership and Temptation to Expropriate.)**.** *Suppose parties have symmetric preferences across targets (condition* (3.3)*). Suppose also that any asset owner g is an active participant in all projects ($g \in \mathcal{P}_k(g)$ for all $k \in \mathcal{K}$) and is the only participant who is armed (condition* (3.4)*). Then, the optimal asset owner is the party with the lowest temptation to expropriate (i.e., the highest safety threshold): $g^* \in \arg\max_g \tau_g^\star$.*

Proof. Condition (3.4) implies that the maximum attainable surplus under ownership g is given by (3.5). Since the owner is an active participant in all projects, the value of projects involving the owner equals the total value ($T_g = V_{\text{total}}$). Consequently, the total surplus simplifies to $TS(g) = \tau_g^\star V_{\text{total}}$. Maximizing this expression with respect to g yields the result. □

Ownership and Involvement in Projects. Next, we show that the optimal ownership depends on the parties' involvement in projects. Intuitively, if the owner is the only armed participant, ownership should not be assigned to a party whose human capital is required in many high-value projects. Making such a highly involved party the owner would "arm" it, necessitating restricted information flows across all those projects.

Corollary 2 (Ownership and Involvement in Projects.)**.** *Suppose all parties have the same temptation to expropriate ($\tau_{ij}^\star = \tau^\star$ for all i and $j \neq i$). Suppose further that for every ownership assignment $g \in \mathcal{N}$ and every project $k \in \mathcal{K}$, every non-owner who is an active participant is disarmed (condition* (3.4)*). Then, the optimal asset owner is the party that is least involved in valuable projects (i.e., the party active in projects with the lowest total value, $g^* \in \arg\min_g T_g$).*

Proof. Condition (3.4) implies that the total attainable surplus under ownership g can be written as (3.5). Since $\tau_i^\star = \tau^\star$, the optimal owner is $g^* \in \arg\max_g \{V_{\text{total}} - (1-\tau^\star)T_g\}$, which is equivalent to $g^* \in \arg\min_g T_g$. □

To see the intuition for Corollary 2 in a minimal setting, consider the following example.

Example 5: Ownership and Involvement in Projects. Suppose there are two parties, $\mathcal{N} = \{1,2\}$, and both require the asset to engage in expropriation, $\mathcal{E} = \{1,2\}$. There are two projects. Project 1 requires alienable actions by the owner, $\mathcal{R}_1 = \{1,2\}$ and $\bar{\mathcal{R}}_1 = \varnothing$. Project 2 does not require the asset, $\mathcal{R}_2 = \varnothing$, but requires party 1's inalienable action to match

party 2's state, $\bar{\mathcal{R}}_2 = \{(1,2)\}$. Suppose both parties have the same temptation for expropriation, $\tau_1^\star = \tau_2^\star = \tau^\star$.

Because $\mathcal{E} = \{1,2\}$, any non-owner is *disarmed* and has no effective expropriation action. Thus, under any ownership assignment $g \in \mathcal{N}$, the *only* armed party is the owner itself.

For Project 1, the only active participant is the asset owner, $\mathcal{P}_1(g) = \{g\}$. For Project 2, the only active participant is party 1 regardless of ownership, $\mathcal{P}_2(1) = \mathcal{P}_2(2) = \{1\}$. Thus, if party 1 owns the asset, it is active in both projects: $T_1 = V_1 + V_2$. If party 2 owns the asset, it is active only in the first project: $T_2 = V_1$.

Ownership determines who is armed and, consequently, whose information must be restricted. If party 1 owns the asset, its information about θ_2 must be capped in every project where it is active. Since it is active in both, both projects suffer:

$$TS(g=1) = \underbrace{V_1 + V_2}_{V_{\text{total}}} - (1-\tau^\star)\underbrace{(V_1 + V_2)}_{T_1} = \tau^\star(V_1 + V_2).$$

In contrast, if party 2 owns the asset, it is active only in Project 1. Thus, its information about θ_1 must be capped. However, party 1 can be fully informed about θ_2, as it is not armed.

$$TS(g=2) = \underbrace{V_1 + V_2}_{V_{\text{total}}} - (1-\tau^\star)\underbrace{V_1}_{T_2} = \tau^\star V_1 + V_2.$$

Allocating ownership to party 2 (the least involved party) yields a strictly higher surplus. This arrangement concentrates the "inefficiency of ownership" on the low-involvement party, preventing the high-involvement party (party 1) from the temptation to expropriate.

Ownership and Inalienable Expropriation. Asset ownership shapes knowledge sharing through two channels: it determines the set of active participants $\mathcal{P}_k(g)$ (production control) and the safety thresholds $\tau_i^{\text{safe}}(g)$ (expropriation control). In the next corollary, we isolate the second channel by assuming ownership does not affect the set of participants. We focus solely on ownership as a tool for reallocating the means of expropriation.

In this setting, it is optimal to allocate the asset to a party that possesses inalienable means of expropriation (i.e., a party that can expropriate regardless of ownership). Since this party is dangerous irrespective of whether they own the asset, giving them the asset does not create a new expropriation risk. Conversely, giving the asset to a party with alienable expropriation motives would turn a previously safe party into a dangerous one.

Corollary 3 (Ownership and Inalienable Expropriation)**.** *Suppose parties have symmetric preferences across targets (condition* (3.3)*). Suppose that for all projects ownership does not affect the set of active participants ($\mathcal{P}_k(g) = \mathcal{P}_k$ for all $g \in \mathcal{N}$ and all $k \in \mathcal{K}$). Then, it is*

weakly optimal to assign asset to a party with inalienable means for expropriation, if the one exists. That is, if $\mathcal{N}\backslash\mathcal{E} \neq \varnothing$, then for any $g \in \mathcal{E}$ and $h \in \mathcal{N}\backslash\mathcal{E}$, $TS(h) \geqslant TS(g)$.

Proof. Fix $g \in \mathcal{E}$ and $h \in \mathcal{N}\backslash\mathcal{E}$. By definition of safety thresholds,

$$\tau_i^{\text{safe}}(g) = \begin{cases} \tau_g^\star & \text{if } i = g, \\ 1 & \text{if } i \in \mathcal{E}\backslash\{g\}, \\ \tau_i^\star & \text{if } i \notin \mathcal{E}, \end{cases} \qquad \tau_i^{\text{safe}}(h) = \begin{cases} 1 & \text{if } i \in \mathcal{E}, \\ \tau_i^\star & \text{if } i \notin \mathcal{E}. \end{cases}$$

Thus, $\tau_i^{\text{safe}}(h) \geqslant \tau_i^{\text{safe}}(g)$. Condition $\mathcal{P}_k(g) = \mathcal{P}_k(h)$ implies $TS(h) \geqslant TS(g)$. □

Bottleneck irrelevance of ownership. Finally, we show that if a single inalienable participant is required for every project and is the most tempted among all potentially armed parties, then reallocating ownership cannot relax the binding information constraint.

Corollary 4 (Bottleneck irrelevance of ownership)**.** *Suppose parties have symmetric preferences across targets (condition* (3.3)*). Suppose there exists party i with inalienable means for expropriation that is active in all projects ($i \notin \mathcal{E}$, $i \in \mathcal{P}_k(g)$ for all $k \in \mathcal{K}$ and all $g \in \mathcal{N}$). Suppose that this party has the highest temptation to expropriate ($\tau_i^\star \leqslant \tau_\ell^\star$ for all $\ell \in \mathcal{N}$). Then, the total surplus is constant regardless of asset ownership: $TS(g) = \tau_i^\star V_{\text{total}}$.*

Proof. Because party i has inalienable means of expropriation ($i \notin \mathcal{E}$), its safety threshold is independent of ownership: $\tau_i^{\text{safe}}(g) = \tau_i^\star$ for any owner g. Since i has the highest temptation to expropriate ($\tau_i^\star \leqslant \tau_\ell^\star$ for all $\ell \in \mathcal{N}$) and is active in all projects ($i \in \mathcal{P}_k(g)$ for all $k \in \mathcal{K}$ and all $g \in \mathcal{N}$), the success probability fo every project is $\min_{j \in \mathcal{P}_k(g)} \tau_j^{\text{safe}}(g) = \tau_i^\star$. □

4 Applications

Our framework highlights that asset ownership may protect knowledge through either control over production or control over expropriation. For example, Apple chose to design the iPad chip in-house rather than partner with Intel because it "just didn't want to teach them everything, which they could go and sell to our competitors" (Steve Jobs, quoted in Isaacson, 2011, p. 493). Here, ownership served to control production and limit knowledge disclosure. In contrast, Apple's ownership choices in its China operations are better interpreted as control over expropriation. Specifically, Apple actively shared knowledge with its Chinese suppliers. Although it did not own most production facilities, it owned key production machines, which limited suppliers' ability to expropriate Apple's know-how (McGee, 2025).

We also show ownership has costs: non-owners may not be willing to fully disclose their information to the owner. For example, China's quid pro quo policy requires multinationals to form joint ventures with local partners to access the Chinese market. This structure gives the local partner control over some manufacturing decisions, forcing the multinational to disclose production-relevant information. One implication of this policy is that it may discourage automakers from bringing their most advanced technology to China due to concerns about expropriation (Bai et al., 2025).

In this section, we illustrate how our results translate into predictions about optimal ownership and information sharing in several industry environments. In all applications, there are two downstream parties, 1 and 2, who possess payoff-relevant information and compete in the sense that they are tempted to expropriate each other's knowledge. A third upstream party, 3, can collaborate with them and is comparatively less tempted to engage in expropriation.

Formally, suppose that parties' expropriation incentives are symmetric across targets and let $\tau_1^\star \equiv \tau_{12}^\star$, $\tau_2^\star \equiv \tau_{21}^\star$, $\tau_3^\star \equiv \tau_{31}^\star = \tau_{32}^\star$ denote their safety thresholds. Party 3 is less tempted to engage in expropriation if

$$\tau_3^\star > \max\{\tau_1^\star, \tau_2^\star\} \tag{4.1}$$

Section 4.1 shows how the upstream party can act as a *knowledge hub*. Two competing downstream parties can benefit from pooling knowledge, but they are too tempted to expropriate each other's secrets. A third party that contributes neither private information nor human capital may nonetheless optimally hold the asset solely because it can safely aggregate the rivals' information. We use this case to rationalize knowledge sharing through third-party suppliers and solution providers, as commonly observed in the auto industry (Küpper et al., 2020).

Section 4.2 studies a setting in which downstream firms do not need to share knowledge with each other, but each project requires an inalienable input from the upstream party. We show why firms are reluctant to share information with a competitor-owned subsidiary, and why a vertically integrated firm may optimally spin off its supplier to unlock value from collaboration with rivals. This logic helps explain the *vertical spin-off* of Delphi from General Motors and provides a rationale for separating Intel's foundry business from its chip-design operations (Fitch, 2025).

Section 4.3 modifies the spin-off setting so that one downstream firm's project does not require the asset owner's active participation. In that case, vertical integration with an informational *Chinese Wall* becomes feasible and optimal: the subsidiary can learn a rival's secrets without those secrets reaching the parent. Comparing this case to Section 4.2 yields a simple technological condition distinguishing when Chinese Walls are feasible from when they are impossible, requiring a spin-off to maximize surplus.

4.1 Knowledge Hubs

Consider an extension of Example 1 involving three parties, $\mathcal{N} = \{1, 2, 3\}$, and a single project, $\mathcal{K} = \{1\}$. The project requires the asset owner to match the states of parties 1 and 2 to succeed, so $\mathcal{R}_1 = \{1, 2\}$. No inalienable actions are required, $\bar{\mathcal{R}}_1 = \varnothing$, and all parties possess the inalienable means to expropriate, so $\mathcal{E} = \varnothing$.

In this environment, only the owner's alienable cooperative actions are payoff-relevant. The planner can therefore keep the two non-owners uninformed, eliminating their expropriation incentives. The remaining constraint is the owner's own temptation: to implement the project, the owner must hold enough information about θ_1 and θ_2 to match them, but not so much as to trigger expropriation. Proposition 1 then implies that the implementable total surplus under owner g is proportional to $\tau_g^\star$, and the optimal owner is the least tempted party: $g^* \in \arg\max_{g \in \mathcal{N}} \tau_g^\star$, so that $g^* = 3$ under (4.1).

The distinctive feature of this example is that the optimal owner (party 3) contributes neither private information nor inalienable inputs to production: θ_3 is payoff-irrelevant and $\bar{\mathcal{R}}_1 = \varnothing$. Party 3 is valuable solely because it has a weaker temptation to expropriate. In this sense, it functions as a *knowledge hub*, enabling greater information pooling among competing firms.

4.2 Optimal Spin-offs

Suppose there are three parties, $\mathcal{N} = \{1, 2, 3\}$, and two projects, $\mathcal{K} = \{1, 2\}$. We interpret parties 1 and 2 as downstream competitors and party 3 as a common upstream supplier. Project $k \in \{1, 2\}$ requires two inputs: an alienable decision by the asset owner and an inalienable input by the supplier. Formally, $\mathcal{R}_1 = \{1\}$, $\bar{\mathcal{R}}_1 = \{(3, 1)\}$, and $\mathcal{R}_2 = \{2\}$, $\bar{\mathcal{R}}_2 = \{(3, 2)\}$. Thus, for project k, the asset owner must match the downstream party k's state, and the supplier must match the same state using its specific human capital.

Expropriation incentives differ by party. The supplier requires the asset to expropriate (alienable expropriation), while the downstream firms can expropriate using inalienable means: $\mathcal{E} = \{3\}$ and $\mathcal{N} \backslash \mathcal{E} = \{1, 2\}$.

We map ownership to organizational form. If a downstream party $i \in \{1, 2\}$ owns the asset, then the supplier is a subsidiary (vertical integration). If the supplier owns the asset ($g = 3$), it is an independent upstream firm.

Given an asset owner $g \in \{1, 2, 3\}$, the cooperative surplus is

$$V_1 \cdot \mathbf{1}\{x_{g1} = \theta_1\} \cdot \mathbf{1}\{\bar{x}_{31} = \theta_1\} + V_2 \cdot \mathbf{1}\{x_{g2} = \theta_2\} \cdot \mathbf{1}\{\bar{x}_{32} = \theta_2\}$$

Because expropriation destroys surplus, the optimal information policy avoids expropriation

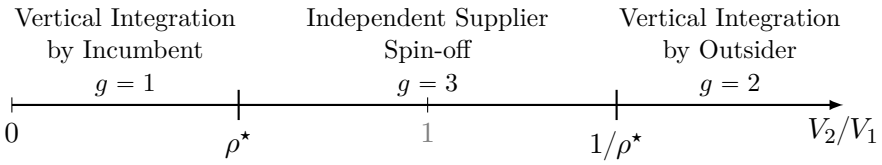


Figure 3: Optimal Organizational Form

The optimal owner varies with the relative market size V_2/V_1. For this figure, we assume that $\rho^\star < 1$.

and trades it off against the probability of successful cooperation. The key difference across ownership structures is *who must learn a rival's information*. If party 1 owns the asset, then project 1 can be implemented with certainty: party 1 knows θ_1 and the supplier is disarmed (so θ_1 can be safely disclosed to party 3 to satisfy $\bar{x}_{31} = \theta_1$). Project 2 is different. To make the payoff-relevant alienable decision for project 2, the owner must respond to θ_2. Since party 1 is armed and has incentives to expropriate party 2's knowledge, information about θ_2 must be capped at party 1's safety threshold $\tau_1^\star$.

The same logic applies symmetrically if party 2 owns the asset. If instead the supplier owns the asset, then the supplier is armed and must receive (capped) information about both downstream parties' states, so both projects succeed with probability $\tau_3^\star$.

Consequently, the maximal attainable surplus under each ownership assignment is

$$\begin{aligned} TS(g=1) &= V_1 + \tau_1^\star V_2, \\ TS(g=2) &= \tau_2^\star V_1 + V_2, \\ TS(g=3) &= \tau_3^\star (V_1 + V_2), \end{aligned}$$

To sharpen the comparative statics, suppose the downstream parties are symmetric in temptation, so $\tau_1^\star = \tau_2^\star \equiv \tau_D^\star$. Recall that the supplier is more trustworthy than a direct competitor, $\tau_3^\star > \tau_D^\star$. Define the threshold

$$\rho^\star \equiv \frac{1 - \tau_3^\star}{\tau_3^\star - \tau_D^\star}.$$

and assume that $\rho^\star < 1$. Then vertical integration by the incumbent ($g = 1$) is optimal when the outside market is not too large, $V_2/V_1 < \rho^\star$. When the outside market is very large, $V_2/V_1 > 1/\rho^\star$, vertical integration by the outsider ($g = 2$) is optimal. Finally, when the outside market is of intermediate size, $V_2/V_1 \in [\rho^\star, 1/\rho^\star]$, the independent supplier (a spin-off, $g = 3$) is optimal.[3]

[3]If $\rho^\star > 1$, the the optimal ownership is $g^* = 1$ if $V_2/V_1 \leqslant 1$, $g^* = 2$ if $1 < V_2/V_1 < \rho^\star$, and $g^* = 3$ if $V_2/V_1 \geqslant \rho^\star$.

Figure 3 summarizes the resulting pattern. When the outside market is small, it is optimal for the incumbent to integrate the supplier, because doing so disarms the supplier and lets the incumbent's own project succeed with certainty. As the outside market grows, the opportunity cost of throttling the outsider's project becomes large. At that point, shifting ownership to the (more trustworthy) supplier relaxes the binding information constraint and makes both projects feasible at the higher probability $\tau_3^\star$. Finally, when the outside market is very large, the logic flips and vertical integration by the outsider becomes optimal.

4.3 Chinese Walls Within Organizations

We now modify the setting of Section 4.2 in one important way. Project 1 remains unchanged, but Project 2 no longer requires an alienable action by the asset owner. Formally, $\mathcal{R}_1 = \{1\}$, $\bar{\mathcal{R}}_1 = \{(3,1)\}$, and $\mathcal{R}_2 = \varnothing$, $\bar{\mathcal{R}}_2 = \{(3,2)\}$. Project 2 now requires *only* the supplier's inalienable input; the owner does not need to match θ_2.

This change alters the feasible information flows. If party 1 owns the asset ($g = 1$), the supplier is disarmed (because $\mathcal{E} = \{3\}$ and $3 \neq g$). The supplier can therefore be fully informed about both θ_1 and θ_2 and can supply inalienable inputs for both projects. Crucially, because the owner is not an active participant in project 2 (the project has no alienable requirements), party 1 does not need to learn θ_2. The organization can therefore sustain a perfect informational partition: the subsidiary (party 3) learns the rival's information to serve it, while the parent (party 1) remains ignorant, eliminating the expropriation risk.

The resulting attainable surplus levels are

$$\begin{aligned} TS(g=1) &= V_1 + V_2, \\ TS(g=2) &= \tau_2^\star V_1 + V_2, \\ TS(g=3) &= \tau_3^\star (V_1 + V_2). \end{aligned}$$

Thus, when $g = 1$, the first-best outcome is achieved.

Comparing this case to Section 4.2 clarifies when Chinese Walls are credible. They are technologically impossible when the owner must be an active participant in the rival's project (because then the owner must be informed about the rival's state). They become credible precisely when the owner can remain operationally passive in the rival's project, so the rival's information can be routed to a disarmed subsidiary without reaching the armed parent.

We note that establishing the theoretical possibility of a Chinese Wall is distinct from the question of how to build it. Questions regarding the specific internal governance mechanisms and incentive systems required to enforce separation between a parent and its subsidiary lie beyond the scope of this framework and represent an exciting avenue for future research.

5 Extensions

This section discusses two extensions of the baseline model. Section 5.1 discusses two conditions under which the first best can be achieved regardless of asset ownership: (i) separability of information, so that information relevant for cooperation can be shared without revealing information useful for expropriation, and (ii) contractibility of expropriation actions. Section 5.2 introduces separation of control rights across multiple assets. Section 5.3 considers the case in which expropriation costs are small enough that tolerating some expropriation can be optimal in order to disclose more information.

5.1 Contracts vs. Ownership

In this subsection, we discuss conditions under which parties can contractually achieve the first best regardless of asset ownership.

Informational Separability. Our baseline assumption is that the same private information that is needed for productive cooperation is also useful for expropriation. In that case, disclosing information to enable cooperation can simultaneously create incentives to expropriate, so the mediator must *ration* information to armed parties, and the attainable surplus may fall below the maximum possible surplus V_{total}.

By contrast, if cooperation-relevant and expropriation-relevant information are *separable*, then the mediator can reveal the information needed for cooperation while keeping beliefs about expropriation targets diffuse, so that no one has an incentive to expropriate. In that case, the first best can be achieved regardless of asset ownership.

Contractible Actions. If some actions are contractible, then whenever the mediator recommends those actions they are enforced, which relaxes the obedience constraints. Thus, contractibility weakly increases the maximum attainable total surplus.

If *expropriation* actions $(z, \bar{z})$ are contractible, then *before an information is revealed* the parties can directly commit to $z_{ij} = \bar{z}_{ij} = 0$ and eliminate the destructive margin. The mediator can then disclose whatever information is needed for cooperation and implement the first best under any ownership structure. In contrast, if only *cooperative* actions $(x, \bar{x})$ or cooperative payoffs (shares α_{ik}) are contractible while expropriation remains non-contractible, then the fundamental obstacle is unchanged: armed parties may still have incentives to deviate toward expropriation after learning too much. As a result, the maximal attainable surplus remains governed by the same obedience-based bound as in the baseline problem (i.e., it cannot exceed $\max_g TS(g)$). The details are provided in Appendix A.4.

5.2 Multiple Assets and Separation of Control Rights

In the baseline model, a single asset bundles multiple alienable decision rights. This bundling generates the central trade-off: ownership protects the owner's knowledge but leaves others exposed to expropriation, discouraging disclosure. This subsection extends the model to $m \geqslant 1$ assets and allows cooperative and expropriation control rights to be unbundled.

Assets. Let $\mathcal{M} = \{1, \ldots, m\}$ be the set of assets. Two assignment functions determine which asset governs which *alienable* decision:

$$\phi^c : \mathcal{N} \to \mathcal{M}, \qquad \phi^e : \mathcal{N} \to \mathcal{M}.$$

Here $\phi^c(j)$ is the asset whose owner controls the alienable *cooperative* action targeting state θ_j, while $\phi^e(j)$ is the asset whose owner controls the alienable *expropriation* action targeting θ_j (for asset-dependent expropriators). The baseline model corresponds to $m = 1$ and $\phi^c = \phi^e \equiv 1$. In general, control over production decisions is bundled when $\phi^c(i) = \phi^c(j)$ for $i \neq j$; control over production and expropriation is bundled when $\phi^c(j) = \phi^e(j)$.

An ownership profile is $g = (g_i)_{i \in \mathcal{M}} \in \mathcal{N}^{\mathcal{M}}$. For each target j, define the induced owners

$$g^c(j) \equiv g_{\phi^c(j)}, \qquad g^e(j) \equiv g_{\phi^e(j)}.$$

Production. Project k succeeds if and only if the relevant alienable and inalienable matching requirements are satisfied:

$$x_{g^c(j)\, j} = \theta_j \quad \forall j \in \mathcal{R}_k, \qquad \bar{x}_{\ell j} = \theta_j \quad \forall (\ell, j) \in \bar{\mathcal{R}}_k.$$

The set of matching tasks for project k under ownership profile g is

$$\mathcal{Q}_k(g) \equiv \{(g^c(j), j) : j \in \mathcal{R}_k\} \ \cup \ \bar{\mathcal{R}}_k. \tag{5.1}$$

Expropriation. A party i can expropriate target $j \neq i$ either via inalienable means ($i \notin \mathcal{E}$) or via an asset if $i \in \mathcal{E}$ and $i = g^e(j)$. Thus the set of parties who are armed *against target* j is

$$\mathrm{Arm}_j(g) \equiv \{i \in \mathcal{N} : i = g^e(j) \text{ or } i \notin \mathcal{E}\}.$$

Define the target-specific safety threshold (and set $\tau_{ii}^{\text{safe}}(g) \equiv 1$)

$$\tau_{ij}^{\text{safe}}(g) \;\equiv\; \begin{cases} \tau_{ij}^{\star} = \dfrac{c_{ij}}{b_{ij} + c_{ij}} & \text{if } i \in \text{Arm}_j(g) \text{ and } i \neq j, \\ 1 & \text{otherwise.} \end{cases} \tag{5.2}$$

Total Surplus. Under the same parameter restrictions as in Proposition 1 (Assumption 1), the characterization of the maximum attainable total surplus carries over:

$$TS(g) = \sum_{k \in \mathcal{K}} V_k \cdot \min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g),$$

where $\mathcal{Q}_k(g)$ and $\tau_{ij}^{\text{safe}}(g)$ are defined in (5.1) and (5.2), respectively.

Since the baseline model corresponds to $m = 1$ and $\phi^c = \phi^e \equiv 1$, allowing separable control rights weakly increases maximum attainable total surplus. The following proposition shows how unbundling control rights across assets can achieve the first best.

Proposition 2 (First best under separable control rights)**.** *The first best V_{total} is achievable in each of the following cases.*

(i) All cooperative requirements are alienable ($\bar{\mathcal{R}}_k = \varnothing$ for all k), and distinct targets are controlled by distinct cooperative assets (ϕ^c is injective).

(ii) All expropriation is asset-dependent ($\mathcal{E} = \mathcal{N}$), distinct targets are governed by distinct expropriation assets (ϕ^e is injective).

Proof. (i) Because ϕ^c is injective, we can choose an ownership profile g such that $g_{\phi^c(j)} = j$ for all $j \in \mathcal{N}$, i.e. $g^c(j) = j$ for all j. Since $\bar{\mathcal{R}}_k = \varnothing$, each project has task set $\mathcal{Q}_k(g) = \{(j, j) : j \in \mathcal{R}_k\}$, which is trivial because party j observes θ_j. Hence $\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) = 1$ for every k and therefore $TS(g) = \sum_k V_k = V_{\text{total}}$.

(ii) Because ϕ^e is injective, choose g such that $g_{\phi^e(j)} = j$ for all j, i.e. $g^e(j) = j$. Fix any $i \neq j$. Since $\mathcal{E} = \mathcal{N}$, party i has no inalienable expropriation; and because $g^e(j) = j \neq i$, party i also does not control the expropriation asset for target j. Hence $i \notin \text{Arm}_j(g)$ and $\tau_{ij}^{\text{safe}}(g) = 1$ for all $i \neq j$. It follows that for every project k, $\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) = 1$, so again $TS(g) = V_{\text{total}}$. □

The intuition behind Proposition 2 is as follows. When cooperation requires only alienable actions and ownership can be arranged so that each action that depends on party i's information is controlled by party i, then no private information needs to be disclosed. Parties can therefore avoid expropriation because they can remain silent while still implementing the required cooperative decisions. This case extends the mechanism in Example 3 of Section 3.2 (ownership as control over production).

When expropriation requires access to assets and ownership can be arranged so that each party controls the asset needed to expropriate *its own* information, then parties can block all expropriation attempts. In that case, the first best is achievable because the means of expropriation are held by the potential victim rather than the potential expropriator. This case extends the mechanism in Example 4 of Section 3.2 (ownership as control over expropriation).

5.3 Marginal Value of Information Under Expropriation

In the baseline model, we assume that expropriation destroys total surplus, $\lambda_{ij} - b_{ij} > V_{\text{total}}$. This assumption ensures that the mediator always prefers to deter expropriation. In this subsection, we discuss the omniscient mediator's problem under the weaker assumption that expropriation is costly but not necessarily prohibitively costly, $\lambda_{ij} - b_{ij} > 0$. Under this weaker condition, it may be optimal to allow expropriation with positive probability for some parties. We give intuition here; Appendix A.5 provides details.

In Appendix A.5, we show that the maximum attainable total surplus under owner $g \in \mathcal{N}$ has the following bound

$$TS(g) \leqslant \max_{\frac{1}{2} \leqslant q_{ij} \leqslant 1} \left\{ \underbrace{\sum_{k \in \mathcal{K}} V_k \cdot \min_{(i,j) \in \mathcal{Q}_k(g)} q_{ij}}_{\text{Cooperation Benefits}} - \underbrace{\sum_{i \in \text{Arm}(g)} \sum_{j \neq i} (\lambda_{ij} - b_{ij}) \max\left\{0, \frac{q_{ij} - \tau^\star_{ij}}{1 - \tau^\star_{ij}}\right\}}_{\text{Expropriation Costs}} \right\}. \tag{5.3}$$

The variable q_{ij} is the precision of the information disclosed to party i about party j's state. In the baseline model, expropriation costs are sufficiently high ($\lambda_{ij} - b_{ij} > V_{\text{total}}$) that the solution involves no expropriation: $q_{ij} = \tau^\star_{ij}$ for armed parties ($i \in \text{Arm}(g)$) and $q_{ij} = 1$ for disarmed parties. The assumption $\lambda_{ij} - b_{ij} > V_{\text{total}}$ is stronger that necessary to ensure that no-expropriation is optimal, and the proof of Proposition 1 derives a weaker bound.

When expropriation costs are small enough, the choice of information precisions q_{ij} is governed by a marginal benefit–marginal cost comparison. Increasing q_{ij} is valuable only for projects in which task (i, j) is a bottleneck; in those projects, the marginal benefit equals the sum of project values for which (i, j) is currently the bottleneck. If i is disarmed, providing information to i is costless. If i is armed, the marginal cost of increasing q_{ij} is zero up to the safety threshold $\tau^\star_{ij}$ and becomes positive for $q_{ij} > \tau^\star_{ij}$, reflecting the increasing probability of expropriation. Figure 4 illustrates this marginal benefit–marginal cost logic for information provision to party i about state j.

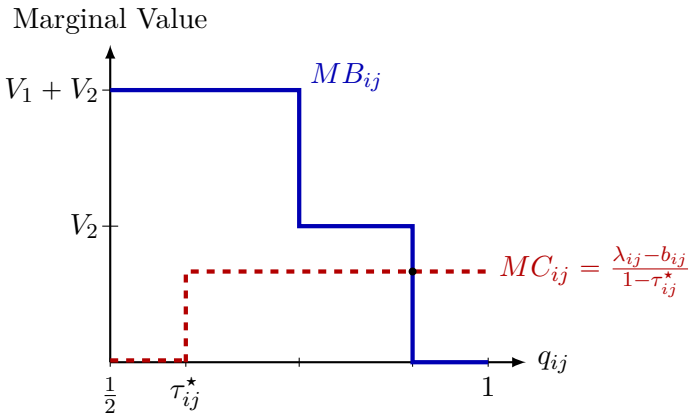


Figure 4: Marginal Value of Information

6 Conclusion

TBD.

References

Alonso, R., Dessein, W., & Matouschek, N. (2008). When Does Coordination Require Centralization? *American Economic Review*, *98*(1), 145–179.

Anton, J. J., & Yao, D. A. (1994). Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights. *The American Economic Review*, *84*(1), 190–209. https://www.jstor.org/stable/2117978

Anton, J. J., & Yao, D. A. (2002). The Sale of Ideas: Strategic Disclosure, Property Rights, and Contracting. *The Review of Economic Studies*, *69*(3), 513–531.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. In R. Nelson (Ed.), *The rate and direction of incentive activity: Economic and social factors* (pp. 609–626). Princeton University Press.

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

Atalay, E., Hortaçsu, A., & Syverson, C. (2014). Vertical Integration and Input Flows. *American Economic Review*, *104*(4), 1120–1148.

Bai, J., Barwick, P. J., Cao, S., & Li, S. (2025). Quid Pro Quo, Knowledge Spillovers, and Industrial Quality Upgrading: Evidence from the Chinese Auto Industry. *American Economic Review*, *115*(11), 3825–3852. https://doi.org/10.1257/aer.20221501

Baliga, S., & Sjöström, T. (2018). *A Theory of the Firm Based on Haggling, Coordination, and Rent-Seeking* (Working Paper). Kellogg School of Management, Northwestern University.

Bergemann, D., & Morris, S. (2019). Information Design: A Unified Perspective. *Journal of Economic Literature*, *57*(1), 44–95.

Bradsher, K. (1998). G.M. Plans to Spin Off Parts Division. *The New York Times*. https://www.nytimes.com/1998/08/04/business/gm-plans-to-spin-off-parts-division.html

Braguinsky, S., Ohyama, A., Okazaki, T., & Syverson, C. (2015). Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry. *American Economic Review*, *105*(7), 2086–2119.

CNN Money. (1999, May). Delphi Spinoff Completed. https://money.cnn.com/1999/05/31/companies/gm/

Demirer, M., & Karaduman, Ö. (2025). Do Mergers and Acquisitions Improve Efficiency? Evidence from Power Plants. *Working Paper*.

Dessein, W. (2002). Authority and Communication in Organizations. *Review of Economic Studies*, *69*(4), 811–838.

Eliason, P., Heebsh, B., McDevitt, R., & Roberts, J. (2020). How Acquisitions Affect Firm Behavior and Performance: Evidence from the Dialysis Industry. *Quarterly Journal of Economics*, *135*(1), 221–267.

Fitch, A. (2025). Is Nvidia Intel's Savior? Not Quite. *The Wall Street Journal*. https://www.wsj.com/tech/intel-nvidia-deal-strategy-two-companies-ec7d7632

Fudenberg, D., & Rayo, L. (2019). Training and Effort Dynamics in Apprenticeship. *American Economic Review*, *109*(11), 3780–3812. https://doi.org/10.1257/aer.20171939

Garicano, L., & Rayo, L. (2017). Relational Knowledge Transfers. *American Economic Review*, *107*(9), 2695–2730. https://doi.org/10.1257/aer.20160194

Gawer, A., & Cusumano, M. A. (2002). *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Harvard Business School Press.

Grossman, S., & Hart, O. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, *94*(4), 691–719.

Hart, O. (1995). *Firms, Contracts, and Financial Structure (Clarendon Lectures in Economics)*. Oxford University Press.

Hart, O., & Holmstrom, B. (2010). A Theory of Firm Scope. *The Quarterly Journal of Economics*, *125*(2), 483–513.

Hart, O., & Moore, J. (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy*, *98*(6), 1119–1158.

Hörner, J., & Skrzypacz, A. (2016). Selling Information. *Journal of Political Economy*, *124*(6), 1515–1562. https://doi.org/10.1086/688874

Isaacson, W. (2011). *Steve Jobs*. Simon & Schuster.

Kamenica, E., & Gentzkow, M. (2011). Bayesian Persuasion. *American Economic Review*, *101*(6), 2590–2615.

Kremer, M. (1993). The O-Ring Theory of Economic Development. *Quarterly Journal of Economics*, *108*, 551–575.

Küpper, D., Okur, A., Betti, F., Bezamat, F., Fendri, M., & Fernandez, B. (2020). How Manufacturers Can Unlock Value from Data Sharing. https://www.bcg.com/publications/2020/manufacturers-unlock-value-from-data-sharing

Legros, P., & Newman, A. F. (2013). A Price Theory of Vertical and Lateral Integration. *The Quarterly Journal of Economics*, *128*(2), 725–770.

Martin, T. W., & Mickle, T. (2017). Why 'Apple Rival' Samsung Also Wins If iPhone X Is a Hit. *The Wall Street Journal*. https://www.wsj.com/articles/why-apple-rival-samsung-also-wins-if-iphone-x-is-a-hit-1506936602

Mathevet, L., Perego, J., & Taneva, I. (2020). On Information Design in Games. *Journal of Political Economy*, *128*(4), 1370–1404. https://doi.org/10.1086/705332

Matouschek, N. (2004). Ex Post Inefficiencies in a Property Rights Theory of the Firm. *The Journal of Law, Economics, and Organization*, *20*(1), 125–147.

McGee, P. (2025). *Apple in China: The Capture of the World's Greatest Company*. Simon; Schuster.

Myerson, R. (1982). Optimal Coordination Mechanisms in Generalized Principal-Agent Problems. *Journal of Mathematical Economics*, *10*(1), 67–81.

Myerson, R. (1983). Mechanism Design by an Informed Principal. *Econometrica*, *51*(6), 1767–1797.

Rantakari, H. (2008). Governing Adaptation. *Review of Economic Studies*, *75*(4), 1257–1285.

Rayo, L., & Segal, I. (2010). Optimal Information Disclosure. *Journal of Political Economy*, *118*(5), 949–987. https://doi.org/10.1086/657922

Steinhilber, S. (2008). *Strategic Alliances: Three Ways to Make Them Work*. Harvard Business Press.

A Appendix

A.1 Implementation Through a Mediated Contract

Recall that each party $i \in \mathcal{N} = \{1, \ldots, n\}$ privately observes a binary state $\theta_i \in \Theta_i \equiv \{-1, 1\}$. Let $\Theta \equiv \prod_{i \in \mathcal{N}} \Theta_i$ and assume the common prior is independent and uniform:

$$p(\theta) = \prod_{i \in \mathcal{N}} p_i(\theta_i), \qquad p_i(\theta_i) = \tfrac{1}{2}.$$

For each party i, the set of feasible action plans is $\mathcal{A}_i \equiv \{-1, 1\}^{2n} \times \{-1, 0, 1\}^{2(n-1)}$. A non-contractible action plan $a_i \in \mathcal{A}_i$ can be written as

$$a_i = \Big((x_{ij})_{j \in \mathcal{N}},\, (\bar{x}_{ij})_{j \in \mathcal{N}},\, (z_{ij})_{j \neq i},\, (\bar{z}_{ij})_{j \neq i} \Big),$$

and we write $\mathcal{A} \equiv \prod_{i \in \mathcal{N}} \mathcal{A}_i$ and $a = (a_i)_{i \in \mathcal{N}} \in \mathcal{A}$.

Ownership $g \in \mathcal{N}$ is contractible. Transfers are contractible and budget balanced:

$$t \in \mathcal{T} \equiv \Big\{ \tilde{t} \in \mathbb{R}^n : \sum_{i \in \mathcal{N}} \tilde{t}_i = 0 \Big\}.$$

Let a non-transfer *outcome* be

$$o = (a, g) \in \mathcal{O} \equiv \mathcal{A} \times \mathcal{N}.$$

Given $o = (a, g)$ and state θ, payoffs are $\tilde{u}_i(o, t; \theta) = \tilde{u}_i(a, g, t; \theta) = u_i(a, g; \theta) + t_i$, and total surplus is

$$TS(a, g; \theta) \;\equiv\; \sum_{i \in \mathcal{N}} u_i(a, g; \theta),$$

where transfers cancel out.

Contracts. We model contracting as a mediated *coordination mechanism* in the sense of Myerson (1982, 1983). Production and expropriation actions $(x, \bar{x}, z, \bar{z})$ are non-contractible, but parties can contract on (i) asset ownership $g \in \mathcal{N}$ (control of payoff-relevant alienable actions), (ii) monetary transfers $t \in \mathcal{T}$, and (iii) the behavior of a neutral mediator who can receive messages from the parties and privately communicate recommendations to them.

A *direct mediated contract* is a pair (φ, t) consisting of:

1. an *allocation rule* $\varphi(\cdot \mid \hat{\theta}) \in \Delta(\mathcal{O})$ for each report profile $\hat{\theta} \in \Theta$; and
2. a *transfer rule* $t : \mathcal{O} \times \Theta \to \mathcal{T}$, where $t(o, \hat{\theta})$ is the transfer vector implemented when the mediator draws outcome $o = (a, g)$ after receiving report profile $\hat{\theta}$.

The interpretation is as follows. After θ is realized, each party i reports $\hat{\theta}_i \in \Theta_i$ to the mediator. Given $\hat{\theta}$, the mediator draws $(a, g) \sim \varphi(\cdot \mid \hat{\theta})$, enforces ownership g and transfers $t((a, g), \hat{\theta})$, and privately recommends a_i to party i. Parties then simultaneously choose their (non-contractible) action plans in $\mathcal{A}_i$.

Bargaining and timing. Fix an initial owner $g_0 \in \mathcal{N}$. The timing of the game is the following:

1. Party 1 makes a take-it-or-leave-it offer of a contract to the other parties.
2. Each party $i \in \{2, \ldots, n\}$ either accepts or rejects the contract.
3. State $\theta \in \Theta$ is realized. Party i observes θ_i.
4. If all participates accepted, they send their reports $\hat{\theta}_i \in \Theta_i$ to the mediator. The mediator draws $o \in \mathcal{O}$ according to $\varphi(\cdot \mid \hat{\theta})$, enforces contractible decisions and sends private recommendations for non-contractible decisions. If at least one party rejects, then no mediator is appointed, the asset owner is the initial owner g_0, no transfers are shared, and no information is shared. In this disagreement outcome, all parties optimally choose no expropriation and set $x_{ij} = 1$ and $\bar{x}_{ij} = 1$ for all (i, j).
5. Parties observe enforced ownership and transfers.
6. Parties choose their actions.
7. Payoffs are realized.

Incentive Constraints. Following Myerson (1982, 1983), we restrict attention to direct mechanisms and impose incentive compatibility in the sense that truthful reporting and obedience to recommendations form a Bayesian Nash equilibrium.

Fix a contract (φ, t). When all parties report truthfully and follow their recommendations, party i's interim expected utility is

$$U_i(\varphi, t \mid \theta_i) \equiv \sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{(a,g) \in \mathcal{O}} \varphi(a, g \mid \theta_i, \theta_{-i}) \Big[u_i(a, g; \theta_i, \theta_{-i}) + t_i\big((a, g), (\theta_i, \theta_{-i})\big) \Big].$$

Parties are not compelled to follow recommendations; a (pure) deviation for party i is a function $\delta_i : D_i \to \mathcal{A}_i$ mapping the available information at the action stage into an action plan, where $D_i \equiv N \times \mathcal{T} \times \mathcal{A}_i \times \Theta_i$, and party i observes $d_i \equiv (g, t, a_i, \theta_i) \in D_i$. For any alternative

report $\theta_i' \in \Theta_i$ and any deviation function $\delta_i : D_i \to \mathcal{A}_i$, define the *double-deviation* payoff

$$U_i^*(\varphi, t, \delta_i, \theta_i' \mid \theta_i) \;\equiv\; \sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{(a,g) \in \mathcal{O}} \varphi(a, g \mid \theta_i', \theta_{-i}) \Big[u_i\big((\delta_i(g, t', a_i, \theta_i), a_{-i}), g; \theta_i, \theta_{-i}\big) + t_i\big((a, g), (\theta_i', \theta_{-i})\big) \Big],$$

where $t' \equiv t((a, g), (\theta_i', \theta_{-i}))$ is the transfer vector implemented after the report (θ_i', θ_{-i}).

The contract (φ, t) is *incentive compatible* if, for every $i \in \mathcal{N}$,

$$U_i(\varphi, t \mid \theta_i) \;\geqslant\; U_i^*(\varphi, t, \delta_i, \theta_i' \mid \theta_i) \qquad \text{for all } \theta_i, \theta_i' \in \Theta_i \text{ and all } \delta_i : D_i \to \mathcal{A}_i. \tag{A.1}$$

This condition subsumes truth-telling (take $\delta_i(d_i) = a_i$) and obedience (take $\theta_i' = \theta_i$) as special cases.

Participation Constraints. Let $o^0 = (\hat{a}^0, g_0)$ denote the disagreement outcome, where $\hat{a}^0$ recommends $x_{ij} = 1$, $\bar{x}_{ij} = 1$ and $z_{ij} = \bar{z}_{ij} = 0$ for all feasible pairs (i, j). The disagreement transfers are zero for all parties, $t_i^0 = 0$. Let $\underline{U}_i$ denote party i's ex-ante expected payoff in disagreement:

$$\underline{U}_i \equiv \sum_{\theta \in \Theta} p(\theta)\, u_i(a^0, g_0; \theta),$$

where $p(\theta) = \frac{1}{2^n}$. A contract (φ, t) is *individually rational* if

$$\sum_{\theta_i \in \Theta_i} p(\theta_i)\, U_i(\varphi, t \mid \theta_i) \;\geqslant\; \underline{U}_i \qquad \text{for all } i \in \mathcal{N} \backslash \{1\}. \tag{A.2}$$

Party 1 can always guarantee itslef $\underline{U}_1$ by offering the null contract; thus we focus on the constraints (A.2) for $i \neq 1$.

The mechanism-design program. Party 1 chooses a contract (φ, t) to maximize her ex-ante expected utility subject to incentive compatibility and participation:

$$\begin{aligned} \max_{\varphi, t} \quad & \sum_{\theta_1 \in \Theta_1} p(\theta_1) U_1(\varphi, t \mid \theta_1) \\ \text{subject to} \quad & \text{(A.1) and (A.2)}, \\ & \varphi(o \mid \theta) \geqslant 0, \quad \sum_{o \in \mathcal{O}} \varphi(o \mid \theta) = 1 \quad \text{for all } \theta \in \Theta, \\ & \sum_{i \in \mathcal{N}} t_i(a, g, \theta) = 0 \quad \text{for all } (a, g, \theta) \in \mathcal{A} \times \mathcal{N} \times \Theta. \end{aligned} \tag{A.3}$$

Relaxed Problem: Obedience Only

Consider the relaxation of (A.3) obtained by dropping the truth-telling constraints (and hence also the double-deviation constraints). In this relaxation, we evaluate incentives under the *truthful* report profile and impose only obedience of the recommended non-contractible actions, together with individual rationality.

Because transfers are enforced before actions are chosen and cannot depend on the realized (non-contractible) actions, a deviation by party i affects only the non-monetary payoff component. Thus, the obedience constraints are: for every $i \in \mathcal{N}$, every $\theta_i \in \Theta_i$, every $g \in \mathcal{N}$, every recommended action plan $a_i \in \mathcal{A}_i$, and every deviation $a_i' \in \mathcal{A}_i$,

$$\sum_{\theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{a_{-i}} \varphi\left(a_i, a_{-i}, g \mid \theta_i, \theta_{-i}\right) \left[u_i\left(a_i, a_{-i}, g; \theta\right) - u_i\left(a_i', a_{-i}, g; \theta\right)\right] \geqslant 0 \tag{A.4}$$

Since transfers are budget balanced and payoffs are quasilinear, we have for any contract

$$\mathbb{E}[\tilde{u}_1] = \mathbb{E}\left[\sum_{i \in \mathcal{N}} u_i(a, g; \theta)\right] - \sum_{i \neq 1} \mathbb{E}[\tilde{u}_i] = \mathbb{E}[W(a, g; \theta)] - \sum_{i \neq 1} \mathbb{E}[\tilde{u}_i].$$

Under ex-ante participation, $\mathbb{E}[\tilde{u}_i] \geqslant \underline{U}_i$ for all $i \neq 1$, so party 1's ex-ante payoff is bounded by

$$\mathbb{E}[\tilde{u}_1] \;\leqslant\; \mathbb{E}[TS(a, g; \theta)] - \sum_{i \neq 1} \underline{U}_i.$$

Moreover, for any obedient allocation rule φ (hence any induced distribution of (θ, g, a) under obedient play), we can choose *constant* balanced transfers $t \in \mathcal{T}$ so that $\mathbb{E}[\tilde{u}_i] = \underline{U}_i$ for every $i \neq 1$. Indeed, letting $\bar{U}_i \equiv \mathbb{E}[u_i(a, g; \theta)]$ denote party i's expected non-monetary payoff under φ, set

$$t_i = \underline{U}_i - \bar{U}_i \quad (i \neq 1), \qquad t_1 = -\sum_{i \neq 1} t_i.$$

These transfers do not affect obedience (A.4) because the same t_i is added to both sides of any action-deviation comparison.

Therefore, maximizing party 1's expected payoff subject to obedience and ex-ante participation is equivalent (up to the additive constant $-\sum_{i \neq 1} \underline{U}_i$) to maximizing expected total surplus subject only to obedience constraint:

$$\begin{aligned} \max_{\varphi} \quad & \sum_{\theta \in \Theta} p(\theta) \sum_{g \in \mathcal{N}} \sum_{a \in \mathcal{A}} \varphi(a, g \mid \theta)\, TS(a, g; \theta) \\ \text{subject to} \quad & \text{(A.4)}, \qquad \varphi(a, g \mid \theta) \geqslant 0, \quad \sum_{a, g} \varphi(a, g \mid \theta) = 1 \;\text{ for all } \theta \in \Theta. \end{aligned} \tag{A.5}$$

For any *fixed* ownership assignment $g \in \mathcal{N}$, if we restrict attention to allocation rules that put probability one on g (i.e., $\varphi(a, g' \mid \theta) = 0$ for $g' \neq g$), then (A.5) coincides with the solution-concept problem (2.2) defining $TS(g)$.

Lemma 1 (Deterministic ownership)**.** *In the relaxed (obedience-only) problem* (A.5)*, allowing randomization of ownership g does not increase the maximal attainable total surplus. There exists an optimal solution with deterministic ownership.*

Proof. Let φ be any feasible allocation rule for (A.5), and let $G \in \mathcal{N}$ denote the (possibly random) owner induced by φ (which is publicly observed before actions are chosen). Let $\psi(a, g) = p(\theta)\varphi(a, g \mid \theta)$ denote the join distribution of (a, g, θ) induced by φ. For each g with $\mathbb{P}_\psi(G = g) > 0$, define the conditional decision rule $\varphi^g(\cdot \mid \theta) \in \Delta(\mathcal{A})$ by

$$\varphi^g(a \mid \theta) \equiv \mathbb{P}_\psi(a \mid \theta,\, G = g) = \frac{\psi(a, g, \theta)}{\sum\limits_a \psi(a, g, \theta)}.$$

Because parties observe $G = g$ before choosing actions, obedience of φ implies that, conditional on $G = g$, the induced rule φ^g satisfies the obedience constraints under ownership g. Hence the conditional expected surplus cannot exceed $TS(g)$:

$$\mathbb{E}_\psi[TS(a, G; \theta) \mid G = g] \leqslant TS(g).$$

Taking expectations over G yields

$$\mathbb{E}_\psi[TS(a, G; \theta)] = \sum_{g \in \mathcal{N}} \Pr(G = g)\, \mathbb{E}_\psi[TS \mid G = g] \;\leqslant\; \sum_{g \in \mathcal{N}} \Pr(G = g)\, TS(g) \;\leqslant\; \max_{g \in \mathcal{N}} TS(g).$$

Finally, choosing $G \equiv g^* = \arg\max_g TS(g)$ and using an optimal obedient recommendation rule that attains $TS(g^*)$ achieves the upper bound. This proves that an optimal solution exists with deterministic ownership. □

The solution to the relaxed (obedience-only) problem in (2.2) with deterministic ownership provides an upper bound for the mechanism-design problem (A.3). Let σ be an optimal decision rule for the problem in (2.2), and let g^* denote an optimal owner, $g^* \in \arg\max_g TS(g)$.

Consider the following contract (φ, t):

$$\varphi(a, g \mid \theta) = \begin{cases} \sigma(a \mid \theta) & \text{if } g = g^* \\ 0 & \text{if } g \neq g^* \end{cases} \quad \text{for all } (a, \theta) \in \mathcal{A} \times \Theta \tag{A.6}$$

$$t_i(a, g, \theta) = t_i = \underline{U}_i - \mathbb{E}_\psi \left[u_i(a, g; \theta) \right] \quad \text{for } i \neq 1 \text{ and all } (a, g, \theta) \in \mathcal{A} \times \mathcal{N} \times \Theta \tag{A.7}$$

$$t_1(a, g, \theta) = -\sum_{i=2}^{n} t_i \tag{A.8}$$

This contract assigns ownership deterministically to g^* and induces the same distribution over (a, θ) as the optimal obedient rule σ that solves (2.2). In Section A.2, we characterize the optimal decision rule (Proposition 1) and show that it attains the maximum expected surplus in (2.2), thereby establishing an upper bound for (A.3). Next, in Section A.3, we show that the contract (A.6)–(A.8) satisfies the incentive-compatibility constraints and therefore achieves this upper bound.

A.2 Proof of Proposition 1

Recall that $\mathcal{R}_k$ and $\bar{\mathcal{R}}_k$ are alienable and inalienable requirements, respectively, and $\mathcal{Q}_k$ is the set of production requirements

$$\mathcal{Q}_k(g) \equiv \{(g, j) : j \in \mathcal{R}_k\} \ \cup \ \bar{\mathcal{R}}_k \ \subseteq \ \mathcal{N} \times \mathcal{N}.$$

Proposition 1 states that the maximum attainable expected total surplus under owner $g \in \mathcal{N}$ is

$$TS(g) = \sum_{k \in \mathcal{K}} V_k \cdot \left(\min_{i \in \mathcal{P}_k(g)} \tau_{ij}^{\text{safe}}(g) \right). \tag{A.9}$$

Moreover, there exists an optimal obedient decision rule that is *nested*. Let $\xi \sim U[0, 1]$ be a common randomization device. For each (ξ, θ, g) define an action profile $a(\xi, \theta, g) = (x, \bar{x}, z, \bar{z})$ as follows:

$$z_{ij} = 0 \ \text{ and } \ \bar{z}_{ij} = 0 \qquad \text{for all } i \in \mathcal{N}, \ j \neq i, \tag{A.10}$$

$$x_{ij} = \bar{x}_{ij} = \begin{cases} \theta_j & \text{if } \xi \leqslant \tau_{ij}^{\text{safe}}(g), \\ -\theta_j & \text{if } \xi > \tau_{ij}^{\text{safe}}(g), \end{cases} \qquad \text{for all } i \in \mathcal{N}. \tag{A.11}$$

Define $\sigma^g(\cdot \mid \theta) \in \Delta(\mathcal{A})$ as the distribution induced by $\xi \sim U[0,1]$ and $a = a(\xi, \theta; g)$; that is,

$$\sigma^g(a \mid \theta) = \mathbb{P}_\xi\big[a(\xi, \theta; g) = a\big].$$

We prove the proposition in two lemmas. First, we show this decision rule is obedient and achieves $TS(g)$. Second, we show that for any obedient decision rule, the expected total surplus does not exceed $TS(g)$.

Lemma 2 (Nested rule is obedient and achieves $TS(g)$)**.** *Fix ownership $g \in \mathcal{N}$. The decision rule σ^g is obedient under ownership g. Moreover, under σ^g each project k succeeds with probability $\min_{(i,j)\in\mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g)$, and hence*

$$\mathbb{E}_\psi\Big[TS(a, g; \theta)\Big] = \sum_{k\in\mathcal{K}} V_k \cdot \min_{(i,j)\in\mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) = TS(g),$$

where $\psi(\theta, a) = p(\theta)\sigma^g(a \mid \theta)$ is the induced distribution over actions and states.

Proof. We verify obedience and compute the induced success probabilities.

Posteriors. For each $j \neq i$, define the *recommended sign* for j contained in recommendation a_i by

$$s_{ij} \equiv x_{ij} = \bar{x}_{ij} \in \{-1, 1\},$$

where $x_{ij} = \bar{x}_{ij}$ under (A.11). By construction of the decision rule, either a recommended sign for target j is correct with probability $\tau_{ij}^{\text{safe}}(g)$ or incorrect with probability $1 - \tau_{ij}^{\text{safe}}(g)$. In particular, for any $j \neq i$,

$$\mathbb{P}_\psi(\theta_j = s_{ij} \mid \theta_i, a_i) = \tau_{ij}^{\text{safe}}(g), \qquad \mathbb{P}_\psi(\theta_j = -s_{ij} \mid \theta_i, a_i) = 1 - \tau_{ij}^{\text{safe}}(g).$$

No expropriation. The recommendation rule always prescribes no expropriation (A.10). If i is disarmed, then every effective expropriation action is 0 and there is no deviation. If i is armed, then $\tau_{ij}^{\text{safe}}(g) = \tau_{ij}^\star = c_{ij}/(b_{ij} + c_{ij})$ by definition, so conditional on (θ_i, a_i) the maximal success probability from attempting expropriation against any target is at most $\tau_{ij}^\star$, implying that the expected private gain from expropriation is weakly non-positive. With the tie-breaking rule favoring the safe action, obeying (A.10) is optimal.

Obedient Cooperation. Now consider deviations that change i's cooperative actions $(x_i, \bar{x}_i)$. For any project k, let $J_{i,k}(g)$ denote the set of indices $j \neq i$ such that $(i, j) \in \mathcal{Q}_k(g)$, i.e. the states that i must match (via either x_{ij} if $i = g$ and $j \in \mathcal{R}_k$, or via $\bar{x}_{ij}$ if $(i, j) \in \bar{\mathcal{R}}_k$). If $J_{i,k}(g) = \varnothing$, then i's cooperative action does not affect project k.

Fix a project k where $J_{i,k}(g) \neq \varnothing$. Under the nested rule, if all parties follow their recommendations, project k succeeds if and only if $\xi \leqslant \tau_{ij}^{\text{safe}}(g)$ for all required tasks (i,j), which is equivalent to:

$$\xi \leqslant \min_{(i,j)\in\mathcal{Q}k(g)} \tau_{ij}^{\text{safe}}(g). \tag{A.12}$$

Since $\tau_{ij}^{\text{safe}}(g) > 1/2$ for all pairs, this probability is strictly greater than $1/2$.

Now suppose party i deviates from its recommendation for a subset of targets $\tilde{J}_{i,k}(g) \subseteq J_{i,k}(g)$. Party i successfully matches the states if and only if the recommendation was correct for the non-deviated targets ($j \notin \tilde{J}_{i,k}(g)$) and incorrect for the deviated targets ($j \in \tilde{J}_{i,k}(g)$). In terms of the randomization variable ξ, this requires:

$$\xi \in \Big(\max_{j\in\tilde{J}_{i,k}(g)} \tau_{ij}^{\text{safe}}(g), \min_{j\in J_{i,k}(g)\setminus\tilde{J}_{i,k}(g)} \tau_{ij}^{\text{safe}}(g) \Big], \quad \text{with the convention } \min\varnothing \equiv 1.$$

Since $\tau_{ij}^{\text{safe}}(g) > 1/2$, the lower bound of this interval is strictly greater than $1/2$. Therefore, the length of the interval — and thus the success probability — is strictly less than $1/2$. Consequently, any deviation yields a strictly lower probability of success than obedience.

Total surplus. Multiplying the probability of the event in (A.12) by V_k and summing over all projects yields the stated expression for $TS(g)$. □

Lemma 3 (Upper bound for any obedient rule)**.** *Fix ownership $g \in \mathcal{N}$. Let $\sigma(\cdot \mid \theta)$ be any obedient decision rule under g, and let $\psi(\theta,a) = p(\theta)\sigma(a \mid \theta)$ be the induced distribution. Then*

$$\mathbb{E}_\psi\left[TS(a,g;\theta)\right] \leqslant TS(g).$$

Proof. For each task $(i,j) \in \mathcal{Q}_k(g)$ define the event that this task is satisfied:

$$C_{k,(i,j)} \equiv \begin{cases} \{(\theta,a) : x_{ij} = \theta_j\} & \text{if } (i,j) = (g,j) \in \{(g,j) : j \in \mathcal{R}_k\} \text{ and } (i,j) \notin \bar{\mathcal{R}}_k, \\ \{(\theta,a) : \bar{x}_{ij} = \theta_j\} & \text{if } (i,j) \in \bar{\mathcal{R}}_k \text{ and } (i,j) \notin \{(g,j) : j \in \mathcal{R}_k\}, \\ \{(\theta,a) : x_{ij} = \bar{x}_{ij} = \theta_j\} & \text{if } (i,j) = (g,j) \in \{(g,j) : j \in \mathcal{R}_k\} \cap \bar{\mathcal{R}}_k. \end{cases} \tag{A.13}$$

Let E_k denote the event that project k succeeds:

$$E_k \equiv \bigcap_{(i,j)\in\mathcal{Q}_k(g)} C_{k,(i,j)}.$$

We write total surplus as

$$TS(a, g; \theta) = TS^{\text{coop}}(a, g; \theta) + TS^{\text{exp}}(a, g; \theta),$$

where $TS^{\text{coop}}(a, g; \theta) = \sum_{k \in \mathcal{K}} V_k \, \mathbf{1}\{E_k\}$, and TS^{exp} is the sum of all expropriation terms.

Upper Bound on Task Probabilities. Let

$$r_{ij} \equiv \mathbb{P}_\psi \left(\zeta_{ij} \neq 0 \right)$$

be the probability that party i attempts an expropriation of party j's information, $j \neq i$. If $i \notin \text{Arm}(g)$, then party i cannot engage in expropriation and $r_{ij} = 0$.

Decompose $\mathbb{P}_\psi(C_{k,(i,j)})$ according to whether i attempts to expropriate j:

$$\mathbb{P}_\psi(C_{k,(i,j)}) = \mathbb{P}_\psi(C_{k,(i,j)}, \zeta_{ij} = 0) + \mathbb{P}_\psi(C_{k,(i,j)}, \zeta_{ij} \neq 0)$$

On the event $\{\zeta_{ij} = 0\}$, obedience implies that attempting expropriation is not strictly profitable for party i. Therefore, conditional on (θ_i, a_i), the probability that any sign choice satisfies the matching task is at most $\tau_{ij}^{\text{safe}}(g)$, so

$$\mathbb{P}_\psi \left(C_{k,(i,j)} \mid \zeta_{ij} = 0 \right) \leqslant \tau_{ij}^{\text{safe}}(g) \;\Rightarrow\; \mathbb{P}_\psi(C_{k,(i,j)}, \zeta_{ij} = 0) \leqslant \tau_{ij}^{\text{safe}}(g) \mathbb{P}_\psi \left(\zeta_{ij} = 0 \right) = \tau_{ij}^{\text{safe}}(g)(1 - r_{ij})$$

In addition,

$$\mathbb{P}_\psi(C_{k,(i,j)}, \zeta_{ij} \neq 0) \leqslant \mathbb{P}_\psi \left(\zeta_{ij} \neq 0 \right) = r_{ij}$$

Combining these two inequalities, we get

$$\mathbb{P}_\psi(C_{k,(i,j)}) \leqslant \tau_{ij}^{\text{safe}}(g)(1 - r_{ij}) + r_{ij}$$

Upper bound on cooperative surplus. The goal is to show that the probability of project k's success has the following upper bound

$$\mathbb{P}_\psi \left(E_k \right) \leqslant \min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g)$$

Fix project $k \in \mathcal{K}$. If $\mathcal{Q}_k(g)$ contains only trivial tasks of the form (i, i), then $\min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) = 1$, so the desired bound holds.

Fix any project k with non-trivial tasks: $(i, j) \in \mathcal{Q}_k(g)$ with $i \neq j$. Pick one bottleneck task:

$$(i_k, j_k) \in \arg \min_{(i,j) \in \mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g) \quad \text{with } j_k \neq i_k.$$

Since $E_k \subseteq C_{k,i_k,j_k}$, we have

$$\mathbb{P}_\psi(E_k) \leqslant \tau^{\text{safe}}_{i_k j_k}(g) + \big(1 - \tau^{\text{safe}}_{i_k j_k}(g)\big) r_{i_k j_k},$$

Multiplying by V_k and summing over projects yields

$$\mathbb{E}_\psi[TS^{\text{coop}}] \leqslant TS(g) + \sum_{i \in \text{Arm}(g)} \sum_{j \neq i} \big(1 - \tau^{\text{safe}}_{ij}(g)\big)\, V^{\min}_{ij}\, r_{ij}, \tag{A.14}$$

where

$$V^{\min}_{ij} \equiv \sum_{k:\ (i_k,j_k)=(i,j)} V_k.$$

Upper bound on expropriation surplus. Fix a pair (i,j) with $i \in \text{Arm}(g)$ and $j \neq i$. On the event $\{\zeta_{ij} \neq 0\}$, obedience implies that attempting expropriation is (weakly) optimal, hence the conditional success probability is at least the private threshold:

$$\mathbb{P}_\psi(\zeta_{ij} = \theta_j \mid \zeta_{ij} \neq 0) \ \geqslant\ \tau^\star_{ij}.$$

Therefore the expected welfare contribution from this ordered pair satisfies

$$\mathbb{E}_\psi\Big[(b_{ij} - \lambda_{ij})\, \mathbf{1}\{\zeta_{ij} = \theta_j\} - c_{ij}\, \mathbf{1}\{\zeta_{ij} = -\theta_j\}\Big] \leqslant \mathbb{E}_\psi\Big[(b_{ij} - \lambda_{ij})\, \mathbf{1}\{\zeta_{ij} = \theta_j\}\Big] \leqslant (b_{ij} - \lambda_{ij})\, \tau^\star_{ij}\, r_{ij},$$

where we the first inequality follows from $-c_{ij} < 0$ and the second inequality follows from the fact the probability of successful expropriation should be at least $\tau^\star_{ij}$. Summing over (i,j) gives

$$\mathbb{E}_\psi[TS^{\text{exp}}] \leqslant \sum_{i \in \text{Arm}(g)} \sum_{j \neq i} \tau^\star_{ij}\, (b_{ij} - \lambda_{ij})\, r_{ij}. \tag{A.15}$$

Upper bound on total surplus. Combining (A.14) and (A.15) gives

$$\begin{aligned}
\mathbb{E}_\psi[TS(a,g;\theta)] &\leqslant TS(g) + \sum_{i \in \text{Arm}(g)} \sum_{j \neq i} r_{ij} \Big[(1 - \tau^\star_{ij}) V^{\min}_{ij} + \tau^\star_{ij}(b_{ij} - \lambda_{ij})\Big] = \\
&= TS(g) - \sum_{i \in \text{Arm}(g)} \sum_{j \neq i} r_{ij} \Big[\underbrace{\tau^\star_{ij}(\lambda_{ij} - b_{ij}) - \big(1 - \tau^\star_{ij}\big) V^{\min}_{ij}}_{>0} \Big].
\end{aligned}$$

By Assumption 1, $\lambda_{ij} - b_{ij} > V_{\text{total}} \geqslant V^{\min}_{ij}$. Also $b_{ij} < c_{ij}$ implies $\tau^\star_{ij} = \frac{c_{ij}}{b_{ij}+c_{ij}} > \frac{1}{2}$, hence $\tau^\star_{ij} > (1 - \tau^\star_{ij})$. Therefore, for every $i \in \text{Arm}(g)$,

$$\tau^\star_{ij}(\lambda_{ij} - b_{ij}) \ >\ \tau^\star_{ij} V^{\min}_{ij} \ >\ (1 - \tau^\star_{ij}) V^{\min}_{ij},$$

so the bracketed term is strictly positive. Therefore, the maximum attainable expected total surplus is $TS(g)$ and can be attained only when there is no expropriation, $r_{ij} = 0$ for all $i \in \mathcal{N}$. □

Conclusion. Lemma 2 shows $TS(g)$ is attainable by the nested no-expropriation rule. Lemma 3 shows no obedient rule can exceed it under Assumption 1. This completes the proof of Proposition 1.

A.3 The Solution to the Mechanism-Design Problem

In this section, we show that the contract described in (A.6)–(A.8) satisfies the incentive-compatibility constraints and therefore solves party 1's problem in (A.3). Intuitively, we prove the following chain of inequalities:

$$U_i(\text{truth, obey}) \geqslant U_i(\text{truth, any deviation } \delta_i) \geqslant U_i(\text{misreport, same deviation } \delta_i),$$

where the first inequality follows from obedience of the optimal nested decision rule (Lemma 2), and the second inequality is proved below.

Proof. Fix the deterministic owner g^* and let $\sigma^*(\cdot \mid \hat{\theta})$ denote the optimal nested decision rule used by the mediator after receiving reports $\hat{\theta} \in \Theta$. Specifically, the mediator draws a $\xi \sim U[0,1]$ and recommends:

$$x_{ij} = \bar{x}_{ij} = \begin{cases} \hat{\theta}_j & \text{if } \xi \leqslant \tau_{ij}^{\text{safe}}(g^*), \\ -\hat{\theta}_j & \text{if } \xi > \tau_{ij}^{\text{safe}}(g^*), \end{cases}, \qquad z_{ij} = \bar{z}_{ij} = 0$$

Fix $i \in \mathcal{N}$, a true type $\theta_i \in \Theta_i$, an alternative report $\theta_i' \in \Theta_i$, and an arbitrary deviation function $\delta_i : D_i \to \mathcal{A}_i$. Transfers are constant under the proposed contract, hence they do not affect incentives and are omitted throughout.

Obedience under truthful reports. Conditional on truthful reports, the mediator uses the nested rule $\sigma^*(\cdot \mid \theta)$. By Lemma 2 (obedience of the nested rule),

$$U_i(\varphi, t \mid \theta_i) \;\geqslant\; U_i^*(\varphi, t, \delta_i, \theta_i \mid \theta_i) \qquad \text{for all } \delta_i : D_i \to \mathcal{A}_i. \tag{A.16}$$

Own-message independence. Under the nested rule, for each $j \neq i$ the recommendation sent to i about matching θ_j depends on $(\hat{\theta}_j, \xi)$ but not on $\hat{\theta}_i$. Hence, holding $\hat{\theta}_{-i}$ fixed, the distribution of the private information d_i (and therefore of the realized action $\delta_i(d_i)$) is the same under $\hat{\theta}_i = \theta_i$ and under $\hat{\theta}_i = -\theta_i$.

Misreporting weakly reduces project success, holding δ_i fixed. Fix any project $k \in \mathcal{K}$. If no requirement in $\mathcal{Q}_k(g^*)$ involves matching party i's state, then i's report does not affect the probability of project k's success. Otherwise, there exists some $\ell \neq i$ with $(\ell, i) \in \mathcal{Q}_k(g^*)$, so success of project k requires that ℓ's matching action for i equals the *true* θ_i. Under $\hat{\theta}_i = \theta_i$, the nested rule recommends to ℓ the correct match with probability $\tau^{\text{safe}}_{\ell i}(g^*)$, while under $\hat{\theta}_i = -\theta_i$ it recommends the correct match with probability $1 - \tau^{\text{safe}}_{\ell i}(g^*)$. Since $\tau^{\text{safe}}_{\ell i}(g^*) > \frac{1}{2}$, we have

$$\mathbb{P}_\psi(E_k \mid \hat{\theta}_i = \theta_i, \delta_i) \geqslant \mathbb{P}_\psi(E_k \mid \hat{\theta}_i = -\theta_i, \delta_i),$$

where E_k is the probability of project k's success under the induced distribution $\psi(a, g^*, \theta) = p(\theta)\sigma^*(a \mid \theta)$ over outcomes and states. Because cooperative utility is $u_i^c(x, \bar{x} \mid g^*, \theta) = \sum_{k \in \mathcal{K}} \alpha_{ik} V_k \mathbf{1}\{E_k\}$, it follows that

$$U_i^*(\varphi, t, \delta_i, \hat{\theta}_i = \theta_i \mid \theta_i) \geqslant U_i^*(\varphi, t, \delta_i, \hat{\theta}_i = -\theta_i \mid \theta_i) \qquad \text{for all } \delta_i : D_i \to \mathcal{A}_i. \tag{A.17}$$

Combining (A.16) and (A.17) yields the double-deviation incentive-compatibility constraint. □

A.4 Contractible Actions and Cooperative Payoffs

In the baseline model, all components of the action plan $a_i = (x_i, \bar{x}_i, z_i, \bar{z}_i) \in \mathcal{A}_i$ are *non-contractible*: the mediator can only send private recommendations, and obedience must be verified against deviations in *all* components. In addition, project shares $\{\alpha_{ik}\}_{i,k}$ are fixed and exogenously given.

This section studies two counterfactual environments. First, we allow expropriation actions to be contractible; this makes it possible to implement the first best. Second, we allow cooperative actions and project shares to be contractible. We show that making cooperative actions and shares contractible does not increase the maximum attainable surplus relative to the baseline.

For notational convenience, decompose each party's action plan into a *cooperation block* and an *expropriation block*:

$$a_i^c \equiv (x_i, \bar{x}_i) \in \mathcal{A}_i^c = \{-1, 1\}^{2n} \qquad a_i^e \equiv (z_i, \bar{z}_i) \in \mathcal{A}_i^e = \{-1, 1\}^{2(n-1)},$$

so that $\mathcal{A}_i = \mathcal{A}_i^c \times \mathcal{A}_i^e$ and $a_i = (a_i^c, a_i^e)$. The mediator observes θ and draws an action profile $a = (a_i)_{i \in \mathcal{N}} \in \mathcal{A}$ from a direct decision rule $\sigma(\cdot \mid \theta) \in \Delta(\mathcal{A})$.

(i) Contractible expropriation actions. Suppose the mediator can *enforce* expropriation actions $(z, \bar{z})$, while cooperative actions $(x, \bar{x})$ remain non-contractible. Then, after receiving

a recommendation $a_i = (a_i^c, a_i^e)$, party i can deviate only in the cooperation block a_i^c; the expropriation block a_i^e is fixed.

Accordingly, σ is *obedient under g with contractible expropriation* if, for every $i \in \mathcal{N}$, every $\theta_i \in \Theta_i$, every recommended $a_i = (a_i^c, a_i^e) \in \mathcal{A}_i$, and every alternative $\tilde{a}_i^c \in \mathcal{A}_i^c$,

$$\sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(a_i, a_{-i} \mid \theta_i, \theta_{-i}) \Big[u_i\big(((a_i^c, a_i^e), a_{-i}), g; \theta\big) - u_i\big(((\tilde{a}_i^c, a_i^e), a_{-i}), g; \theta\big) \Big] \geqslant 0. \tag{A.18}$$

Define the corresponding omniscient benchmark value:

$$\begin{aligned} TS^e(g) \equiv \max_{\sigma} & \sum_{\theta} p(\theta) \sum_{a} \sigma(a \mid \theta) \cdot TS(a, g; \theta) \\ & \text{subject to (A.18)}, \quad \sigma(a \mid \theta) \geqslant 0, \quad \sum_{a} \sigma(a \mid \theta) = 1 \text{ for all } \theta \in \Theta \end{aligned} \tag{A.19}$$

(ii) Contractible actions and cooperative payoffs. Suppose instead that the mediator can *enforce* cooperative actions $(x, \bar{x})$, while expropriation actions $(z, \bar{z})$ remain non-contractible. Then party i can deviate only in the expropriation block a_i^e; the cooperation block a_i^c is fixed.

Accordingly, σ is *obedient under g with contractible cooperation* if, for every $i \in \mathcal{N}$, every $\theta_i \in \Theta_i$, every recommended $a_i = (a_i^c, a_i^e) \in \mathcal{A}_i$, and every alternative $\tilde{a}_i^e \in \mathcal{A}_i^e$,

$$\sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i} \mid \theta_i) \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(a_i, a_{-i} \mid \theta_i, \theta_{-i}) \Big[u_i\big(((a_i^c, a_i^e), a_{-i}), g; \theta\big) - u_i\big(((a_i^c, \tilde{a}_i^e), a_{-i}), g; \theta\big) \Big] \;\geqslant\; 0. \tag{A.20}$$

Suppose the mediator can also choose and enforce the division of cooperative payoffs $\{\alpha_{ik}\}_{i,k}$. Thus, the corresponding omniscient mediator's benchmark value is

$$\begin{aligned} TS^c(g) \equiv \max_{\sigma} & \sum_{\theta} p(\theta) \sum_{a} \sigma(a \mid \theta) \cdot TS(a, g; \theta) \\ & \text{s.t. (A.20)}, \quad \alpha_{ik} \geqslant 0, \quad \sum_{i \in \mathcal{N}} \alpha_{ik} = 1 \text{ for all } k \in \mathcal{K} \\ & \sigma(a \mid \theta) \geqslant 0, \quad \sum_{a} \sigma(a \mid \theta) = 1 \text{ for all } \theta \in \Theta \end{aligned} \tag{A.21}$$

Proposition 3 (Contractability and the attainable surplus)**.** *Fix any ownership assignment $g \in \mathcal{N}$. If expropriation actions $(z, \bar{z})$ are contractible, then*

$$TS^e(g) = V_{\text{total}} \qquad \textit{for every } g \in \mathcal{N}.$$

If cooperative actions $(x, \bar{x})$ and projects' shares $\{\alpha_{ik}\}_{i,k}$ are contractible, then

$$TS^c(g) = TS(g) \qquad \textit{for every } g \in \mathcal{N},$$

where $TS(g)$ is the value characterized in Proposition 1.

Proof. **(i) Contractible expropriation actions:** $TS^e(g) = V_{\text{total}}$. Define a (deterministic) decision rule σ^{FB} as follows. For every $\theta \in \Theta$:

1. Set *all* expropriation actions to zero: $z_{ij} = 0$ and $\bar{z}_{ij} = 0$ for all i and all $j \neq i$;
2. Recommend $x_{ij} = \bar{x}_{ij} = \theta_j$ for all $i \in \mathcal{N}$ and $j \in \mathcal{N}$.

Under σ^{FB}, there is no expropriation, and every project succeeds surely. Thus the total surplus is V_{total}. Since $(z, \bar{z})$ are enforced, party i can deviate only in $a_i^c = (x_i, \bar{x}_i)$. If i deviates to some $\tilde{a}_i^c \neq a_i^c$, then at least one matching requirement that uses i's action is violated, which can only weakly decrease i's cooperative payoff (since $\alpha_{ik} \geqslant 0$ for all k). Thus, σ^{FB} is obedient.

(ii) Contractible cooperation and shares: $TS^c(g) = TS(g)$. First, we show that $TS^c(g) \geqslant TS(g)$. Let σ^g be the nested no-expropriation decision rule constructed in the proof of Proposition 1 (see (A.10)-(A.11)). Under σ^g we have $\zeta_{ij} = 0$ for all $i \neq j$ almost surely, and each project k succeeds with probability $\min_{(i,j)\in\mathcal{Q}_k(g)} \tau_{ij}^{\text{safe}}(g)$, so it achieves $TS(g)$. This decision rule is obedient and satisfies (A.20).

Second, we show that $TS^c(g) \leqslant TS(g)$. This inequality follows from the same steps as in the proof of Lemma 3. The binding restriction comes from the expropriation obedience constraints. Since expropriation actions remain non-contractible, the same upper bound continues to apply even when cooperative actions are contractible.

Finally, allowing the mediator to choose the shares $\{\alpha_{ik}\}$ does not affect the objective W, which depends only on total project values. It also does not affect the obedience constraints (A.20): because cooperative actions are fixed in both terms, cooperative payoffs cancel when comparing deviations in expropriation. Hence, making shares contractible is immaterial for the welfare bound. □

A.5 Optimal Decision Rule with Expropriation

The goal of this section is to prove the following result.

Proposition 4. *Suppose $\lambda_{ij} - b_{ij} > 0$ and $b_{ij} < c_{ij}$. The maximum attainable total surplus*

under an owner $g \in \mathcal{N}$ is bounded by

$$TS(g) \leqslant \max_{\frac{1}{2} \leqslant q_{ij} \leqslant 1} \left\{ \sum_{k \in \mathcal{K}} V_k \cdot \min_{(i,j) \in \mathcal{Q}_k(g)} q_{ij} \; - \sum_{i \in \mathrm{Arm}(g)} \sum_{j \neq i} (\lambda_{ij} - b_{ij}) \max\left\{0, \; \frac{q_{ij} - \tau^\star_{ij}}{1 - \tau^\star_{ij}}\right\} \right\}. \quad \text{(A.22)}$$

We will prove this result in two lemmas. Fix an owner $g \in \mathcal{N}$. For each project k and each task $(i,j) \in \mathcal{Q}_k(g)$, let $C_{k,(i,j)}$ denote the event that the task (i,j) is satisfied (as defined in (A.13)), and let $E_k \equiv \bigcap_{(i,j) \in \mathcal{Q}_k(g)} C_{k,(i,j)}$ denote project success. Let $\psi(\theta, a) \equiv p(\theta)\sigma(a \mid \theta)$ be the joint distribution induced by an obedient decision rule $\sigma(\cdot \mid \theta)$.

Task-level accuracy. For each pair (i,j) define

$$q_{ij} \equiv \max\left\{ \mathbb{P}_\psi\left(x_{ij} = \theta_j\right), \mathbb{P}_\psi\left(\bar{x}_{ij} = \theta_j\right) \right\} \quad \text{(A.23)}$$

Expropriation probabilities. For each pair (i,j) with $j \neq i$, define expropriation attempt and success probabilities

$$r_{ij} \equiv \mathbb{P}_\psi(\zeta_{ij} \neq 0), \qquad s_{ij} \equiv \mathbb{P}_\psi(\zeta_{ij} = \theta_j).$$

By feasibility, $0 \leqslant s_{ij} \leqslant r_{ij} \leqslant 1$.

Lemma 4. *Consider any obedient decision rule $\sigma(\cdot \mid \theta)$. Fix an ordered pair (i,j) with $j \neq i$, $i \in \mathrm{Arm}(g)$, and $(i,j) \in \mathcal{Q}_k(g)$ for some $k \in \mathcal{K}$. Then*

$$s_{ij} \geqslant \max\left\{0, \frac{q_{ij} - \tau^\star_{ij}}{1 - \tau^\star_{ij}}\right\}$$

Proof. Let $I_i \equiv (\theta_i, a_i)$ denote i's information at the action stage, and define the *best-guess accuracy* about θ_j:

$$p_{ij}(I_i) \;\equiv\; \max\left\{\mathbb{P}_\psi(\theta_j = 1 \mid I_i), \; \mathbb{P}_\psi(\theta_j = -1 \mid I_i)\right\} \;\in\; [1/2, 1].$$

Step 1: We show that any matching task cannot be satisfied with conditional probability exceeding $p_{ij}(I_i)$. By construction,

$$\mathbb{P}_\psi(\theta_j = x_{ij} \mid I_i) \leqslant p_{ij}(I_i) \quad \text{and} \quad \mathbb{P}_\psi(\theta_j = \bar{x}_{ij} \mid I_i) \leqslant p_{ij}(I_i).$$

Taking the expectations yields

$$\mathbb{E}_\psi\left[\theta_j = x_{ij} \mid I_i\right] \leqslant \mathbb{E}_\psi\left[p_{ij}(I_i)\right] \quad \text{and} \quad \mathbb{E}_\psi\left[\theta_j = \bar{x}_{ij} \mid I_i\right] \leqslant \mathbb{E}_\psi\left[p_{ij}(I_i)\right]$$

Hence,

$$q_{ij} \leqslant \mathbb{E}_\psi[p_{ij}(I_i)]. \tag{A.24}$$

Step 2: Obedience pins down when expropriation is attempted, and ties success to $p_{ij}(I_i)$. Because $i \in \mathrm{Arm}(g)$, it can choose $\zeta_{ij} \in \{-1, 0, 1\}$. Given I_i, the private payoff from choosing sign $w \in \{-1, 1\}$ is

$$b_{ij}\,\mathbb{P}_\psi(\theta_j = w \mid I_i) - c_{ij}\,\mathbb{P}_\psi(\theta_j = -w \mid I_i) = (b_{ij} + c_{ij})\,\mathbb{P}_\psi(\theta_j = w \mid I_i) \;-\; c_{ij}.$$

Therefore the optimal expropriation sign is a most-likely sign, and attempting expropriation yields nonnegative private payoff if and only if $p_{ij}(I_i) \geqslant \tau^\star_{ij} \equiv \frac{c_{ij}}{b_{ij}+c_{ij}}$. With the tie-break favoring $\zeta_{ij} = 0$ at indifference, obedience implies:

$$\zeta_{ij} = 0 \quad\Longleftrightarrow\quad p_{ij}(I_i) \leqslant \tau^\star_{ij} \qquad \text{and} \qquad \zeta_{ij} \in \{-1, 1\} \quad\Longleftrightarrow\quad p_{ij}(I_i) > \tau^\star_{ij}.$$

Moreover, on $\{\zeta_{ij} \neq 0\}$ the recommended sign must be a most-likely sign, so conditional on I_i the success probability of the attempted expropriation equals $p_{ij}(I_i)$. Hence the success probability $s_{ij} \equiv \mathbb{P}_\psi(\zeta_{ij} = \theta_j)$ satisfies

$$s_{ij} = \mathbb{E}_\psi\Big[p_{ij}(I_i) \cdot \mathbf{1}\{p_{ij}(I_i) > \tau^\star_{ij}\}\Big]. \tag{A.25}$$

Step 3: We derive a lower bound on successful expropriation in terms of q_{ij}. Fix $\tau \in (1/2, 1)$ and note the pointwise inequality, valid for all $p \in [1/2, 1]$:

$$p - \tau \;\leqslant\; (1 - \tau)\, p\, \mathbf{1}\{p > \tau\}. \tag{A.26}$$

If $p \leqslant \tau$, the left-hand side is $\leqslant 0$ while the right-hand side is 0; if $p > \tau$, then $p - \tau \leqslant (1 - \tau)p$ because $p \leqslant 1$.

Apply (A.26) with $p = p_{ij}(I_i)$ and $\tau = \tau^\star_{ij}$ and take expectations:

$$\mathbb{E}_\psi[p_{ij}(I_i)] - \tau^\star_{ij} \;\leqslant\; (1 - \tau^\star_{ij})\,\mathbb{E}_\psi\Big[p_{ij}(I_i)\mathbf{1}\{p_{ij}(I_i) > \tau^\star_{ij}\}\Big] = (1 - \tau^\star_{ij})\, s_{ij},$$

where the last equality uses (A.25). Rearranging gives

$$s_{ij} \;\geqslant\; \frac{\mathbb{E}_\psi[p_{ij}(I_i)] - \tau^\star_{ij}}{1 - \tau^\star_{ij}}.$$

Combining with (A.24) yields

$$s_{ij} \;\geqslant\; \max\Big\{0,\; \frac{q_{ij} - \tau^\star_{ij}}{1 - \tau^\star_{ij}}\Big\}. \tag{A.27}$$

□

Lemma 5. *Fix g and any obedient decision rule σ with induced ψ. The upper bound on the expected total surplus is*

$$\mathbb{E}_\psi[TS] \leqslant \sum_{k\in\mathcal{K}} V_k \cdot \min_{(i,j)\in\mathcal{Q}_k(g)} q_{ij} \;-\; \sum_{i\in\mathrm{Arm}(g)} \sum_{j\neq i} (\lambda_{ij} - b_{ij}) \max\Big\{0,\ \frac{q_{ij}-\tau^\star_{ij}}{1-\tau^\star_{ij}}\Big\}, \tag{A.28}$$

where q_{ij} is defined in (A.23)*.*

Proof. Write $TS = TS^{\mathrm{coop}} + TS^{\mathrm{exp}}$ as in Lemma 3.

Cooperation. Recall that $E_k \equiv \bigcap_{(i,j)\in\mathcal{Q}_k(g)} C_{k,(i,j)}$ is the probability that project k succeeds. We have the following inequalities

$$\mathbb{E}_\psi[TS^{\mathrm{coop}}] = \sum_{k\in\mathcal{K}} V_k\, \mathbb{P}_\psi(E_k) \leqslant \sum_{k\in\mathcal{K}} V_k \cdot \min_{(i,j)\in\mathcal{Q}_k(g)} \mathbb{P}_\psi\left(C_{k,(i,j)}\right) \leqslant \sum_{k\in\mathcal{K}} V_k \cdot \min_{(i,j)\in\mathcal{Q}_k(g)} q_{ij}$$

where the first inequality follows from the proof in Lemma 3, and the second inequality follows from the definition of q_{ij}.

Expropriation. For each armed $i \in \mathrm{Arm}(g)$ and $j \neq i$, the welfare contribution from expropriation is

$$\mathbb{E}_\psi\left[TS^{\mathrm{exp}}_{ij}\right] \equiv \mathbb{E}_\psi\left[(b_{ij}-\lambda_{ij})\,\mathbf{1}\{\zeta_{ij}=\theta_j\} - c_{ij}\,\mathbf{1}\{\zeta_{ij}=-\theta_j\}\right] = (b_{ij}-\lambda_{ij})s_{ij} - c_{ij}(r_{ij}-s_{ij}).$$

Dropping the non-positive $-c_{ij}$ term yields

$$\mathbb{E}_\psi[W^{\mathrm{exp}}] \leqslant \sum_{i\in\mathrm{Arm}(g)} \sum_{j\neq i} (b_{ij}-\lambda_{ij}) s_{ij}$$

Since $b_{ij} - \lambda_{ij} < 0$, Lemma 4 implies

$$(b_{ij}-\lambda_{ij})\, s_{ij} \;\leqslant\; (b_{ij}-\lambda_{ij}) \max\Big\{0,\ \frac{q_{ij}-\tau^\star_{ij}}{1-\tau^\star_{ij}}\Big\} = -(\lambda_{ij}-b_{ij}) \max\Big\{0,\ \frac{q_{ij}-\tau^\star_{ij}}{1-\tau^\star_{ij}}\Big\}.$$

Summing over (i,j) and combining with the cooperative bound yields (A.28). □