

Control and Equity Tradeoff in a Partnership ^{*}

Sijia Zhang [†]

April 8, 2026

Abstract

This paper studies control reallocation in a partnership, where the partners hold asymmetric information about whose expertise should direct the firm. Using a mechanism design approach, it characterizes the first- and second-best mechanisms that incorporate equity transfers as part of the design. The analysis provides novel insights into how the interdependence between partners' expertise shapes the optimal equity transfer. It shows that both partial and full equity transfers may emerge as part of the optimal mechanism. Specifically, with a control shift, both the first- and second-best solutions may prescribe full equity transfer under substitute expertise, but minimal equity transfer under complementary expertise.

Keywords: Control and ownership, Mechanism design, Interdependent values, Partnership dissolution

JEL Codes: D21, D28, G32, G34

^{*}We are indebted to Silvana Krasteva and Alexander L. Brown for their constant support and invaluable guidance. We thank Yunus Aybas, Marco Castillo, Huiyi Guo, Allan Hernández-Chanto, Eric Maskin and Guoqiang Tian for helpful comments and suggestions. We have also benefited from comments by the participants of the Texas Theory Camp (2024). We gratefully acknowledge support from Mary and Charles H. Gregory '64 Hagler Institute for Advanced Study Student Fellowship.

[†]Texas A&M University; teresasijia@tamu.edu.

1 Introduction

When business conditions change, firms often need to restructure. Sometimes the change runs deeper, requiring not just new strategies but new leadership. Yet whether such a transition is truly warranted is rarely common knowledge: the initial controller, who operates the firm day to day, holds private information that outside partners cannot easily access or verify. This creates a fundamental problem: the partner who knows the most about the business is often the one least willing to relinquish control. This problem is particularly acute in startups, where founders have superior knowledge about their products compared to outside investors. For example, WeWork’s founder Adam Neumann retained control despite mounting governance problems, with investors intervening only after the failed 2019 IPO. Theranos illustrates an even more severe outcome: Elizabeth Holmes’s super-voting shares prevented timely intervention despite core technology failures, ultimately leading to dissolution with hundreds of millions in losses. As these examples illustrate, such misalignment of incentives could leave the firm locked in an inefficient arrangement long after a control transition would have been beneficial. Since restructuring generally comes with equity transfers, such transfers emerge as natural instruments for aligning incentives. This paper studies how equity transfers could be optimally designed so that control transitions happen exactly when they should.

To investigate this question, we consider a closely held firm with two partners: an initial controller I and a passive shareholder O , each with expertise in a distinct area. The firm’s value depends on which partner controls the firm, each partner’s type (representing their area’s potential), and the interdependence between the two areas, which captures whether they function as complements or substitutes. Efficiency requires allocating control to the partner whose expertise is currently more valuable. However, under asymmetric information, the identity of the efficient controller is not common knowledge, making it unclear whether a control transition should happen.

We use a mechanism design approach to study the control reallocation process in the presence of such asymmetric information. A key feature of our model is that equity transfers are part of the mechanism design: we allow ownership structure to vary with the identity of the assigned controller. Specifically, let s_i denote the shares held by partner $i \in \{I, O\}$ when assigned control, with the remaining shares $1 - s_i$ held by the other partner. When $s_i = 1$, the assigned controller i acquires full ownership, implying a complete buyout of the other partner. A value of $s_i < 1$ indicates that i retains only partial ownership, leaving the other partner a stake in the firm’s value.

Our analysis focuses primarily on the case where only I holds private information about her type, as is typical in venture capital settings where founders hold stronger private information than investors.¹ The direct mechanism must incentivize I ’s truthful reporting and induce voluntary participation by both partners. However, the information rents required

¹In Section 5, we consider a more general setup where both partners hold private information, while I ’s private information is still more pivotal than O ’s.

by I may outweigh the net welfare gains from reallocation, making efficiency unattainable without external subsidies. We characterize the first-best mechanism, which achieves efficient reallocation with minimal subsidy (if any), and the second-best mechanism, which maximizes *ex ante* welfare without incurring any subsidy.

The first-best mechanism minimizes the subsidy required to implement efficient control reallocation. Since the subsidy increases with I 's information rents, the first-best optimal shares (s_I, s_O) are those that minimize information rents. To see how shares affect information rents, note that the mechanism operates as a share trading process: the partner assigned control purchases shares from the other partner at a fixed price—the firm's value at I 's worst-off type. The worst-off type is the one whose net payoff from the mechanism is lowest relative to the status quo. Under the efficient reallocation rule, this type coincides with O 's type, where the identity of the controller is inconsequential for efficiency and I earns zero information rents.

When I 's type exceeds the worst-off type (i.e., O 's type), efficiency requires I to retain control, and s_I determines the number of shares I must buy. In this case, I values shares more than the worst-off type, so she purchases shares at a price below her valuation. This price gap accrues information rents to I , which increase with the number of shares purchased. Therefore, it is optimal to set s_I at the lowest feasible level (fewest shares purchased by I). When I 's type is lower than the worst-off type, efficiency requires transferring control to O , and s_O determines the number of shares O must buy. Whether these shares are under- or overvalued from I 's perspective depends on the interdependence between the partners' expertise. Under positive interdependence, the price exceeds I 's valuation, so the shares are overvalued. Thus, selling more shares at the fixed price generates additional rents for I , making it optimal to set s_O at the lowest feasible level (fewest shares sold by I). In contrast, under negative interdependence, shares are undervalued, making $s_O = 1$ optimal as it requires I to sell all shares at a below-valuation price.

When efficient control reallocation is unattainable without external subsidies, we characterize the second-best mechanism that maximizes *ex ante* welfare without incurring subsidies. We ask how this different objective shapes the optimal (s_I, s_O) . In contrast to the first-best, which focuses solely on minimizing information rents, the second-best mechanism balances two goals: reducing information rents and maintaining high *ex ante* welfare.

To achieve this, the second-best mechanism sacrifices efficient allocation for the types where the benefit from identifying the efficient controller is smallest, i.e., when the potential value from each partner's expertise is sufficiently close. The mechanism instead applies a specific lottery rule that leads these types to become the worst-off, thereby creating a zero information rent region. Consequently, the second-best mechanism incorporates two types of regions with distinct allocation rules: efficient allocation regions with positive information rents, as in the first-best, and a lottery region with zero information rents.

The highest achievable welfare is therefore affected by (s_I, s_O) through two channels. The "information rent channel" operates in the efficient allocation region, in the same way as in

the first-best. This channel favors (s_I, s_O) that minimize information rents, as lower rents allow the efficient region to expand without incurring subsidy. The “lottery loss channel” operates in the lottery region. Here, since the lottery probability must be a specific function of (s_I, s_O) to induce the zero information rent region, the choice of (s_I, s_O) directly affects the welfare loss relative to efficiency. This channel favors (s_I, s_O) that help to maintain high expected welfare in the inefficient region.

The balance of these two channels determines the optimal (s_I, s_O) . A key finding is that when a partner i is the preferred controller,² the “lottery loss channel” (the new channel arising in the second-best) prescribes the same equity structure as the “information rent channel” (the only channel in the first best). As a result, the optimal shares align across the first-best and second-best when the preferred controller retains control, so that the equity prescription from the first-best continues to hold. However, for the non-preferred controller j , the two channels conflict. Therefore, the optimal s_j depends on their relative strength, and may deviate substantially from the first-best prescription.

We further explore two extensions to provide broader insights. First, we allow O to hold private information with varying importance, measured by how pivotal it is in determining the efficient allocation. When O ’s private information has minimal importance, our first-best results regarding the optimal (s_I, s_O) continue to hold. However, at higher importance levels, new forces emerge. In addition to affecting the amount of shares traded, (s_I, s_O) now affect I ’s information rents by shifting the worst-off type as well. Moreover, their effect on O ’s information rents (which increase with the importance level) must also be accounted for. These additional forces can lead the optimal shares to deviate substantially from the one-sided case. Second, we generalize the second-best analysis to allow the interdependence to be captured by a non-linear function. Our results continue to hold when the function is concave, but may not hold when it is convex. This is because ensuring zero information rents for a range of types is infeasible with a convex function, making the effect of the information rent channel ambiguous.

These findings yield concrete guidance for the design of equity contracts for restructuring. Beyond conventional considerations such as incentive alignment or liquidity constraints, the optimal equity structure should also be shaped based on how partners’ skill sets interact (complements or substitutes). When the potential gains from control transition are compelling, this interdependence could drive the optimal equity structure in distinct directions. In particular, complementary expertise calls for continued co-ownership by the initial controller when relinquishing control, as retained shares reduce the informational frictions of the transition. By contrast, substitute expertise calls for a clean dissolution, as retained shares in this case would cause the firm to incur excessive compensation just to secure the transition. This prescription extends to the case where both partners hold private information, as long as the initial controller holds substantially stronger information advantages.

²We define the preferred controller formally in Section 4. Intuitively, such a controller exists when the average firm value under this partner’s control is always higher than under the other partner’s control in the lottery region.

When the gains from reallocation are more modest, achieving efficient control transition in all cases may no longer be feasible without external support. The optimal contract then deliberately keeps some ambiguity, to preserve the viability of transitions when they matter most. Specifically, for cases where the values of the two partners’ expertise are close, the optimal contract assigns control through a predetermined lottery, without resolving who the better controller truly is. When control transition is more likely to benefit the firm on average, the complement-substitute prescription continues to hold upon control transition. Otherwise, the prescription can differ, as the two channels pull in opposing directions: reducing information costs versus preserving the firm’s value under the lottery rule.

Related Literature Our model is closely related to the partnership dissolution literature. Extensive literature has studied partnership dissolution using mechanism design (Cramton et al. (1987), Fieseler et al. (2003), Jehiel and Paudner (2006), Ornelas and Turner (2007), Galavotti et al. (2011), and Loertscher and Wasser (2019)), considering different information structures (one-sided or two-sided), valuation setups (independent or interdependent), and outside options (one partner automatically retains control, or a lottery determines allocation by default). Fershtman et al. (2023) extend the classical dissolution model to allow remaining in partnership as one possible efficient outcome. We consider a different extension by enriching the compensation method. Specifically, in contrast to the classical models where departing partners must be bought out using lump-sum payments, we allow equity to be part of the compensation package (i.e., the partner not assigned control can retain some shares). We show that a richer compensation package is not guaranteed to be strictly beneficial. Instead, a clear dissolution (with lump-sum only payments) can still be optimal, especially when partners’ skill sets are substitutes.

Among the dissolution literature, the closest paper to ours is Jehiel and Paudner (2006), which features a similar setting but asks a fundamentally different question: how should the initial ownership shares be set to facilitate efficient dissolution? Instead, we take initial ownership as given and ask how post-reallocation shares should be designed in the restructured contract to facilitate efficient control transition. Though the two papers are methodologically similar, the different focus yields substantially different economic insights, and the results are not directly comparable. Specifically, changing the initial ownership share causes the traded shares to move mechanically together. The post-reallocation shares, by contrast, govern the relative attractiveness of each reallocation outcome independently. This separability introduces a richer interaction between equity structure, interdependence, and information rents, allowing us to disentangle when and why prescriptions differ sharply or align across the two reallocation outcomes. Additionally, in Jehiel and Paudner (2006), a lottery proportional to initial ownership shares serves as the status quo, while in our setup the default is simply the continuation of the current arrangement—the initial controller retains full control while the other partner remains passive. This status quo closely resembles real-life control transitions and shapes the optimal contract design in a non-trivial way: inducing zero information rents for some types necessarily comes at a welfare cost. Moreover, this status quo naturally makes

efficient reallocation harder to achieve, and such difficulty allows us to clearly tease out how different restructured shares can facilitate or hinder efficient control transition.

In our model, partners’ payoffs remain tied to the firm value through retained shares even when not receiving control. Our model thus contributes to the literature on how contingent payments (i.e., payments tied to *ex post* values directly) affect informational frictions in an allocation process. Seminal works of Hansen (1985) and DeMarzo et al. (2005) show that using securities (such as shares, debt, and options) as payment can reduce information rents in the allocation of a project or firm.³ In bilateral trading setups, Yoon (2020) and Pan and Wang (2021) show that allowing the seller or a third party to retain partial ownership can overcome Myerson and Satterthwaite (1983)’s impossibility result. A key distinction of our model is that the identities of seller and buyer (i.e., which partner receives control) arise endogenously, rather than being fixed as in these papers. This endogeneity creates a two-way contingent payment structure, where whether payments in each direction facilitate or hinder efficient transition depends critically on the nature of expertise interdependence between partners.⁴

The paper closest to ours is Ferreira et al. (2015), where both partners hold equally important private information and firm value depends solely on the controller’s expertise. While that paper finds that the departing controller optimally retains the maximum available shares, our results show that the optimal prescription depends critically on how the partners’ expertise interact. Under complementary expertise, retaining maximum shares upon relinquishing control is optimal, which is consistent with Ferreira et al. (2015). Under substitute expertise, however, the opposite holds: a clean dissolution with a complete buyout is optimal. This divergence arises because our setup allows firm value to be shaped by both partners’ expertise and considers an information structure where the initial controller holds stronger private information, features particularly relevant in venture capital settings. In the extreme case, when only the initial controller holds private information, the trading price of shares is uniquely determined and invariant to private information, driving the optimal equity structure to corner solutions rather than the interior ones found in their model when the initial controller retains control. When the other partner starts to hold private information, corner solution continues to hold as long as the asymmetry in private information remains sufficiently strong. Moreover, which corner is optimal depends on the nature of expertise interdependence—a force absent in their framework that can reverse the prescription entirely. In addition to the first-best, we explore the second-best mechanism when ensuring efficient restructuring is too costly. We establish that the complement-substitute prescription continues to hold upon control transition, when such transition is on average more beneficial

³See Skrzypacz (2013) for an overview. Subsequent work considers different security types and their effects on efficiency or rent extraction (e.g., Hernandez-Chanto and Fioriti (2019); Liu and Bernhardt (2021); Fioriti and Hernandez-Chanto (2022)).

⁴A separate strand of literature studies CEO replacement and executive compensation, which also explores tradeoffs between paying with lump-sum payments and stock options (see, e.g., Edmans et al. (2017)). In contrast to the principal-agent framework commonly used in that literature, our framework focuses on asset allocation under asymmetric information.

for marginal cases where the values of the two partners' expertise are sufficiently close.

Ekmekci et al. (2016) study the problem of selling a firm's control rights to a buyer where the seller can choose how many shares to sell and at what price. They characterize the conditions under which the optimal mechanism must be either a take-it-or-leave-it offer to sell the minimum fraction of shares necessary for control transition or a take-it-or-leave-it offer to sell all shares. Which of the two is optimal depends on what information the seller wants to screen. While we obtain a similar binary result in our main setup (the assigned controller either acquires all or minimum shares from the other partner), our results are driven by different underlying forces: two partners' expertise relationship and the mechanism's objective (the first-best or second-best).

The remainder of the paper is organized as follows. We set up the model and preliminaries in Section 2. In Section 3, we characterize when efficient reallocation is attainable without subsidy, and, in cases where it is not, identify the minimum subsidy required to restore the first-best. In Section 4, we characterize the second-best mechanism that maximizes welfare without incurring subsidy. In addition to these characterizations, both Sections 3 and 4 focus on identifying the optimal restructured shares (s_I, s_O) . In the discussion (Section 5), we generalize the main setup and examine when previous results hold and when deviations can occur. Finally, we conclude the paper and discuss possible future directions (Section 6).

2 Model and preliminaries

Consider a firm jointly owned by two risk-neutral partners: an insider I , who holds a fraction $r \in [0, 1]$ of ownership (shares), and an outsider O , who holds the remaining shares $1 - r$. Initially, I has full control over the firm's operations, while O is a passive shareholder not involved in any decision-making process. Due to some exogenous factors (e.g., changes in the business environment), the firm needs to reconsider whether partner I remains the optimal controller, or control should be shifted entirely to O . The two partners engage in negotiations to determine who should control the firm.

The two partners have different skill sets. Let θ_i denote partner i 's type ($i \in \{I, O\}$), which captures the firm's potential in the business area aligned with i 's expertise. As a specific example, in a venture capital setting, θ_I captures the value the firm can generate through the founder's technical expertise and product knowledge, while θ_O captures the value the firm can generate through the investor's expertise in scaling the business, accessing broader networks, and implementing financial restructuring. The firm's value under partner i 's control is given by $v_i(\theta_i, \theta_j)$, where the arguments are ordered with the controller's index first.⁵

⁵As noted by Ferreira et al. (2015), we can understand the actual value under i 's control as $v_i(\theta_i, \theta_j) + \epsilon$, where ϵ follows a distribution with zero mean, is independent of private information, and has unbounded support. With the existence of ϵ , I 's private information is not *ex post* verifiable: efficiency cannot be achieved by imposing harsh punishments for misreporting based on the partners' observed payoffs after reallocation (considered in papers including Mezzetti (2004), Mezzetti (2007), Mylovanov and Zapechelnuyk (2017), Kunimoto and Zhang (2026)). Since only $v_i(\theta_i, \theta_j)$ matters at the time of reallocation, we use

We allow for a general functional form for $v_i(\theta_i, \theta_j)$ with the following intuitive properties.

The firm's value under i 's control $v_i(\theta_i, \theta_j)$

- is continuously differentiable and strictly increasing in θ_i ;
- satisfies the single-crossing property $\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_i} > \frac{\partial v_j(\theta_j, \theta_i)}{\partial \theta_i}$;
- is higher than $v_j(\theta_j, \theta_i)$ if and only if $\theta_i > \theta_j$.

The first condition ensures that the controller's own expertise contributes positively to the firm's value. The second condition, the single-crossing property, guarantees that the firm's value is the same under each partner's control (i.e., the identity of the controller is inconsequential) for at most one value of θ_I . The third condition is not necessary for the analysis. With only the first two conditions, there exists at most one threshold θ_I^T such that $v_I(\theta_I, \theta_O) < v_O(\theta_O, \theta_I)$ when $\theta_I < \theta_I^T$ and $v_I(\theta_I, \theta_O) > v_O(\theta_O, \theta_I)$ when $\theta_I > \theta_I^T$. The third condition simply sets $\theta_I^T = \theta_O$ for convenience.

Another important feature of $v_i(\theta_i, \theta_j)$ is that it depends not only on the controller's type θ_i but also on the other partner's type θ_j , capturing possible interdependence between the partners' areas of expertise. These two areas could, in principle, function as complements ($\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_j} > 0$) or substitutes ($\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_j} < 0$). Continuing the venture capital example, if the founder's deep product knowledge and the investor's broad market access reinforce each other, the two areas are complements. If instead the firm is pivoting toward a direction that makes the founder's original expertise compete with the investor's new strategic vision, the two areas are substitutes.⁶

When engaged in the control reallocation negotiation, the two partners hold asymmetric information relevant to the firm's value. This makes it challenging to ensure control is reallocated efficiently. Formally, we assume that each partner $i \in \{I, O\}$ holds information about their respective expertise-aligned business areas (θ_i). It is commonly known that θ_i is independently distributed according to a continuous and differentiable cumulative distribution function $F_i(\theta_i)$. We start by analyzing the case of one-sided private information. Specifically, we assume that at the time of negotiation, the realization θ_O is common knowledge, while θ_I is privately observed by I . This information structure is particularly natural in venture capital settings: while the founder's knowledge of the firm's technical feasibility, product roadmap, and operational challenges tends to be inherently private, the investor's capabilities are largely observable through public signals such as portfolio performance, reputation, and past investments. It is therefore reasonable to treat the outsider's type as common knowledge while maintaining that the insider's type is private. In practice, the investor may also hold some private information, though typically less pivotal than the founder's. We

$v_i(\theta_i, \theta_j)$ for analysis throughout the paper.

⁶We can also think of additional examples. Positive interdependence: advancements in AI enhance the value of traditional software products, as seen with Microsoft's integration of AI into Office applications. Negative interdependence: digital expansion at firms like The New York Times may divert resources and readership from its traditional print business.

later extend the analysis to cover this more general case in Section 5, providing insights on when and to what extent this additional private information affects the optimal results.

2.1 Mechanism for control reallocation

The timeline for the model is as follows. I , as the initial controller, learns her realization θ_I privately while θ_O is common knowledge. Negotiation occurs between the two partners to determine: (1) the reallocation outcome, (2) the monetary transfers that occur between the two partners, and (3) the new ownership structure. After a partner i is assigned control, the firm operates under i 's control and generates value $v_i(\theta_i, \theta_j)$.

Reallocation Mechanism: Following the partnership dissolution literature and the control contest model of Ferreira et al. (2015), we cast the negotiation as a mechanism designed to maximize the *ex ante* welfare of I and O (i.e., the firm's value) subject to incentive compatibility, individual rationality, and budget balance constraints. The mechanism design approach allows us to directly characterize the optimal feasible contract without committing to a specific bargaining procedure, thereby providing guidance on contract design that applies across different bargaining settings.

Invoking the Revelation Principle, without loss of generality, we restrict attention to a direct mechanism where partner I 's message is simply her type θ_I . Formally, given I 's reported type $\tilde{\theta}_I$, the mechanism $\langle c, t, s_I, s_O \rangle$ consists of:

- (1) A control reallocation rule $c(\tilde{\theta}_I)$, specifying the probability that I retains control;
- (2) Ownership shares of partner i upon receiving control, s_i , with the remaining shares $1 - s_i$ held by the other partner;
- (3) A lump-sum monetary transfer rule $t(\tilde{\theta}_I)$, specifying the monetary transfer from O to I , which can be positive or negative.

We impose that the assigned controller i must hold at least $\underline{s}$ shares, i.e., $s_i \geq \underline{s}$. The parameter $\underline{s}$ represents the exogenously imposed minimum ownership required to maintain controlling power. When $\underline{s} = 1$ and thus $s_I = s_O = 1$, the mechanism reduces to the partnership dissolution framework, in which the controlling partner must acquire all shares from the other partner and retain full ownership (i.e., a complete buyout). In our mechanism, however, we allow the partner assigned control to acquire only partial shares from the other partner, i.e., $\underline{s} < 1$. This is a natural extension, as letting the non-controlling partner retain some shares is commonly observed and utilized in closely-held firms in practice.⁷ The threshold $\underline{s}$ should also be strictly greater than 0. First, it reflects legal restrictions

⁷One commonly-observed example is the vesting schedule in venture capital contracts, where shares are gradually distributed to the founders. If a founder steps down from management, they typically retain the shares already vested.

imposed on firms' governance structures. Second, agency problems become severe when the controller's ownership is too small, substantially reducing firm value.⁸ A strictly positive $\underline{s}$ can be understood as a convenient way to abstract from this problem by assuming the controller exerts full effort once holding at least $\underline{s}$ shares.⁹

Payoff of each partner: The payoff function of each partner takes the form $sv + t$, where s denotes the shares held by the partner, v denotes the firm's valuation, and t denotes the received transfers. Given a reallocation outcome $c(\tilde{\theta}_I) \in \{0, 1\}$, I 's and O 's *ex post* payoffs in the mechanism are as follows:

$$\begin{aligned} U_I(\theta_I, \tilde{\theta}_I, s_I, s_O) &= c(\tilde{\theta}_I)s_I v_I(\theta_I, \theta_O) + (1 - c(\tilde{\theta}_I))(1 - s_O)v_O(\theta_O, \theta_I) + t(\tilde{\theta}_I) \\ U_O(\theta_I, \tilde{\theta}_I, s_I, s_O) &= (1 - c(\tilde{\theta}_I))s_O v_O(\theta_O, \theta_I) + c(\tilde{\theta}_I)(1 - s_I)v_I(\theta_I, \theta_O) - t(\tilde{\theta}_I) \end{aligned}$$

The reallocation outcome affects the partners' *ex post* utility directly through the ownership benefits. For example, when $c(\tilde{\theta}_I) = 1$ (I receives control), partner I retains shares s_I and thus earns $s_I v_I(\theta_I, \theta_O)$ from them. Conversely, when $c(\tilde{\theta}_I) = 0$ (O receives control), partner I keeps the residual shares $(1 - s_O)$ and her retained shares yield $(1 - s_O)v_O(\theta_O, \theta_I)$.

Objective: The mechanism maximizes the *ex ante* welfare of I and O (i.e., the firm's value), subject to I 's incentive compatibility, both partners' individual rationality, and *ex post* budget balance (*BB*) constraints. The incentive compatibility and individual rationality constraints apply to the interim stage, after I learns her private type but before the mechanism is executed. Since monetary transfers only occur between the two partners in the mechanism, *BB* holds automatically. The following constraints are therefore required:

Partner I prefers to truthfully report her private type θ_I :

$$U_I(\theta_I, \theta_I, s_I, s_O) \geq U_I(\theta_I, \tilde{\theta}_I, s_I, s_O) \quad \forall \theta_I, \tilde{\theta}_I \quad (ICI)$$

Partner I agrees to participate in the mechanism for any type:

$$U_I(\theta_I, \theta_I, s_I, s_O) \geq r v_I(\theta_I, \theta_O) \quad \forall \theta_I \quad (IRI)$$

Partner O agrees to participate in the mechanism:

$$\mathbb{E}_{\theta_I}[U_O(\theta_I, \theta_I, s_I, s_O) - (1 - r)v_I(\theta_I, \theta_O)] \geq 0 \quad (IRO)$$

⁸Extensive literature has focused on this kind of agency problem in firms. See the seminal paper by Jensen and Meckling (1976).

⁹We do not impose $r \geq \underline{s}$. That is, I 's initial shares may no longer be sufficient to maintain controlling power after reallocation. When firms are reallocating control to adapt to changing environments, it is natural to assume $\underline{s}$ captures a different standard for the firm after reallocation (and a change in $\underline{s}$ itself can be one of the reasons the firm needs to consider reallocating control).

The optimal mechanism maximizes the expected *ex ante* welfare:

$$EW = \int_0^1 [c(\theta_I)v_I(\theta_I, \theta_O) + (1 - c(\theta_I))v_O(\theta_O, \theta_I)]dF_I(\theta_I)$$

subject to *ICI*, *IRI*, *IRO*, and *BB* constraints.

Model interpretation: Following Ferreira et al. (2015), we do not allow (s_I, s_O) to depend flexibly on I 's private information. Instead, the restructured ownership only depends on the identity of the assigned controller. This choice is made for both empirical and theoretical reasons. In practice, partners typically pre-specify ownership adjustments upon reallocation of control, while leaving more flexibility for lump-sum payments.¹⁰ From a technical standpoint, allowing restructured shares to vary freely with private information significantly complicates the analysis. In particular, since equity transfers between partners can occur in both directions, it becomes difficult to meaningfully disentangle control allocation (c) from restructured ownership (s).

Although we frame the model as a control reallocation process, it can equivalently be interpreted as a partnership dissolution procedure. Under this interpretation, the mechanism determines which partner retains the firm's control while the other exits. Our model extends the partnership dissolution framework by allowing the departing partner to receive compensation through both monetary payments and retained shares, rather than monetary payments alone. All results and implications derived in this paper apply equally to this partnership dissolution interpretation.

2.2 Preliminaries

Efficiency (the first-best outcome) requires allocating control to the partner with the higher θ_i . As in Ferreira et al. (2015), efficiency can sometimes become trivially feasible. This occurs when the initial shares of each partner already satisfy the minimum threshold for managerial control, i.e., when $\min\{r, 1 - r\} \geq \underline{s}$. In this case, the mechanism only needs to allocate control to the partner with the higher type without any equity or monetary transfers, and all constraints are automatically satisfied. Clearly, I has no incentive to misreport θ_I (*ICI*): her payoff monotonically increases with the firm's value, making efficient control reallocation aligned with her interests. Moreover, both *IRI* and *IRO* are trivially satisfied because each partner's net payoff from the mechanism must equal or exceed their outside option.

However, the condition $\min\{r, 1 - r\} \geq \underline{s}$ often fails in practice. For instance, some firms require a majority shareholding as the controlling stake. In this case, the condition never

¹⁰For example, in commonly-used Valuation Adjustment Mechanism (VAM) contracts, equity adjustments are generally predetermined when control shifts due to unmet performance targets. This raises another interpretation of our modeling: the exploration of the optimal (s_I, s_O) can be understood as identifying which pre-specification of shares best facilitates efficient control transition.

holds.¹¹ When the condition fails to hold (i.e., $\underline{s}$ is binding), efficient reallocation necessitates monetary and equity transfers. In many cases, a binding $\underline{s}$ can render efficiency unattainable, even under an optimally designed mechanism. This paper restricts attention to these more interesting cases.

We derive the following lemma based on I 's incentive compatibility condition (ICI). It generalizes Lemma 1 of Jehiel and Pauzner (2006), which considers the special case of $s_I = s_O = 1$.

Lemma 1 *For any (s_I, s_O) with $s_I \geq r$ and $s_O \geq 1 - r$ ¹²:*

- (i) $c(\theta_I)$ is weakly increasing in θ_I .
- (ii) For any θ_I , $\lim_{\varepsilon \rightarrow 0} t(\theta_I + \varepsilon) - t(\theta_I) = \lim_{\varepsilon \rightarrow 0} (c(\theta_I + \varepsilon) - c(\theta_I))[(1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]$.
- (iii) If c is differentiable at θ_I , then $t'(\theta_I) \equiv c'(\theta_I)[(1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]$.
- (iv) There is a constant m such that for all θ_I , $t(\theta_I) = \int_0^{\theta_I} c(x)[s_I \frac{\partial v_I(x, \theta_O)}{\partial x} - (1 - s_O) \frac{\partial v_O(\theta_O, x)}{\partial x}] dx + c(\theta_I)[(1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)] + m$.

Lemma 1(i) states that the probability of I retaining control should be increasing in her type, θ_I , to induce truthful reporting. Lemma 1(ii) provides a necessary condition for incentive compatibility (ICI), determining how monetary transfers must adjust with changes in control allocation outcomes. Specifically, the term $[(1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]$ is the difference in I 's ownership payoff under O 's control versus her own control. Monetary transfers must change proportionally to this term whenever control allocation probability changes. A direct implication of Lemma 1(ii) is that the transfer is identical for the same allocation outcome, i.e., $t(\theta_I) = t(\theta'_I)$ whenever $c(\theta_I) = c(\theta'_I)$. Lemma 1(iii) restates Lemma 1(ii) assuming that c is differentiable. Lemma 1(iv) characterizes the general form of transfers, which together with Lemma 1(i) provides necessary and sufficient conditions for ICI .

As mentioned above, we focus on situations in which efficiency cannot be achieved through control reallocation alone (i.e., $\min\{r, 1 - r\} < \underline{s}$). We also impose the assumptions required by Lemma 1 for tractability. Formally, we define the lower bounds for feasible restructured shares as $\underline{s}_I = \max\{r, \underline{s}\}$ and $\underline{s}_O = \max\{1 - r, \underline{s}\}$. The feasible set of restructured ownership is then given by

$$S^F = \{(s_I, s_O) \mid s_I \in [\underline{s}_I, 1], s_O \in [\underline{s}_O, 1]\}.$$

¹¹Additional sources of a high $\underline{s}$ may include: (1) Legal restrictions in jurisdictions like Singapore and Hong Kong that strictly limit unbundling control from shares. (2) Agency concerns in early-stage startups (especially in technology and biotech) requiring high ownership to maintain incentives.

¹²This assumption ($s_I \geq r$ and $s_O \geq 1 - r$) implies that the assigned controller cannot divest shares to the other partner. We impose it for tractability. Efficiency does not become trivially attainable by relaxing this assumption.

Throughout the remainder of the paper, we restrict attention to shares (s_I, s_O) in this feasible set S^F .

3 (Im)possibility of the first-best outcome

In this section, we identify the condition under which efficient reallocation is possible. When this condition fails, we characterize the minimum subsidy required to restore this first-best outcome.

As previously established, efficiency requires that managerial control be allocated to the partner with the higher type. That is:¹³

$$c^E(\theta_I) = \begin{cases} 0 & \text{for } \theta_I < \theta_O \\ 1 & \text{for } \theta_I > \theta_O \end{cases}. \quad (1)$$

Definition 1. Efficient reallocation is *possible* if there exists a mechanism $\langle c^E, t, s_I, s_O \rangle$ that satisfies *ICI*, *IRI*, *IRO*, and *BB*.

To understand how restructured shares (s_I, s_O) affect the possibility of efficient reallocation, we begin by characterizing the possibility condition given fixed (s_I, s_O) , then identifying the optimal shares (s_I^E, s_O^E) that minimize the required subsidy.

3.1 First-best mechanism under fixed shares

To satisfy *ICI*, a mechanism must fulfill the conditions in Lemma 1. The efficient reallocation rule $c^E(\theta_I)$ is increasing in θ_I , and thus satisfies Lemma 1(i). By Lemma 1(ii), the lump-sum transfer t must remain the same for any given allocation outcome $c^E(\theta_I) \in \{0, 1\}$. Therefore, to identify the minimum transfer I must receive to satisfy both *IRI* and *ICI*, it suffices to focus on the type for which *IRI* binds most tightly under each allocation outcome $c^E(\theta_I) \in \{0, 1\}$.

When $\theta_I < \theta_O$, efficiency requires reallocating control to partner O ($c^E = 0$). In this case, I must sell $(r - (1 - s_O))$ shares in exchange for a lump-sum transfer. Then, *IRI* can be rewritten as $t(\theta_I) \geq rv_I(\theta_I, \theta_O) - (1 - s_O)v_O(\theta_O, \theta_I)$. The right-hand side is increasing in θ_I since $\frac{\partial v_I(\theta_I, \theta_O)}{\partial \theta_I} > \frac{\partial v_O(\theta_O, \theta_I)}{\partial \theta_I}$ and $s_O \geq 1 - r$. Therefore, *IRI* is hardest to satisfy for the type $\theta_I = \theta_O$ in this region, and the minimum transfer is $t_{min}^E = (r - 1 + s_O)v_I(\theta_O, \theta_O)$.

Conversely, when $\theta_I > \theta_O$, efficiency requires that I retains control ($c^E = 1$). In this case, I must acquire $(s_I - r)$ shares from O and *IRI* reduces to $t(\theta_I) \geq rv_I(\theta_I, \theta_O) - s_I v_I(\theta_I, \theta_O)$. The right-hand side is decreasing in θ_I because $s_I \geq r$. Thus, *IRI* is again hardest to satisfy when $\theta_I = \theta_O$ and the minimum transfer is $t_{min}^E = (r - s_I)v_I(\theta_O, \theta_O)$.

¹³A tie occurs with zero probability and is therefore ignored.

As described above, *IRI* is hardest to satisfy at the type $\theta_I = \theta_O$ across all types, we therefore refer to this type as the “worst-off type”. Combining both cases, we have the minimum lump-sum transfer:

$$t_{min}^E(\theta_I, s_I, s_O) = \begin{cases} (r - 1 + s_O)v_I(\theta_O, \theta_O) & \text{for } \theta_I < \theta_O \\ (r - s_I)v_I(\theta_O, \theta_O) & \text{for } \theta_I > \theta_O \end{cases} \quad (2)$$

It can be verified that $t_{min}^E(\theta_I, s_I, s_O)$ together with efficient allocation c^E satisfies the *ICI* condition.¹⁴ Since *BB* is satisfied automatically in the mechanism, the remaining condition for efficient reallocation to be possible is *IRO*. Substituting c^E and t_{min}^E into O ’s expected payoff, we can represent O ’s expected loss from participating in the mechanism $\langle c^E, t_{min}^E, s_I, s_O \rangle$ as follows:

$$\begin{aligned} EL(s_I, s_O) = & \int_0^{\theta_O} s_O(v_O(\theta_O, \theta_O) - v_O(\theta_O, x)) - (1 - r)(v_I(\theta_O, \theta_O) - v_I(x, \theta_O))dF_I(x) \\ & + \int_{\theta_O}^1 (s_I - r)(v_I(x, \theta_O) - v_I(\theta_O, \theta_O))dF_I(x) \end{aligned} \quad (3)$$

Note that *IRO* reduces to $EL(s_I, s_O) \leq 0$. Since this inequality incorporates the minimum transfer that ensures *ICI* and *IRI*, it represents a necessary and sufficient condition for the possibility of efficient reallocation. Whenever this condition fails, $EL(s_I, s_O)$ represents the minimum subsidy required to restore the first-best outcome. Hence, for any fixed (s_I, s_O) , $EL(s_I, s_O)$ fully characterizes the possibility of implementing the first-best outcome. We denote the mechanism $\langle c^E, t_{min}^E, s_I, s_O \rangle$ as the first-best mechanism under (s_I, s_O) , as it implements the first-best outcome whenever possible.

Two observations follow from the analysis above. First, equation (3) reveals that s_i affects possibility only through its impact on O ’s payoff in states where partner i is assigned control. Specifically, s_I matters only when I retains control ($\theta_I > \theta_O$), while s_O matters only when O is allocated control ($\theta_I < \theta_O$). Second, equation (2) reveals that the minimum transfer t_{min}^E is solely determined by I ’s worst-off type ($\theta_I = \theta_O$) and the number of shares traded. Here, $v_I(\theta_O, \theta_O)$ can be interpreted as the per-unit price at which shares are traded.

3.2 Optimal shares in the first-best mechanism

Among all feasible ownership structures $(s_I, s_O) \in S^F$, we seek the shares that minimize the subsidy required for the first-best outcome: (s_I^E, s_O^E) .

¹⁴Minimizing transfers for each allocation outcome independently under *IRI* does not automatically ensure that the allocation jump at point $\theta_I = \theta_O$ satisfies the condition in Lemma 1(ii), hence additional verification is needed.

Definition 2. $(s_I^E, s_O^E) \equiv \arg \min_{(s_I, s_O) \in S^F} EL(s_I, s_O).$

The following proposition characterizes the optimal shares (s_I^E, s_O^E) .

Proposition 1 *The optimal shares (s_I^E, s_O^E) are uniquely determined by the interdependence between θ_I and θ_O .*

- (i) *Under positive interdependence ($\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_j} > 0$), it is optimal for the assigned controller to retain minimum shares upon receiving control: $(s_I^E, s_O^E) = (\underline{s}_I, \underline{s}_O)$.*
- (ii) *Under negative interdependence ($\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_j} < 0$), it is optimal for I to hold minimum shares when she retains control, while O should acquire full ownership when receiving control: $(s_I^E, s_O^E) = (\underline{s}_I, 1)$.*

The intuition is as follows. Minimizing the subsidy required to restore efficiency (if any) is equivalent to minimizing I 's *ex ante* information rents under the efficient allocation rule. Thus, the optimal shares (s_I^E, s_O^E) minimize I 's *ex ante* information rents. To understand how (s_I, s_O) affect information rents, note that the operation of the mechanism can be interpreted as follows: the partner assigned control needs to buy a certain amount of shares from the other partner, with the amount determined by (s_I, s_O) . Information rents are affected by (s_I, s_O) through the number of shares traded.¹⁵

When $\theta_I > \theta_O$, I retains control and must purchase $(s_I - r)$ shares from O at a per-unit price equal to the valuation at the worst-off type, $v_I(\theta_O, \theta_O)$. Since the shares are undervalued from I 's perspective ($v_I(\theta_I, \theta_O) > v_I(\theta_O, \theta_O)$), I 's interim information rents increase proportionally with the number of shares she buys. Therefore, $s_I^E = \underline{s}_I$ is optimal, as it minimizes the number of shares traded.

When $\theta_I < \theta_O$, O is assigned control and must purchase $s_O - (1 - r)$ shares from I , paying $(s_O - (1 - r))v_I(\theta_O, \theta_O) = (s_O - (1 - r))v_O(\theta_O, \theta_O)$. In this case, whether these traded shares are under- or overvalued from I 's perspective depends on the interdependence relationship. Under positive interdependence, shares are overvalued for I ($v_O(\theta_O, \theta_O) > v_O(\theta_O, \theta_I)$), and I earns additional rents by selling more shares. Thus, $s_O^E = \underline{s}_O$ is optimal. Under negative interdependence, the opposite holds: shares are undervalued. A complete buyout of I ($s_O^E = 1$) is thus optimal, forcing I to sell all shares at the below-valuation price and thereby minimizing her interim information rents.

These results provide practical insights for firms seeking to ensure efficient control allocation. To facilitate smooth reallocation and minimize I 's reluctance to relinquish control, contracts should: (1) rely on monetary transfers rather than substantial share grants when I retains control, as granting additional shares increases reallocation costs; (2) when the two partners' skill sets are complementary (positive interdependence), grant O limited ownership

¹⁵Shares (s_I, s_O) affect interim information rents through the number of shares traded, which in turn affects *ex ante* information rents.

when shifting control, since allowing I to retain shares makes her accept reallocation more readily; and (3) when the two skill sets are substitute (negative interdependence), implement a complete buyout of I when shifting control, as allowing I to retain shares would incur excessive compensation relative to lump-sum payment.

4 Second-best mechanism

When efficient reallocation is not possible, it is natural to consider the second-best mechanism that maximizes the *ex ante* welfare while still satisfying *ICI*, *IRI*, *IRO*, and *BB*. Analogous to the previous section, we first identify the second-best mechanism (and its corresponding maximum achievable welfare) under fixed restructured shares (s_I, s_O) . Then, we identify the optimal $(s_I, s_O) \in S^F$ that achieve the highest welfare.

In the first-best analysis, the allocation is fixed to be efficient, and the optimal shares (s_I, s_O) are identified by minimizing information rents alone. By contrast, the second-best mechanism must allow some degree of inefficiency to satisfy all the constraints. As a result, the optimal shares in the second-best entail a trade-off between reducing information rents and maintaining high *ex ante* welfare. As we discuss below, this trade-off can give rise to substantially different optimal shares compared to the first-best.

Throughout this section, we restrict attention to separable value functions of the form $v_i(\theta_i, \theta_j) = \theta_i + \eta(\theta_j)$ where $\eta(\cdot)$ is linear (i.e., $\eta'(\cdot)$ is constant). This serves as a natural benchmark and provides transparent insights. The Online Appendix provides a generalization that considers nonlinear value functions. For most of this section, we assume positive interdependence (i.e., $\eta'(\cdot) > 0$) since efficiency is generally harder to achieve in this case. In Subsection 4.2.3, we characterize the outcome under negative interdependence.

4.1 Second-best mechanism under fixed shares

As a starting point for our analysis, we define the following probability threshold:

$$TH(s_I, s_O) = \frac{r - (1 - s_O)\eta'(\cdot)}{s_I - (1 - s_O)\eta'(\cdot)}.$$

This threshold plays a crucial role in identifying the worst-off types, as the lemma below shows.

Lemma 2 *For any mechanism $\langle c, t, s_I, s_O \rangle$ satisfying *ICI*, if type θ_I satisfies $c(\theta_I) = TH(s_I, s_O)$, then θ_I is a worst-off type.*

To understand the above result, note that substituting $t(\theta_I)$ from Lemma 1(iv) into *IRI* yields:

$$\int_0^{\theta_I} c(x)[s_I - (1 - s_O)\eta'(x)]dx + (1 - s_O)v_O(\theta_O, \theta_I) - rv_I(\theta_I, \theta_O) + m \geq 0 \quad (4)$$

The left-hand side represents I 's net payoff from any mechanism that satisfies ICI . Minimizing this expression with respect to θ_I yields $c(\theta_I) = TH(s_I, s_O)$. Thus, any type θ_I satisfying $c(\theta_I) = TH(s_I, s_O)$ must be a worst-off type. Setting $m = r\eta(\theta_O) - (1 - s_O)\theta_O$ makes IRI binding for all such worst-off types, thereby making them earn zero information rents. This means that the mechanism can induce a region of types earning zero information rents by uniformly assigning control to I with probability $TH(s_I, s_O)$, accompanied by the appropriate fixed monetary transfers.

The next proposition shows that the characterization of the second-best mechanism can be reduced to an optimization problem with a single constraint (referred to as the “implementability constraint”) that depends on $TH(s_I, s_O)$ in an intuitive way.

Proposition 2 *The second-best allocation rule $c(\theta_I)$ is weakly increasing in θ_I and solves the following program:*

$$EW(s_I, s_O) = \max_{c(\cdot)} \int_{x=0}^1 [c(x)v_I(x, \theta_O) + (1 - c(x))v_O(\theta_O, x)]dF_I(x)$$

subject to

$$\begin{aligned} & \int_0^{\theta_I^*} (c(x) - TH(s_I, s_O))[\Delta_{NW}(x, \theta_O) + (s_I - (1 - s_O)\eta'(\cdot))\frac{F_I(x)}{f_I(x)}]dF_I(x) \\ & + \int_{\theta_I^{**}}^1 (c(x) - TH(s_I, s_O))[\Delta_{NW}(x, \theta_O) - (s_I - (1 - s_O)\eta'(\cdot))\frac{1 - F_I(x)}{f_I(x)}]dF_I(x) \quad (5) \\ & - \int_0^1 (1 - TH(s_I, s_O))\Delta_{NW}(x, \theta_O)dF_I(x) \geq 0 \end{aligned}$$

where:¹⁶

$$\theta_I^* = \sup\{\theta_I : c(\theta_I) < TH(s_I, s_O)\}$$

$$\theta_I^{**} = \inf\{\theta_I : c(\theta_I) > TH(s_I, s_O)\}$$

$$\Delta_{NW}(\theta_I, \theta_O) = v_I(\theta_I, \theta_O) - v_O(\theta_O, \theta_I) = \theta_I + \eta(\theta_O) - \theta_O - \eta(\theta_I).$$

As shown by Lemma 1(i), the allocation rule $c(\cdot)$ must be weakly increasing to satisfy ICI . In addition, combining IRI and IRO with the transfer rule t in Lemma 1(iv) yields the implementability constraint given by (5). Together, $c(\cdot)$ being weakly increasing and the implementability constraint holding are necessary and sufficient for the existence of a mechanism implementing $c(\cdot)$ that satisfies ICI , IRI , IRO , and BB .

Besides playing a central role in identifying the worst-off type, $TH(s_I, s_O)$ also acts as a threshold that determines whether I is a net seller or buyer relative to the “virtual status

¹⁶If $c(\theta_I) \geq TH(s_I, s_O)$ for all θ_I , let $\theta_I^* = 0$. If $c(\theta_I) \leq TH(s_I, s_O)$ for all θ_I , let $\theta_I^{**} = 1$.

quo.” To define the virtual status quo, note that in the partnership dissolution framework of Jehiel and Paudner (2006), setting $c(\theta_I) = r$ directly reproduces I ’s status quo, naturally generating a zero information rent region at no welfare cost. However, the mechanism cannot directly replicate such a status quo in our setup, for two reasons. First, our status quo has I always in control. The allocation rule that induces I ’s indifference with the status quo has to assign control to O with positive probability, meaning that achieving a zero-rent region must accompany a welfare change. Second, the restructured ownership must take the form $(s_I, 1 - s_I)$ or $(1 - s_O, s_O)$, leading to ownership benefits that differ from those in status quo. The lottery probability must therefore be adjusted to account for such changed ownership structure. As established in Lemma 2, we can use an allocation probability $c(\cdot) = TH(s_I, s_O)$ with appropriate monetary compensation to induce I to become the worst-off and earn zero rents, as if she were in the status quo, for any types. We therefore refer to allocating control to I with this probability as the virtual status quo within the mechanism.

We can now provide an intuitive explanation of the left-hand side of the implementability constraint (5). $\theta_I < \theta_I^*$ can be interpreted as the value realizations for which I is a net seller compared to the virtual status quo (since $c(\theta_I) \leq c(\theta_I^*) < TH(s_I, s_O)$). Then, $c(x) - TH(s_I, s_O)$ denotes I ’s allocation probability relative to the virtual status quo. The first term in brackets $(\Delta_{NW}(\theta_I, \theta_O))$ captures the welfare difference between reallocating control to I and to O , while the second term involving $\frac{F_I(x)}{f_I(x)}$ reflects I ’s information rents as a net seller. Analogously, $\theta_I > \theta_I^{**}$ corresponds to the realizations for which I is a net buyer, with the corresponding welfare change and information rents as a buyer. Finally, the last term $-\int_0^1 (1 - TH(s_I, s_O)) \Delta_{NW}(x, \theta_O) dF_I(x)$ adjusts for the difference in expected welfare between adopting the virtual status quo and the true status quo.

To glean further insights into the structure of the second-best mechanism, we impose the following two assumptions:

Assumption 1. $\frac{F_I(\theta_I)}{f_I(\theta_I)}$ is increasing in θ_I and $\frac{1-F_I(\theta_I)}{f_I(\theta_I)}$ is decreasing in θ_I .

Assumption 2. $\int_0^1 v_O(\theta_O, x) dF_I(x) \geq \int_0^1 v_I(x, \theta_O) dF_I(x)$.

Assumption 1 is a standard regularity condition in implementation theory that guarantees the monotonicity of the virtual valuations. Assumption 2 requires that the average firm value under O ’s control exceeds that under I ’s control, ruling out the trivial case where *ex ante* welfare is maximized by simply maintaining the status quo (i.e., I always retains control).

The following proposition characterizes the second-best mechanism under (s_I, s_O) :

Proposition 3 *When Assumptions 1-2 hold, the second-best mechanism under (s_I, s_O) follows the control reallocation rule:*

$$c^{SB}(\theta_I, s_I, s_O) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^* \\ TH(s_I, s_O) & \text{for } \theta_I^* \leq \theta_I \leq \theta_I^{**} \\ 1 & \text{for } \theta_I > \theta_I^{**} \end{cases} \quad (6)$$

where $(\theta_I^*, \theta_I^{**})$ is the smallest interval¹⁷ s.t. $0 < \theta_I^* < \theta_O$ and $1 > \theta_I^{**} > \theta_O$ satisfy

$$\begin{aligned} & \int_0^{\theta_I^*} \left[s_O v_O(\theta_O, x) - (1-r)v_I(x, \theta_O) - r v_I(\theta_I^*, \theta_O) + (1-s_O)v_O(\theta_O, \theta_I^*) \right] dF_I(x) \\ & + \int_{\theta_I^*}^{\theta_I^{**}} (1 - TH(s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x) \\ & + \int_{\theta_I^{**}}^1 (r - s_I)(v_I(x, \theta_O) - v_I(\theta_I^{**}, \theta_O)) dF_I(x) = 0. \end{aligned} \quad (7)$$

and

$$(v_O(\theta_O, \theta_I^*) - v_I(\theta_I^*, \theta_O)) \frac{f_I(\theta_I^*)}{F_I(\theta_I^*)} = (v_I(\theta_I^{**}, \theta_O) - v_O(\theta_O, \theta_I^{**})) \frac{f_I(\theta_I^{**})}{1 - F_I(\theta_I^{**})}. \quad (8)$$

It is evident from (6) that the second-best mechanism sacrifices efficiency in the interval $\theta_I \in (\theta_I^*, \theta_I^{**})$, replacing the efficient allocation with a lottery $c(\cdot) = TH(s_I, s_O)$. Intuitively, the mechanism sacrifices efficiency precisely where the benefit of identifying the optimal controller is smallest: when the two partners' abilities are sufficiently close to each other. By instead applying the lottery $c(\cdot) = TH(s_I, s_O)$ in the region, all types within it earn zero information rents, effectively reducing the overall cost of implementing the mechanism. Condition (7) ensures that information rents are reduced just enough to restore implementability. Condition (8) balances the two ways of reducing information rents—via increasing θ_I^{**} or decreasing θ_I^* . At the margin, the cost of misallocation from either approach (corrected by the rate at which the implementability constraint (5) is relaxed) should be the same.

4.2 Optimal shares in the second-best mechanism

Having characterized the second-best mechanism under fixed restructured shares (s_I, s_O) , we now turn to identifying the optimal (s_I^{SB}, s_O^{SB}) that correspond to the highest *ex ante* expected welfare.

From Proposition 3, the second-best mechanism is pinned down by two thresholds, θ_I^* and θ_I^{**} . Efficiency is guaranteed outside the interval $[\theta_I^*, \theta_I^{**}]$, while the lottery $c(\theta_I) = TH(s_I, s_O)$ is applied inside. Through conditions (7) and (8), each pair of shares (s_I, s_O) uniquely identifies the corresponding thresholds $(\theta_I^*(s_I, s_O)$ and $\theta_I^{**}(s_I, s_O))$ and thus the second-best mechanism with shares fixed.

Given the structure of the second-best mechanism, we formally define (s_I^{SB}, s_O^{SB}) as follows:

¹⁷When θ_I^* approaches 0, the corresponding θ_I^{**} satisfying (8) must approach 1. Given Assumption 2, there must exist values θ_I^* and θ_I^{**} satisfying (7) and (8). Taking the smallest such interval ensures its uniqueness.

Definition 3. $(s_I^{SB}, s_O^{SB}) \equiv \arg \max_{(s_I, s_O) \in S^F} EW(s_I, s_O)$, where

$$\begin{aligned} EW(s_I, s_O) &= \int_0^{\theta_I^*(s_I, s_O)} v_O(\theta_O, \theta_I) dF_I(\theta_I) + \int_{\theta_I^{**}(s_I, s_O)}^1 v_I(\theta_I, \theta_O) dF_I(\theta_I) \\ &+ \int_{\theta_I^*(s_I, s_O)}^{\theta_I^{**}(s_I, s_O)} \left[TH(s_I, s_O) v_I(\theta_I, \theta_O) + (1 - TH(s_I, s_O)) v_O(\theta_O, \theta_I) \right] dF_I(\theta_I) \end{aligned}$$

We first consider an extreme case in which efficiency always prescribes assigning control to O . This case provides clear-cut trade-offs from adjusting restructured shares (s_I, s_O) and allows for a closed-form characterization of (s_I^{SB}, s_O^{SB}) . Thus, it serves as a useful benchmark before turning to the general case.

4.2.1 Optimal shares in the extreme case

Consider an extreme case where O is always the efficient controller (i.e., $\theta_O > 1$). Note that even when O is commonly known to be the efficient controller, the mechanism still needs to compensate I to ensure her participation (IRI). Efficient reallocation thus can still be impossible in this extreme case: namely, when I 's required compensation outweighs the net welfare gains from the efficient reallocation. The second-best mechanism is then required and takes the following form (by Proposition 3).

Corollary 1 *If $\theta_O > 1$ and Assumptions 1-2 hold, then for a given (s_I, s_O) , the second-best mechanism follows the control reallocation rule:*

$$c^{SB}(\theta_I, s_I, s_O) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^* \\ TH(s_I, s_O) & \text{for } \theta_I \geq \theta_I^* \end{cases} \quad (9)$$

where $\theta_I^* = \sup\{\theta_I : 0 < \theta_I < 1\}$ is uniquely determined by:

$$\begin{aligned} &\int_0^{\theta_I^*} \left[s_O v_O(\theta_O, x) - (1 - r) v_I(x, \theta_O) - r v_I(\theta_I^*, \theta_O) + (1 - s_O) v_O(\theta_O, \theta_I^*) \right] dF_I(x) \\ &+ \int_{\theta_I^*}^1 (1 - TH(s_I, s_O)) (v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x) = 0. \end{aligned} \quad (10)$$

In this extreme case, the second-best allocation sacrifices efficiency in favor of a lottery for sufficiently high values of θ_I (i.e., those closest to θ_O). The following proposition characterizes the optimal second-best restructured shares.

Proposition 4 *If Assumptions 1-2 hold and $\theta_O > 1$, then under positive interdependence ($\eta'(\cdot) > 0$), $(s_I^{SB}, s_O^{SB}) = (1, \underline{s}_O)$.*

We sketch the key steps of the proof, starting with s_O^{SB} . We denote the second-best mechanism under fixed shares (s_I, s_O) as $\langle c^{SB}(\theta_I, s_I, s_O), t, s_I, s_O \rangle$, and construct an alternative mechanism $\langle c(\theta_I, s_I, \underline{s}_O), t', s_I, \underline{s}_O \rangle$ based on it, with the following allocation rule:¹⁸

$$c(\theta_I, s_I, \underline{s}_O) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^*(s_I, s_O) \\ TH(s_I, \underline{s}_O) & \text{for } \theta_I \geq \theta_I^*(s_I, s_O) \end{cases}. \quad (11)$$

The allocation rule above adopts the threshold $\theta_I^*(s_I, s_O)$ from the second-best allocation under (s_I, s_O) , thereby maintaining the same division of types into the efficient region and the lottery region. It then adjusts the lottery probability to the level that generates the worst-off types under $(s_I, \underline{s}_O)$: $c(\cdot) = TH(s_I, \underline{s}_O)$. As a result, the implementability constraint for the constructed rule is:

$$\begin{aligned} & \int_0^{\theta_I^*(s_I, s_O)} \left[\underline{s}_O v_O(\theta_O, x) - (1-r)v_I(x, \theta_O) \right. \\ & \quad \left. - r v_I(\theta_I^*(s_I, s_O), \theta_O) + (1-\underline{s}_O)v_O(\theta_O, \theta_I^*(s_I, s_O)) \right] dF_I(x) \\ & \quad + \int_{\theta_I^*(s_I, s_O)}^1 (1-TH(s_I, \underline{s}_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x) \geq 0. \end{aligned} \quad (12)$$

Note that the left-hand side of inequality (12) collapses to that of equation (10) when $\underline{s}_O$ is replaced with s_O . Since replacing $\underline{s}_O$ with $s_O > \underline{s}_O$ decreases the left-hand side of (12), we can conclude that the constructed mechanism $\langle c(\theta_I, s_I, \underline{s}_O), t', s_I, \underline{s}_O \rangle$ has a more relaxed implementability constraint relative to $\langle c^{SB}(\theta_I, s_I, s_O), t, s_I, s_O \rangle$. Since the implementability constraint for the latter mechanism (the second-best mechanism under (s_I, s_O)) binds by definition, the constructed mechanism must be implementable.

This constructed mechanism is not only implementable, but also delivers strictly higher expected welfare: it allocates control to the efficient controller O with higher probability in the lottery region. Since this constructed mechanism satisfies the implementability constraint with slack, the second-best mechanism under $(s_I, \underline{s}_O)$ must be able to achieve even higher welfare by expanding the efficient region. Therefore, for any feasible $s_O \neq \underline{s}_O$, the second-best mechanism under $(s_I, \underline{s}_O)$ must dominate that under (s_I, s_O) .

A parallel construction establishes $s_I^{SB} = 1$. Thus, under this extreme case, the optimal shares are $s_O^{SB} = \underline{s}_O$ and $s_I^{SB} = 1$. That is, the efficient controller O retains the minimum feasible ownership whenever assigned control, while I acquires all shares from O whenever she retains control. Notably, the optimal share $s_O^{SB} = \underline{s}_O$ coincides with that in the first-best, i.e., $s_O^{SB} = s_O^E$. We now explain the intuition behind this result.

¹⁸ t and t' are uniquely determined by the other elements of the mechanism, and we omit their expressions here as they are not directly relevant to the proof.

According to the structure of the second-best mechanism, each type is either in the efficient reallocation region with positive information rents (if $\theta_I < \theta_I^*(s_I, s_O)$), or in the lottery region with zero information rents (if $\theta_I \geq \theta_I^*(s_I, s_O)$). The restructured shares (s_I, s_O) thus intuitively affect the maximum achievable welfare through two distinct channels.

The “information rent channel” operates in the efficient region, in the same way as in the first-best: restructured shares (s_I, s_O) affect the information rents via the number of shares traded. This channel favors (s_I, s_O) that induce the lowest interim information rents, since lower information rents allow the efficient region to be expanded while preserving the mechanism’s implementability (i.e., a higher $\theta_I^*(s_I, s_O)$). The “lottery loss channel” operates in the lottery region and affects the achievable welfare by determining the expected welfare loss relative to the efficient outcome. Specifically, it operates through affecting lottery probability $TH(s_I, s_O)$. This channel favors (s_I, s_O) that minimize the expected welfare loss.

The balance of these two channels determines the optimal (s_I, s_O) . Consider s_O first. The information rent channel favors $\underline{s}_O$ (for the same reason as in the first-best). The lottery loss channel also favors $\underline{s}_O$ because it corresponds to $TH(s_I, s_O)$ assigning control to the efficient controller O with the highest probability. Therefore, $\underline{s}_O$ must be optimal as it is endorsed by both channels. This also implies that the prescription for optimal s_O is consistent across the first-best and the second-best analysis, because the new channel in the second-best (lottery loss channel) operates in the same direction as the only channel in the first-best (information rent channel). As for s_I , the information rent channel is irrelevant (since efficiency never requires I to be the controller), so only the lottery loss channel matters. Therefore, $s_I = 1$ is optimal because it maximizes the likelihood of assigning control to the efficient controller O in the lottery region.

4.2.2 Optimal shares in the general case

In the extreme case above, efficiency always prescribes allocating control to O . Consequently, the lottery loss channel clearly prefers (s_I, s_O) that are associated with the lottery allocating control to O with the highest probability. In the general case where both partners can possibly be the efficient controller, how the lottery loss channel operates becomes ambiguous.

However, the insights from the extreme case extend to the general case when one partner emerges as the “preferred controller.” Partner i is the preferred controller if, in the lottery region of any second-best mechanism under $(s_I, s_O) \in S^F$, the average firm value under i ’s control exceeds that under j ’s control. The existence of a preferred controller resolves the ambiguity in how the lottery loss channel operates. Specifically, the lottery loss channel prefers shares (s_I, s_O) corresponding to the lottery allocating control to the preferred controller with the highest probability.

We define the expression $PP(s_I, s_O)$, which captures the expected difference between the

firm's value under I 's and O 's control in the second-best lottery region:

$$PP(s_I, s_O) = \int_{\theta_I^*(s_I, s_O)}^{\theta_I^{**}(s_I, s_O)} v_I(x, \theta_O) - v_O(\theta_O, x) dF_I(x).$$

By definition, partner I is the preferred controller if $PP(s_I, s_O) > 0$ for all $(s_I, s_O) \in S^F$, while partner O is the preferred controller if $PP(s_I, s_O) < 0$ for all $(s_I, s_O) \in S^F$. When one partner i is the preferred controller, the insights from the extreme case extend naturally.

Proposition 5 *Suppose Assumptions 1-2 hold. Under positive interdependence ($\eta'(\cdot) > 0$):*

- (i) *If O is the preferred controller, i.e., $PP(s_I, s_O) < 0 \quad \forall (s_I, s_O) \in S^F$, then $s_O^{SB} = \underline{s}_O$.*
- (ii) *If I is the preferred controller, i.e., $PP(s_I, s_O) > 0 \quad \forall (s_I, s_O) \in S^F$, then $s_I^{SB} = \underline{s}_I$.*

Proposition 5 states that under positive interdependence, whenever the preferred controller i retains control, it is optimal to restrict this partner's restructured shares to the lowest level $\underline{s}_i$. The proof closely mirrors the construction described in the extreme case, so we leave it to the Appendix. The intuition also parallels the extreme case: when i is the preferred controller (i.e., on average it is more efficient to allocate control to i), both the information rent channel (the only channel in the first-best analysis) and the lottery loss channel (the new channel in the second-best) favor this level, making $\underline{s}_i$ optimal in both the first-best and the second-best mechanisms.

One might notice that Proposition 5 is silent on the optimal share of the partner who is not the preferred controller. This is because the two channels conflict over this partner's optimal shares. For the non-preferred controller j , the information rent channel favors $s_j = \underline{s}_j$ since it minimizes information rents in the efficient region, while the lottery loss channel favors $s_j = 1$ since it minimizes expected welfare loss in the lottery region. The optimal level of s_j then depends on the relative strength of these two channels, which cannot be determined without additional information.

We now present an example to illustrate that the optimal restructured shares for the non-preferred controller can vary substantially depending on the specific setup, including the common knowledge about O 's type.

Example 1 *Suppose the value of the firm is $v_i(\theta_i, \theta_j) = \theta_i + 0.7\theta_j$, I 's initial share is $r = 0.7$, and I 's private information θ_I follows a uniform distribution $F_I(x) \sim U[0, 1]$. We assume that after control reallocation, the assigned controller must hold at least 0.8 ownership of the firm, i.e., $S^F = \{(s_I, s_O) \mid s_I \in [0.8, 1], s_O \in [0.8, 1]\}$. We consider two distinct cases: when $\theta_O = 0.55$ and when $\theta_O = 0.8$.*

We can establish numerically that in both cases, O is the preferred controller, i.e., $PP(s_I, s_O) < 0$ for $\forall (s_I, s_O) \in S^F$. Thus, based on Proposition 5(i), $s_O^{SB} = \underline{s}_O = 0.8$.

However, s_I^{SB} is indeterminate, as it depends on the relative strength of the information rent channel and the lottery loss channel.

Figure 1 depicts how the achievable welfare responds to changes in s_I for the two distinct cases, with s_O fixed at its optimal level. As shown in the figure, when $\theta_O = 0.55$, $s_I^{SB} = \underline{s_I} = 0.8$; whereas when $\theta_O = 0.8$, $s_I^{SB} = 1$. The intuition for why complete dissolution becomes optimal when θ_O is higher is as follows: the lottery region generated by the second-best mechanism under any (s_I, s_O) must always include $\theta_I = \theta_O$, that is, we must have $\theta_I^* \leq \theta_O \leq \theta_I^{**}$. The information rent channel for s_I operates only through affecting the information rents of types above θ_I^* (where it is efficient for I to control). In this example, a higher θ_O leads to a higher θ_I^* , reducing the range of types where the information rent channel operates for s_I . As a result, as θ_O increases from 0.55 to 0.8, the lottery loss channel changes from being dominated by the information rent channel to dominating it, leading to $s_I^{SB} = 1$.

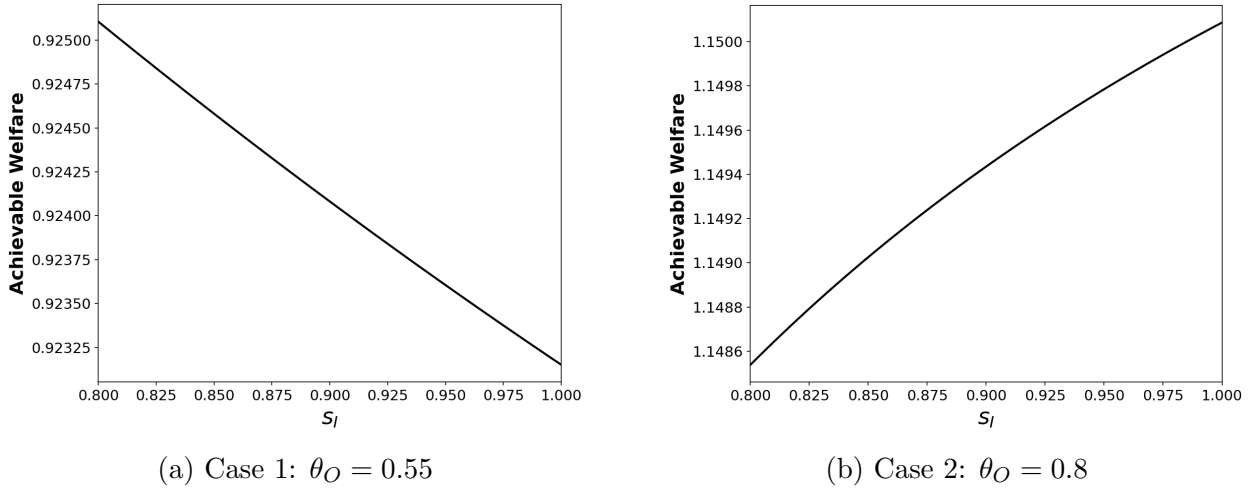


Figure 1: Achievable Welfare vs s_I
($s_O = \underline{s_O} = 0.8$)

The example illustrates the key intuition: the optimal shares for the non-preferred controller tend to depend on the size of the lottery region. When the lottery region is small (i.e., the second-best mechanism does not require sacrificing much efficiency for all (s_I, s_O)), the first-best and second-best prescriptions for the non-preferred controller tend to be aligned, as the primary focus for both is minimizing information rents under the efficient allocation. However, when the lottery region is large enough, maintaining high welfare in the inefficient region becomes more important. This leads to a divergence between the first-best and second-best prescriptions. In particular, upon receiving control, the second-best may call for the non-preferred controller to acquire all shares, while the first-best prescribes retaining minimum shares. Factors such as the value of θ_O and the distribution of I 's private information all affect the relative strength of the two channels, thereby determining whether or not the two prescriptions align.

4.2.3 Optimal shares under negative interdependence

Under negative interdependence ($\eta'(\cdot) < 0$), the characterization of the second-best mechanism in Subsection 4.1 still holds. While the optimal shares (s_I^{SB}, s_O^{SB}) change with the interdependence, similar intuition regarding a preferred controller applies, as shown below.

Proposition 6 *Suppose Assumptions 1-2 hold. Under negative interdependence $\eta'(\cdot) < 0$:*

- (i) *If O is the preferred controller, i.e., $PP(s_I, s_O) < 0 \quad \forall (s_I, s_O) \in S^F$, then $s_O^{SB} = 1$.*
- (ii) *If I is the preferred controller, i.e., $PP(s_I, s_O) > 0 \quad \forall (s_I, s_O) \in S^F$, then $s_I^{SB} = \underline{s}_I$.*

The proof follows a similar construction as in the positive interdependence case, so we omit it here and focus on the intuition. The results regarding s_I remain unchanged, since the interdependence relationship does not affect how each channel operates for s_I . However, s_O affects the two channels in the opposite direction: (1) in the information rent channel, $s_O = 1$ minimizes information rents under negative interdependence, consistent with the first-best under negative interdependence in Section 3; (2) in the lottery loss channel, $1 - TH(s_I, s_O)$ increases with s_O , so $s_O = 1$ corresponds to the lottery assigning control to O with the highest probability. As a result, when O is the preferred controller, $s_O^{SB} = 1$ since this level is optimal for both channels. When I is the preferred controller, the optimal second-best s_O^{SB} depends on the relative strength of the two channels. Interestingly, when O is the preferred controller and the lottery loss channel dominates for s_I , it can even be optimal to have $s_I^{SB} = s_O^{SB} = 1$. In such a case, it is always optimal for dissolution (a complete buyout) to occur regardless of who retains control, and our second-best mechanism would coincide with the second-best dissolution mechanism with the same outside option.

Summarizing across results under both positive and negative interdependence, our analysis provides practical guidance: When a preferred controller exists, their optimal retained shares upon receiving control align with the first-best prescription. The non-preferred controller's optimal shares upon receiving control, however, depend on the trade-off between efficiency and its cost. When achieving efficiency requires substantial subsidies and becomes too costly (i.e., when the lottery region is large), I as the non-preferred controller should buy out O completely upon receiving control. For O as the non-preferred controller, the prescription to avoid substantial subsidies depends on interdependence: under positive interdependence, O should acquire all shares; under negative interdependence, O should retain only the minimum shares required for control.

5 Discussion

In this section, we explore the robustness of our results when we deviate from the baseline setup. First, we consider a more general setup in which both I and O hold private information, with the importance of O 's private information varying across different levels. We show

that when O 's private information is relatively unimportant, the results in Section 3 remain robust. However, as its importance increases, new forces come into play, leading to substantial deviations from our baseline findings. Second, we generalize the second-best analysis to allow for a nonlinear interdependence relationship, where the relationship is captured by function $\eta(\cdot)$. We show that our second-best results still hold when $\eta(\cdot)$ is concave but may deviate when $\eta(\cdot)$ is convex.

5.1 Two-sided private information

In practice, both partners may hold relevant private information. For example, in startups, while founders have superior knowledge about their products' technical feasibility, investors often have better insights into market opportunities and scaling strategies that founders may not be aware of. We extend the model to incorporate this possibility. Specifically, we now assume O 's type is given by $\theta_O = \epsilon\omega + (1 - \epsilon)a$.¹⁹ The type θ_O consists of two components: a commonly-known term a and a privately-known random variable ω , which is distributed according to $F_O(\cdot)$ on $[0, 1]$. The parameter ϵ determines the weight of each component.

When $\epsilon = 0$, we have $\theta_O = a$, implying O 's private information is irrelevant to efficient reallocation. The setup then reduces to the baseline one-sided private information case. When $\epsilon = 1$, we have $\theta_O = \omega$, which corresponds to a case where O 's private information is as pivotal as I 's in implementing efficient reallocation. For intermediate values $\epsilon \in (0, 1)$, O 's type lies in the interval $[\underline{\theta}_O, \overline{\theta}_O] = [(1 - \epsilon)a, \epsilon + (1 - \epsilon)a]$. In this case, O 's private information is relevant only when I 's type also falls within this range.

To characterize the (im)possibility condition for efficient reallocation under this general setup, we consider a direct mechanism which operates in the following way:

1. Each partner $i \in \{I, O\}$ pays a participation fee k_i .
2. Each partner i announces his or her type $\tilde{\theta}_i$.
3. Depending on I 's report $\tilde{\theta}_I$:
 - (a) If $\tilde{\theta}_I \leq \underline{\theta}_O$, O is allocated control, and is required to transfer $T = (s_I + s_O - 1)\mathbb{E}[v_O(\theta_O, \theta_O)]$ to I .
 - (b) If $\tilde{\theta}_I \in (\underline{\theta}_O, \overline{\theta}_O)$, partner i is assigned control if $\tilde{\theta}_i > \tilde{\theta}_j$. No direct monetary transfer occurs between the two partners. Partner j receives a transfer $t_j = (s_I + s_O - 1)v_i(\tilde{\theta}_i, \tilde{\theta}_i)$ from the mechanism.
 - (c) If $\tilde{\theta}_I \geq \overline{\theta}_O$, I retains the firm and no monetary transfer occurs.
4. The partner assigned control keeps shares s_i , and the other partner keeps $1 - s_i$.

¹⁹We continue to assume the valuation function $v_i(\theta_i, \theta_j) = \theta_i + \eta(\theta_j)$ with $\theta_I \sim F_I(\cdot)$ on $[0, 1]$.

Lemma 3 in the Online Appendix establishes that to characterize payoff structures that satisfy incentive compatibility under efficient mechanisms, we can restrict attention to this mechanism without loss of generality (similar to the property of M-mechanisms in Ferreira et al. (2015)). By designing the mechanism to further satisfy both individual rationality and the budget balance condition, we can characterize the possibility condition.

We define the expression $EL(s_I, s_O)$ under two-sided private information:

$$\begin{aligned} EL(s_I, s_O) &= (s_I + s_O - 1)\mathbb{E}[\max\{v_I(\theta_I, \theta_I), v_I(\theta_O, \theta_O)\} \mid \theta_I \in (\underline{\theta}_O, \overline{\theta}_O)]Pr(\theta_I \in (\underline{\theta}_O, \overline{\theta}_O)) \\ &\quad - \widehat{NE}_I(\theta_I^*, s_I, s_O) - \widehat{NE}_O(\theta_O^*, s_I, s_O) \end{aligned} \quad (13)$$

Efficient reallocation is possible if and only if $EL(s_I, s_O) \leq 0$ (as established in Proposition 7 in the Online Appendix). In the expression $EL(s_I, s_O)$, $\widehat{NE}_i(\theta_i, s_I, s_O)$ denotes the net expected payoff of partner i excluding the participation fee k_i , and θ_i^* denotes the worst-off type of partner i . The worst-off types of the two partners are given by²⁰:

$$\begin{aligned} \theta_I^* &= G^{-1}\left(\frac{r - (1 - s_O)\eta'(\theta_I^*)}{s_I - (1 - s_O)\eta'(\theta_I^*)}\right) \\ \theta_O^* &= \begin{cases} \underline{\theta}_O & \text{if } \eta'(\cdot) \leq 0 \\ \min\{\max\{F_I^{-1}(\frac{(s_I - r)\eta'(\theta_O^*)}{s_O - (1 - s_I)\eta'(\theta_O^*)}), \underline{\theta}_O\}, \overline{\theta}_O\} & \text{if } \eta'(\cdot) > 0 \end{cases} \end{aligned} \quad (14)$$

By definition, (s_I^E, s_O^E) is identified as the pair minimizing the expression $EL(s_I, s_O)$ —the subsidy required to restore the first-best. Equivalently, it minimizes the sum of information rents. Here, we focus on the intuition for how and why (s_I^E, s_O^E) can differ from the baseline case, leaving all other details to the Online Appendix.

A key change arises from the fact that the minimization objective becomes different. Recall that in the one-sided private information case ($\epsilon = 0$), the optimal (s_I^E, s_O^E) minimizes I 's *ex ante* information rents. Under the two-sided private information case ($\epsilon > 0$), the optimal (s_I^E, s_O^E) instead minimizes the sum of I 's and O 's *ex ante* information rents. Partner O 's information rents respond differently to changes in (s_I, s_O) than I 's information rents do, with their magnitude generally increasing with ϵ .

Beyond bringing in O 's information rents, $\epsilon > 0$ also changes how (s_I, s_O) shape each partner's *ex ante* information rents. In the one-sided private information case ($\epsilon = 0$), I 's worst-off type is independent of (s_I, s_O) : I knows exactly when she earns rents as a seller or buyer. In this case, (s_I, s_O) affect I 's information rents only through the size of traded shares, as established in Section 3. Specifically, the size of the traded shares determines how quickly interim information rents increase as I 's types move away from the worst-off type, and thus shaping I 's *ex ante* information rents.

²⁰For convenience, we express θ_O as following distribution $G(\cdot)$ on $[\underline{\theta}_O, \overline{\theta}_O] = [(1 - \epsilon)a, \epsilon + (1 - \epsilon)a]$, where $G(\cdot)$ is an affine transformation of $F_O(\cdot)$.

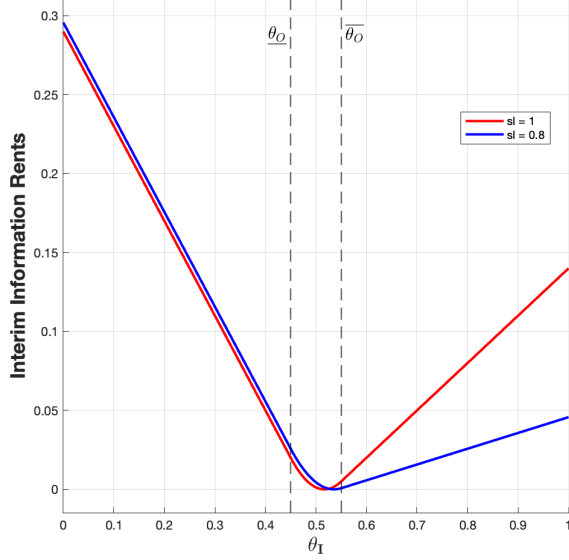
When both partners' private information is relevant ($\epsilon > 0$), (s_I, s_O) affect each partner's *ex ante* information rents through two channels. The first channel, the "trade size channel," is similar to the only channel in the one-sided case: it changes the rate at which interim information rents increase as types move away from the worst-off type, via the size of traded shares. The second channel, the "worst-off type channel," is new: changes in (s_I, s_O) also shift the identity of the worst-off type, thereby changing which types earn rents as sellers versus buyers. Specifically, when $\epsilon > 0$, I 's worst-off type is still confined to the range where O 's private information is relevant, but now varies with (s_I, s_O) , with the value depending on the tradeoff between the expected payoff from each reallocation outcome (as reflected in (14)).

With these differences, two main points can be established. First, as $\epsilon \rightarrow 0$, the optimal (s_I^E, s_O^E) converges to the results in Section 3. This follows because the expression $EL(s_I, s_O)$ in the two-sided case (13) converges to that in the one-sided case ((3) in Section 3). This convergence is intuitive: with small ϵ , O 's *ex ante* information rents are negligible, since he can only obtain information rents when I 's private information is insufficient to determine the efficient allocation (i.e., when $\theta_I \in [\underline{\theta}_O, \overline{\theta}_O]$). In this case, minimizing the sum of information rents effectively reduces to minimizing I 's information rents. Moreover, with small ϵ , I 's information rents behave similarly to the one-sided case. The reason is that I 's worst-off type is confined to the narrow range $[\underline{\theta}_O, \overline{\theta}_O]$ of length ϵ , which significantly constrains the worst-off type channel.

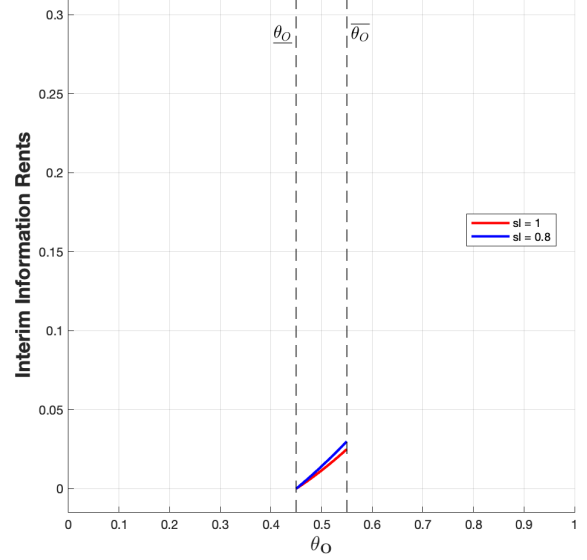
Second, as ϵ grows, the optimal (s_I^E, s_O^E) depends on the balance of all these forces, which can lead to various deviations from the one-sided case. The optimal shares s_i^E may settle at an interim level, $\underline{s}_i < s_i^E < 1$, or, in some cases, even shift more dramatically. For instance, it may move from the lowest feasible level $\underline{s}_i$ to 1 as ϵ increases, as reflected in the following example.

Example 2 Consider the valuation function $v_i(\theta_i, \theta_j) = \theta_i + 0.5\theta_j$. Suppose I 's initial share is $r = 0.7$ and her private information θ_I follows a uniform distribution $\theta_I \sim U[0, 1]$. O 's type is given by $\theta_O = \epsilon\omega + 0.5(1 - \epsilon)$ with $\omega \sim U[0, 1]$. We assume that after reallocation, the assigned controller must hold at least 0.8 ownership of the firm: $S^F = \{(s_I, s_O) \mid s_I \in [0.8, 1], s_O \in [0.8, 1]\}$. We fix $s_O = 0.8$ and explore how changes in s_I affect information rents for each partner. Figures (2a) and (2b) display the two partners' interim information rents when $\epsilon = 0.1$, and Figures (2c) and (2d) display those when $\epsilon = 0.9$. Each figure provides a comparison between $s_I = 0.8$ (red line) and $s_I = 1$ (blue line).

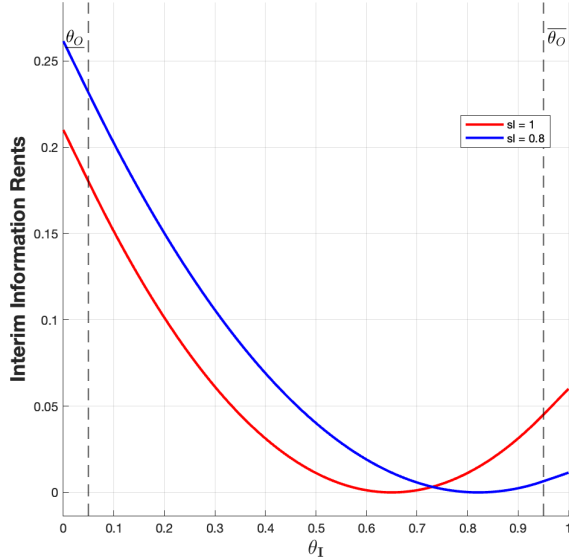
When $\epsilon = 0.1$, we have $s_I^E = 0.8$. As reflected in Figure (2b), O 's information rents are minimal, since O earns information rents only within $[\underline{\theta}_O, \overline{\theta}_O] = [0.45, 0.55]$. For I 's information rents, the worst-off type channel has minimal effect: Figure (2a) shows almost no shift in the worst-off type. As a result, s_I^E can be regarded as determined by its effect on I 's information rents through the trade size channel, which is reflected in the change in slope for types higher than the worst-off type in Figure (2a). Since the trade size channel is the only channel in the one-sided case, the result with $\epsilon = 0.1$ aligns with that case ($s_I^E = \underline{s}_I$).



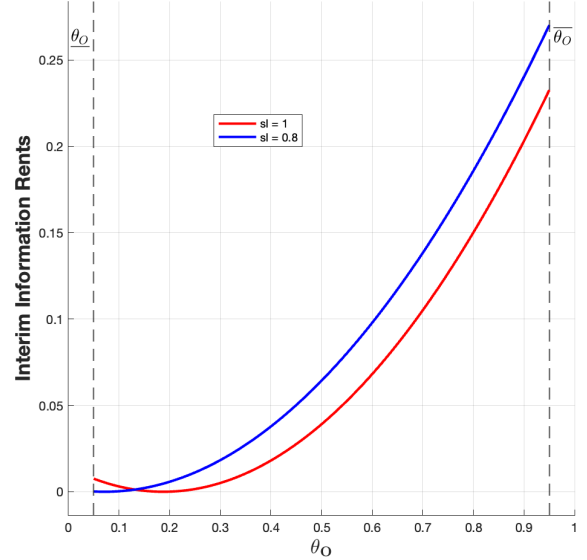
(a) I 's Interim Information Rents, $\epsilon = 0.1$



(b) O 's Interim Information Rents, $\epsilon = 0.1$



(c) I 's Interim Information Rents, $\epsilon = 0.9$



(d) O 's Interim Information Rents, $\epsilon = 0.9$

Figure 2: Interim Information Rents under Different Values of ϵ

When $\epsilon = 0.9$, we have $s_I^E = 1$ instead. When ϵ is large, O 's information rents become important as well. The trade size channel for both I and O has limited impact: changes in s_I mainly affect I 's interim information rents for types above the worst-off type (who tend to be buyers) and O 's interim information rents for types below his worst-off type (who tend to be sellers). However, under the current setup, few types satisfy these conditions. In contrast, the worst-off type channels for the two partners not only have a large effect, but also operate in the same direction: increasing s_I shifts both partners' worst-off types (thereby restructuring interim information rents) in a direction that reduces their ex ante information rents. As shown in Figures (2c) and (2d), the parallel movement caused by the shifting of the

worst-off type (reflecting the worst-off type channel) has a larger effect on information rents than the change in slope (reflecting the trade size channel), leading to a dramatic deviation from the one-sided result— $s_I^E = 1$.

To further illustrate how ϵ and interdependence can yield substantially different (s_I^E, s_O^E) , Figures (3a) and (3b) plot the optimal (s_I^E, s_O^E) (y-axis) against ϵ (x-axis) under two interdependence structures: negative interdependence ($v_i(\theta_i, \theta_j) = \theta_i - 0.5\theta_j$) and positive interdependence ($v_i(\theta_i, \theta_j) = \theta_i + 0.5\theta_j$).

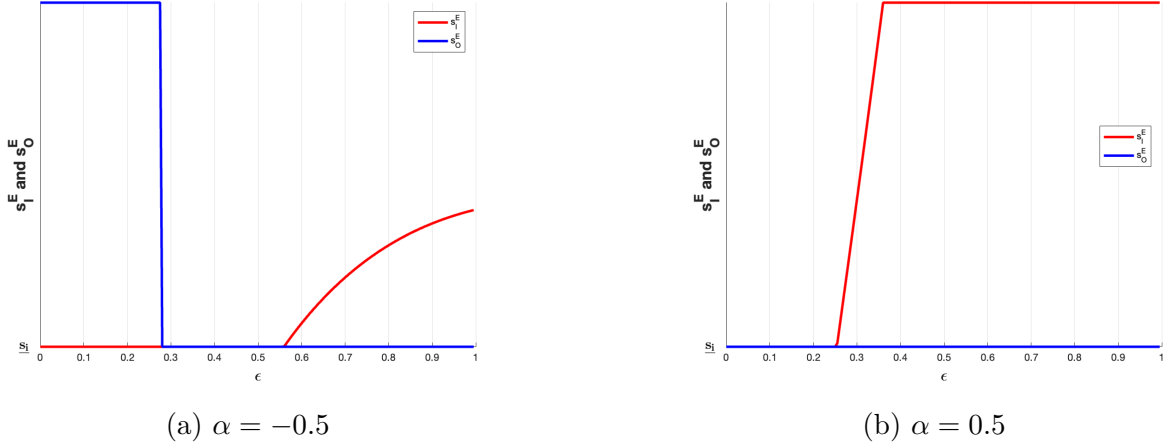


Figure 3: Effect of ϵ on (s_I^E, s_O^E) under Different Interdependence Relationships

5.2 Second-best mechanism with nonlinear $\eta(\cdot)$

In this subsection, we explore how the form of the second-best mechanism and the optimal shares (s_I^{SB}, s_O^{SB}) change when the interdependence relationship is no longer constrained to the linear case (i.e., when $\eta(\cdot)$ can be nonlinear). In the Online Appendix, we characterize the second-best mechanism for the cases when $\eta(\cdot)$ is convex or concave. Here, we focus on the main intuitions.

Under the appropriate assumptions, the construction of the second-best mechanism follows the same principle: give up efficient allocation for a range of I 's types closest to θ_O and replace it with specific allocation(s) ($c(\cdot) = TH(s_I, s_O, \theta_I)$), so all these types become the worst-off. When $\eta(\cdot)$ is linear, $c(\cdot) = TH(s_I, s_O, \theta_I)$ is a constant value, so a single lottery rule can be used to induce a range of worst-off types. When $\eta(\cdot)$ is concave, $TH(s_I, s_O, \theta_I)$ becomes increasing in θ_I . Following the same principle, the lottery rule $c(\cdot) = TH(s_I, s_O, \theta_I)$ is applied to generate a region of worst-off types, but now higher types of I receive control with strictly higher probabilities. However, when $\eta(\cdot)$ is convex, $TH(s_I, s_O, \theta_I)$ is a decreasing function of θ_I , which creates a conflict: incentive compatibility requires $c(\cdot)$ to be weakly increasing in θ_I , making it impossible to induce a range of worst-off types. Since constructing a zero-rent region is infeasible, the second-best mechanism instead chooses a type θ_I^C ,

such that applying $c(\theta_I^C) = TH(s_I, s_O, \theta_I^C)$ most effectively loosens the implementability constraint.

With a concave $\eta(\cdot)$, although the second-best mechanism takes a different form, the intuition from the linear case carries over, and Propositions 4 and 5 continue to apply. This is because the mechanism still maintains one lottery region with zero information rents and remaining efficient regions with positive information rents, allowing the two channels (information rent channel and lottery loss channel) to operate in the same way as in the linear case. Under convex $\eta(\cdot)$, however, these propositions no longer hold. The reason is that, except for the unique worst-off type θ_I^C , all types in the lottery region earn positive information rents. As a result, the information rent channel operates not only in the efficient region but also in the lottery region, with an undetermined direction. This makes the overall effect of (s_I, s_O) ambiguous, even in the extreme case where efficiency always assigns control to O .

6 Conclusion

This paper studies how equity transfers should be optimally designed to facilitate efficient control transitions under asymmetric information. Using a mechanism design approach, we show that the nature of expertise interdependence is a key determinant of optimal equity structure. Specifically, complementary skill sets favor the initial controller retaining partial ownership after transition, while substitute skill sets call for complete dissolution—suggesting the initial controller be bought out entirely. This complement-substitute prescription is robust: it extends to cases where the gains from restructuring are more modest, particularly when control transition is on average more beneficial; and extends to cases where the other partner also holds private information, as long as the initial controller’s information advantage remains substantially stronger.

In practice, many other factors shape optimal ownership restructuring. For example, ownership structure itself affects the effort the newly assigned controller invests, thereby influencing the firm’s future value. Our model abstracts from this channel by assuming full effort above a minimum ownership threshold. Liquidity constraints may also prevent the partner assigned control from purchasing all shares, providing an additional rationale for partial ownership. Future work could incorporate these factors, thereby further enriching our understanding of how different frictions could jointly shape restructuring contract design.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

During the preparation of this work, the author used Claude to proofread existing content. After using this tool/service, the author reviewed and edited the content as needed. The author takes full responsibility for the content of the publication.

References

- Cramton, P., Gibbons, R., and Klemperer, P. (1987). Dissolving a partnership efficiently. *Econometrica*, pages 615–632.
- d’Aspremont, C. and Gérard-Varet, L.-A. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11(1):25–45.
- DeMarzo, P. M., Kremer, I., and Skrzypacz, A. (2005). Bidding with securities: Auctions and security design. *American Economic Review*, 95(4):936–959.
- Edmans, A., Gabaix, X., and Jenter, D. (2017). Executive compensation: A survey of theory and evidence. *The Handbook of the Economics of Corporate Governance*, 1:383–539.
- Ekmekci, M., Kos, N., and Vohra, R. (2016). Just enough or all: Selling a firm. *American Economic Journal: Microeconomics*, 8(3):223–256.
- Ferreira, D., Ornelas, E., and Turner, J. L. (2015). Unbundling ownership and control. *Journal of Economics & Management Strategy*, 24(1):1–21.
- Fershtman, D., Szabadi, B., and Wasser, C. (2023). Efficient resolution of partnership disputes. *The RAND Journal of Economics*, 54(4):543–569.
- Fieseler, K., Kittsteiner, T., and Moldovanu, B. (2003). Partnerships, lemons, and efficient trade. *Journal of Economic Theory*, 113(2):223–234.
- Fioriti, A. and Hernandez-Chanto, A. (2022). Leveling the playing field for risk-averse agents in security-bid auctions. *Management Science*, 68(7):5441–5463.
- Galavotti, S., Muto, N., and Oyama, D. (2011). On efficient partnership dissolution under ex post individual rationality. *Economic Theory*, 48(1):87–123.
- Hansen, R. G. (1985). Auctions with contingent payments. *American Economic Review*, 75(4):862–865.
- Hernandez-Chanto, A. and Fioriti, A. (2019). Bidding securities in projects with negative externalities. *European Economic Review*, 118:14–36.
- Jehiel, P. and Pauzner, A. (2006). Partnership dissolution with interdependent values. *The RAND Journal of Economics*, 37(1):1–22.
- Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.
- Kunimoto, T. and Zhang, C. (2026). Efficient bilateral trade with interdependent values: The use of two-stage mechanisms. *Journal of Mathematical Economics*, 122:103191.
- Liu, T. and Bernhardt, D. (2021). Rent extraction with securities plus cash. *The Journal of Finance*, 76(4):1869–1912.
- Loertscher, S. and Wasser, C. (2019). Optimal structure and dissolution of partnerships. *Theoretical Economics*, 14(3):1063–1114.

- Mezzetti, C. (2004). Mechanism design with interdependent valuations: Efficiency. *Econometrica*, 72(5):1617–1626.
- Mezzetti, C. (2007). Mechanism design with interdependent valuations: Surplus extraction. *Economic Theory*, 31(3):473–488.
- Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281.
- Mylovanov, T. and Zapechelnyuk, A. (2017). Optimal allocation with ex post verification and limited penalties. *American Economic Review*, 107(9):2666–2694.
- Ornelas, E. and Turner, J. L. (2007). Efficient dissolution of partnerships and the structure of control. *Games and Economic Behavior*, 60(1):187–199.
- Pan, L. and Wang, D. (2021). The broker-optimal bilateral trading mechanisms with linear contracts. *Economics Letters*, 208:110055.
- Skrzypacz, A. (2013). Auctions with contingent payments—an overview. *International Journal of Industrial Organization*, 31(5):666–675.
- Williams, S. R. (1999). A characterization of efficient, bayesian incentive compatible mechanisms. *Economic Theory*, 14(1):155–180.
- Yoon, K. (2020). Bilateral trading with contingent contracts. *International Journal of Game Theory*, 49(2):445–461.

Appendix

Proof of Lemma 1

The proof follows a similar approach to Jehiel and Pauzner (2006) (Lemma 1). The main difference is that in our setup, I 's ownership payoff is $s_I v_I(\theta_I, \theta_O)$ upon retaining control and $(1 - s_O)v_O(\theta_O, \theta_I)$ upon losing control, rather than full ownership $v_I(\theta_I, \theta_O)$ and zero respectively.

(i) From *ICI*, for any θ_I and $\tilde{\theta}_I$, the following two inequalities must hold:

$$t(\theta_I) - t(\tilde{\theta}_I) \geq (c(\theta_I) - c(\tilde{\theta}_I))((1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)) \quad (15)$$

$$t(\theta_I) - t(\tilde{\theta}_I) \leq (c(\theta_I) - c(\tilde{\theta}_I))((1 - s_O)v_O(\theta_O, \tilde{\theta}_I) - s_I v_I(\tilde{\theta}_I, \theta_O)) \quad (16)$$

Combining them together, we have

$$(c(\theta_I) - c(\tilde{\theta}_I))[s_I(v_I(\theta_I, \theta_O) - v_I(\tilde{\theta}_I, \theta_O)) - (1 - s_O)(v_O(\theta_O, \theta_I) - v_O(\theta_O, \tilde{\theta}_I))] \geq 0.$$

The expression $s_I v_I(\theta_I, \theta_O) - (1 - s_O)v_O(\theta_O, \theta_I)$ is increasing in θ_I by assumption. Therefore, $c(\theta_I)$ must be (weakly) increasing in θ_I .

(ii) For inequalities (15) and (16), we replace θ_I with $\theta_I + \varepsilon$ and replace $\tilde{\theta}_I$ with θ_I . Combining them together, we have:

$$\begin{aligned} & (c(\theta_I + \varepsilon) - c(\theta_I))(1 - s_O)v_O(\theta_O, \theta_I) - (c(\theta_I + \varepsilon) - c(\theta_I))s_I v_I(\theta_I, \theta_O) \\ & \geq t(\theta_I + \varepsilon) - t(\theta_I) \geq \\ & (c(\theta_I + \varepsilon) - c(\theta_I))(1 - s_O)v_O(\theta_O, \theta_I + \varepsilon) - (c(\theta_I + \varepsilon) - c(\theta_I))s_I v_I(\theta_I + \varepsilon, \theta_O) \end{aligned} \quad (17)$$

Taking limit $\varepsilon \rightarrow 0$ yields (ii).

(iii) Assuming $c(\theta_I)$ is continuous at θ_I , if we divide the inequality (17) by ε and take the limit $\varepsilon \rightarrow 0$, we have:

$$\begin{aligned} t'(\theta_I) &= \lim_{\varepsilon \rightarrow 0} \frac{t(\theta_I + \varepsilon) - t(\theta_I)}{\varepsilon} \\ &= ((1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)) \frac{\partial(c(\theta_I + \varepsilon) - c(\theta_I))}{\partial \varepsilon} \\ &= ((1 - s_O)v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O))c'(\theta_I) \end{aligned}$$

(iv) We denote the points at which $c(\cdot)$ is not differentiable in the interval $(0, \theta_I)$ as $\theta_{I,1}, \theta_{I,2}, \theta_{I,3}, \dots, \theta_{I,n}$. Also, $\theta_{I,0} = 0$ and $\theta_{I,n+1} = \theta_I$.

For ε sufficiently close to 0, the transfer $t(\theta_I)$ can be expressed as:

$$\begin{aligned}
t(\theta_I) &= t(0) + t(0 + \varepsilon) - t(0) + \sum_{i=1}^{n+1} [t(\theta_{I,i} - \varepsilon) - t(\theta_{I,i-1} + \varepsilon)] \\
&\quad + \sum_{i=1}^n [t(\theta_{I,i} + \varepsilon) - t(\theta_{I,i} - \varepsilon)] + t(\theta) - t(\theta_I - \varepsilon) \\
&= t(0) + t(0 + \varepsilon) - t(0) + \sum_{i=1}^{n+1} \left[\int_{x=\theta_{I,i-1}+\varepsilon}^{\theta_{I,i}-\varepsilon} t'(x) dx \right] + \sum_{i=1}^n [t(\theta_{I,i} + \varepsilon) - t(\theta_{I,i} - \varepsilon)] \\
&\quad + t(\theta_I) - t(\theta_I - \varepsilon).
\end{aligned}$$

Taking the limit $\varepsilon \rightarrow 0$ and applying lemma (ii) and (iii), we have: (we use $c(\theta_I^-)$ and $c(\theta_I^+)$) to denote $\lim_{\tilde{\theta}_I \rightarrow \theta_I, \tilde{\theta}_I < \theta_I} c(\tilde{\theta}_I)$ and $\lim_{\tilde{\theta}_I \rightarrow \theta_I, \tilde{\theta}_I > \theta_I} c(\tilde{\theta}_I)$

$$\begin{aligned}
t(\theta_I) &= t(0) + (c(0^+) - c(0)) \cdot ((1 - s_O) v_O(\theta_O, 0) - s_I v_I(0, \theta_O)) \\
&\quad + \sum_{i=1}^{n+1} \int_{\theta_{I,i-1}}^{\theta_{I,i}} c'(x) [(1 - s_O) v_O(\theta_O, x) - s_I v_I(x, \theta_O)] dx \\
&\quad + \sum_{i=1}^n (c(\theta_{I,i}^+) - c(\theta_{I,i}^-)) \cdot [(1 - s_O) v_O(\theta_O, \theta_{I,i}) - s_I v_I(\theta_{I,i}, \theta_O)] \\
&\quad + (c(\theta_I) - c(\theta_I^-)) \cdot [(1 - s_O) v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]
\end{aligned}$$

Through integration by part, the expression $\int_{\theta_{I,i-1}}^{\theta_{I,i}} c'(x) \cdot [(1 - s_O) v_O(\theta_O, x) - s_I v_I(x, \theta_O)] dx$ can be written as:

$$\begin{aligned}
&\int_{\theta_{I,i-1}}^{\theta_{I,i}} c(x) \left[s_I \frac{\partial v_I(x, \theta_O)}{\partial x} - (1 - s_O) \frac{\partial v_O(\theta_O, x)}{\partial x} \right] dx + c(\theta_{I,i}^-) [(1 - s_O) v_O(\theta_O, \theta_{I,i}^-) - s_I v_I(\theta_{I,i}^-, \theta_O)] \\
&- c(\theta_{I,i-1}^+) [(1 - s_O) v_O(\theta_O, \theta_{I,i-1}^+) - s_I v_I(\theta_{I,i-1}^+, \theta_O)]
\end{aligned}$$

Substituting it back to the expression of $t(\theta_I)$, we have:

$$\begin{aligned}
t(\theta_I) = & t(0) + (c(0^+) - c(0)) \cdot ((1 - s_O) v_O(\theta_O, 0) - s_I v_I(0, \theta_O)) \\
& + \sum_{i=1}^{n+1} \left[\int_{\theta_{I,i-1}}^{\theta_{I,i}} c(x) \left[s_I \frac{\partial v_I(x, \theta_O)}{\partial x} - (1 - s_O) \frac{\partial v_O(\theta_O, x)}{\partial x} \right] dx \right. \\
& + c(\theta_{I,i}^-) [(1 - s_O) v_O(\theta_O, \theta_{I,i}^-) - s_I v_I(\theta_{I,i}^-, \theta_O)] - c(\theta_{I,i-1}^+) [(1 - s_O) v_O(\theta_O, \theta_{I,i-1}^+) - s_I v_I(\theta_{I,i-1}^+, \theta_O)] \left. \right] \\
& + \sum_{i=1}^n \left[(c(\theta_{I,i}^+) - c(\theta_{I,i}^-)) \cdot [(1 - s_O) v_O(\theta_O, \theta_{I,i}) - s_I v_I(\theta_{I,i}, \theta_O)] \right. \\
& + (c(\theta_I) - c(\theta_I^-)) \cdot [(1 - s_O) v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)] \left. \right]
\end{aligned}$$

After re-arranging:

$$\begin{aligned}
t(\theta_I) = & t(0) - c(0) [(1 - s_O) v_O(\theta_O, 0) - s_I v_I(0, \theta_O)] + \int_0^{\theta_I} c(x) \left[s_I \frac{\partial v_I(x, \theta_O)}{\partial x} - (1 - s_O) \frac{\partial v_O(\theta_O, x)}{\partial x} \right] dx \\
& + c(\theta_I) [(1 - s_O) v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]
\end{aligned}$$

Therefore, the mechanism is incentive compatible if and only if $t(\theta_I) = m + \int_0^{\theta_I} c(x) \left[s_I \frac{\partial v_I(x, \theta_O)}{\partial x} - (1 - s_O) \frac{\partial v_O(\theta_O, x)}{\partial x} \right] dx + c(\theta_I) [(1 - s_O) v_O(\theta_O, \theta_I) - s_I v_I(\theta_I, \theta_O)]$, where $m = t(0) - c(0) [(1 - s_O) v_O(\theta_O, 0) - s_I v_I(0, \theta_O)]$ and $c(\theta_I)$ is weakly increasing in θ_I .

Proof of Proposition 1

Take derivative of $EL(s_I, s_O)$ with respect to s_O , we have:

$$\int_0^{\theta_O} (v_O(\theta_O, \theta_O) - v_O(\theta_O, x)) dF_I(x) \quad (18)$$

Under positive interdependence, $\frac{\partial v_i(\theta_i, \theta_j)}{\partial \theta_j} > 0$, the expression (18) is positive, implying the $EL(s_I, s_O)$ increase with s_O . To minimize the $EL(s_I, s_O)$, the optimal $s_O^E = \underline{s_O}$. By contrast, under negative interdependence, the expression (18) is negative, implying $EL(s_I, s_O)$ decreases with s_O . To minimize $EL(s_I, s_O)$, $s_O^E = 1$.

Take derivative of $EL(s_I, s_O)$ with respect to s_I yields:

$$\int_{\theta_O}^1 (v_I(x, \theta_O) - v_I(\theta_O, \theta_O)) dF_I(x)$$

Since $v_i(\theta_i, \theta_j)$ is increasing in θ_i , expression above must be larger than zero. To minimize EL , $s_I^E = \underline{s_I}$ regardless of positive or negative interdependence.

Proof of Lemma 2

Recall that the value of the firm is captured by $v_i(\theta_i, \theta_j) = \theta_i + \eta(\theta_j)$. Based on Lemma 1 (i) and (iv), the *ICI* holds if and only if :

$$t(\theta_I) = \int_0^{\theta_I} c(x)[s_I - (1 - s_O)\eta'(x)]dx + c(\theta_I)[(1 - s_O)(\theta_O + \eta(\theta_I)) - s_I(\theta_I + \eta(\theta_O))] + m$$

for some constant m and $c(\theta_I)$ is weakly increasing in θ_I .

Substituting the expression into *IRI* yields:

$$\int_0^{\theta_I} c(x)[s_I - (1 - s_O)\eta'(x)]dx + (1 - s_O)(\theta_O + \eta(\theta_I)) - r(\theta_I + \eta(\theta_O)) + m \geq 0 \quad (19)$$

To guarantee *IRI* being satisfied for all types, it is equivalent to ensure that *IRI* is satisfied for the worst-off type. The worst-off type must minimize the left-hand side expression of inequality (19). Taking derivative of the left-hand side with respect to θ_I gives: $c(\theta_I)[s_I - (1 - s_O)\eta'(\theta_I)] - (r - (1 - s_O)\eta'(\theta_I))$. The monotonicity depends on the relationship between $c(\theta_I)[s_I - (1 - s_O)\eta'(\theta_I)]$ and $(r - (1 - s_O)\eta'(\theta_I))$, or equivalently, depends on the relationship between allocation rule $c(\theta_I)$ and function $TH(\theta_I, s_I, s_O) = \frac{r - (1 - s_O)\eta'(\theta_I)}{s_I - (1 - s_O)\eta'(\theta_I)}$. Under the assumption that $\eta(\cdot)$ is a linear function, $TH(\theta_I, s_I, s_O)$ is irrelevant to the value of θ_I and reduces to $TH(s_I, s_O)$.

Suppose in a mechanism satisfying *ICI*, $c(\theta_I) = TH(s_I, s_O)$ for types $\theta_I \in (\theta_I^*, \theta_I^{**})$, then these types are the worst-off types. The left-hand side of inequality (19) is decreasing function of θ_I for types $\theta_I < \theta_I^*$ (since for them, $c(\theta_I) < TH(s_I, s_O)$, as *ICI* implies that $c(\theta_I)$ must be increasing in θ_I) and increasing function of θ_I for types $\theta_I > \theta_I^{**}$ (for them, $c(\theta_I) > TH(s_I, s_O)$).

Proof of Proposition 2

The proof follows a similar approach to Jehiel and Pauzner (2006) (Proposition 3). The key difference is that the implementability constraint now depends on the shares (s_I, s_O) through the threshold $TH(s_I, s_O)$, which induces the worst-off types, and through information rent terms. Additionally, the status quo in our setup would incur a welfare adjustment when inducing worst-off types (reallocating control to I with probability $TH(s_I, s_O)$).

IRI is satisfied for all types of I if and only if it is satisfied for the worst-off type θ_I^* . Then, for a mechanism that satisfies both *IRI* and *ICI*, we must have

$$\int_0^{\theta_I^*} c(x)[s_I - (1 - s_O)\eta'(\cdot)]dx + (1 - s_O)(\theta_O + \eta(\theta_I^*)) - r(\theta_I^* + \eta(\theta_O)) + m \geq 0 \quad (I)$$

for some constant m , where $\theta_I^* = \sup\{\theta_I : c(\theta_I) < TH(s_I, s_O)\}$.

Substituting transfer rule from Lemma (iv) to the *IRO* condition yields:

$$\begin{aligned} & \int_0^1 \left[((1 - c(\theta_I))s_O - c(\theta_I)(1 - s_O))(\theta_O + \eta(\theta_I)) + (c(\theta_I) - (1 - r))(\theta_I + \eta(\theta_O)) - r\theta_I \right. \\ & \left. + (1 - s_O)\eta(\theta_I) - \int_0^{\theta_I} [(c(x) - TH(s_I, s_O))(s_I - (1 - s_O)\eta'(\cdot))] dx \right] dF_I(\theta_I) - m \geq 0 \end{aligned} \quad (\text{II})$$

Therefore, there exists a constant m that can both satisfy I and II if and only if:

$$\begin{aligned} & \int_0^{\theta_I^*} c(x)[s_I - (1 - s_O)\eta'(\cdot)]dx + (1 - s_O)(\theta_O + \eta(\theta_I^*)) - r(\theta_I^* + \eta(\theta_O)) \\ & + \int_0^1 \left[[(1 - c(\theta_I))s_O - c(\theta_I)(1 - s_O)](\theta_O + \eta(\theta_I)) + (c(\theta_I) - (1 - r))(\theta_I + \eta(\theta_O)) \right. \\ & \left. - r\theta_I + (1 - s_O)\eta(\theta_I) - \int_0^{\theta_I} (c(x) - TH(s_I, s_O))(s_I - (1 - s_O)\eta'(\cdot))dx \right] dF_I(\theta_I) \geq 0 \end{aligned} \quad (\text{III})$$

We can decompose the left-hand side of (III) into two parts, and rewrite them as: 1) $\int_0^{\theta_I^*} (c(x) - TH(s_I, s_O))[s_I - (1 - s_O)\eta'(\cdot)]dx - \int_0^1 \int_0^{\theta_I} (c(x) - TH(s_I, s_O))(s_I - (1 - s_O)\eta'(\cdot))dx dF_I(\theta_I)$, and 2) $\int_0^1 (1 - c(\theta_I))(\theta_O + \eta(\theta_I)) + (c(\theta_I) - 1)(\theta_I + \eta(\theta_O))dF_I(\theta_I)$.

Through integration by parts, the part 1) can be further rewritten as:

$$\begin{aligned} & \int_0^{\theta_I^*} (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] dx \\ & - \int_0^1 (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] (1 - F_I(x)) dx \\ & = \int_0^{\theta_I^*} (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] F_I(x) dx \\ & - \int_{\theta_I^{**}}^1 (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] (1 - F_I(x)) dx \\ & = \int_0^{\theta_I^*} (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] \frac{F_I(x)}{f_I(x)} dF_I(x) \\ & - \int_{\theta_I^{**}}^1 (c(x) - TH(s_I, s_O)) [s_I - (1 - s_O) \eta'(\cdot)] \frac{1 - F_I(x)}{f_I(x)} dF_I(x) \end{aligned}$$

As for the part 2):

$$\begin{aligned}
& \int_0^1 (1 - c(\theta_I))(\theta_O + \eta(\theta_I)) + (c(\theta_I) - 1)(\theta_I + \eta(\theta_O)) dF_I(\theta_I) \\
&= \int_0^1 (TH(s_I, s_O) - c(\theta_I))(\theta_O + \eta(\theta_I)) + (c(\theta_I) - TH(s_O, s_O))(\theta_I + \eta(\theta_O)) dF_I(\theta_I) \\
&+ \int_0^1 (1 - TH(s_I, s_O))(\theta_O + \eta(\theta_I)) + (TH(s_I, s_O) - 1)(\theta_I + \eta(\theta_O)) dF_I(\theta_I)
\end{aligned}$$

Combining them together, (III) can be rewritten as:

$$\begin{aligned}
& \int_0^{\theta_I^*} (c(\theta_I) - TH(s_I, s_O))[\theta_I + \eta(\theta_O) + (s_I - (1 - s_O))\eta'(\cdot)] \frac{F_I(\theta_I)}{f_I(\theta_I)} - \theta_O - \eta(\theta_I)] dF_I(\theta_I) \\
&+ \int_{\theta_I^{**}}^1 (c(\theta_I) - TH(s_I, s_O))[\theta_I + \eta(\theta_O) - (s_I - (1 - s_O))\eta'(\cdot)] \frac{1 - F_I(\theta_I)}{f_I(\theta_I)} - \theta_O - \eta(\theta_I)] dF_I(\theta_I) \\
&+ (1 - TH(s_I, s_O)) \int_0^1 (\theta_O + \eta(\theta_I) - \theta_I - \eta(\theta_O)) dF_I(\theta_I) \geq 0
\end{aligned}$$

We have the implementability constraint 5 once we substitute the expression $\Delta_{NW}(\theta_I, \theta_O) = v_I(\theta_I, \theta_O) - v_O(\theta_O, \theta_I) = \theta_I + \eta(\theta_O) - \theta_O - \eta(\theta_I)$ into the inequality above. And the second-best mechanism must both satisfy the inequality 5 and that $c(\theta_I)$ is weakly increasing in θ_I so that *ICI*, *IRI* and *IRO* are all satisfied.

Proof of Proposition 3

The Hamiltonian of the maximization problem of Proposition 2 can be written as below. We use $\lambda \geq 0$ as the Lagrange multiplier of the constraint:

$$\begin{aligned}
& (1 + \lambda) \int_0^{\theta_I^*} (c(\theta_I) - TH(s_I, s_O))[\theta_I + \eta(\theta_O) + \frac{\lambda}{1 + \lambda}(s_I - (1 - s_O))\eta'(\cdot)] \frac{F_I(\theta_I)}{f_I(\theta_I)} - \theta_O - \eta(\theta_I)] dF_I(\theta_I) \\
&+ (1 + \lambda) \int_{\theta_I^{**}}^1 (c(\theta_I) - TH(s_I, s_O))[\theta_I + \eta(\theta_O) - \frac{\lambda}{1 + \lambda}(s_I - (1 - s_O))\eta'(\cdot)] \frac{1 - F_I(\theta_I)}{f_I(\theta_I)} - \theta_O - \eta(\theta_I)] dF_I(\theta_I) \\
&+ \int_0^1 [(TH(s_I, s_O) - \lambda + \lambda TH(s_I, s_O))(\theta_I + \eta(\theta_O)) + (1 + \lambda)(1 - TH(s_I, s_O))(\theta_O + \eta(\theta_I))] dF_I(\theta_I)
\end{aligned}$$

We want to identify the allocation rule $c(\theta_I)$ that maximizes the above expression, but we have to restrict attention to allocation rule $c(\theta_I)$ that is weakly increasing in θ_I , as this is a necessary condition for *ICI*. As in Jehiel and Paudner (2006), we could consider separately what $c(\theta_I)$ should look like for types $\theta_I < \theta_I^*$ and $\theta_I > \theta_I^{**}$ in the second-best mechanism. For types $\theta_I \in (\theta_I^*, \theta_I^{**})$, $c(\theta_I) = TH(s_I, s_O)$ by definition.

For the region $\theta_I < \theta_I^*$: suppose there exists a type $\tilde{\theta}_I$ such that $0 < c(\tilde{\theta}_I) < TH(s_I, s_O)$. We show that modifying the allocation rule for $\tilde{\theta}_I$ and its neighborhood strictly improves welfare without violating any constraints.

Define

$$N(\theta_I) = \Delta_{NW}(\theta_I, \theta_O) + \frac{\lambda}{1+\lambda}(s_I - (1-s_O)\eta'(\cdot))\frac{F_I(\theta_I)}{f_I(\theta_I)}$$

Since $\Delta_{NW}(\theta_I, \theta_O) = \theta_I + \eta(\theta_O) - \theta_O - \eta(\theta_I)$ is increasing in θ_I , and $\frac{F_I(\theta_I)}{f_I(\theta_I)}$ is increasing in θ_I as well by Assumption 1, $N(\theta_I)$ is increasing for all types $\theta_I < \theta_I^*$. Then consider two cases:

Case 1: $N(\tilde{\theta}_I) > 0$. For all $\theta_I \in [\tilde{\theta}_I, \theta_I^*]$, we have $N(\theta_I) > 0$ and $c(\theta_I) < TH(s_I, s_O)$, so the contribution to the Lagrangian for these types is $(c(\theta_I) - TH(s_I, s_O)) \cdot N(\theta_I) < 0$. Setting $c(\theta_I) = TH(s_I, s_O)$ for types on $[\tilde{\theta}_I, \theta_I^*]$ strictly increases the Lagrangian, while preserving the monotonicity in $c(\theta_I)$.

Case 2: $N(\tilde{\theta}_I) \leq 0$. Since $N(\cdot)$ is increasing, $N(\theta_I) \leq 0$ for all $\theta_I \leq \tilde{\theta}_I$. For these types, the contribution $(c(\theta_I) - TH(s_I, s_O)) \cdot N(\theta_I) > 0$. Setting $c(\theta_I) = 0$ on $[0, \tilde{\theta}_I]$ strictly increases this product. Again, the monotonicity is preserved.

As illustrated above, having $0 < c(\tilde{\theta}_I) < TH(s_I, s_O)$ cannot be optimal. Instead, $c(\theta_I) = 0$ is optimal for the range $[0, \theta_I^*)$.

A similar argument can be used to show that it is optimal $c(\theta_I) = 1$ for types $\theta_I \in (\theta_I^{**}, 1]$. Therefore, second-best allocation given (s_I, s_O) must follow:

$$c(\theta_I) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^* \\ TH(s_I, s_O) & \text{for } \theta_I^* \leq \theta_I \leq \theta_I^{**} \\ 1 & \text{for } \theta_I > \theta_I^{**} \end{cases}$$

We substitute this allocation rule into the implementability constraint 5. By definition of the second-best mechanism, this constraint must be binding, so we have equation (7).

With the existence of an interior solution, differentiating with respect to θ_I^* and θ_I^{**} would yield:

$$\theta_O + \eta(\theta_I^*) - \theta_I^* - \eta(\theta_O) = \frac{\lambda}{1+\lambda}(s_I - (1-s_O)\eta'(\cdot))\frac{F_I(\theta_I^*)}{f_I(\theta_I^*)}$$

and

$$\theta_I^{**} + \eta(\theta_O) - \theta_O - \eta(\theta_I^{**}) = \frac{\lambda}{1+\lambda}(s_I - (1-s_O)\eta'(\cdot))\frac{1 - F_I(\theta_I^{**})}{f_I(\theta_I^{**})}$$

Combining them together, we have the second equation (8).

Proof of Proposition 4

We define the following two expressions:

$$\begin{aligned} & EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O)) \\ &= \int_0^{\theta_I^*(s_I, s_O)} \left[s_O v_O(\theta_O, x) - (1-r)v_I(x, \theta_O) \right. \\ &\quad \left. - r v_I(\theta_I^*(s_I, s_O), \theta_O) + (1-s_O)v_O(\theta_O, \theta_I^*(s_I, s_O)) \right] dF_I(x) \end{aligned} \quad (20)$$

$$\begin{aligned} & EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O)) \\ &= \int_{\theta_I^*(s_I, s_O)}^1 (1 - TH(s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x) \end{aligned} \quad (21)$$

We can decompose the left-hand side of equation (10) into two parts: part relevant to the efficient allocation region (denoted by $EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O))$), and part relevant to the lottery region (denoted by $EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O))$). We use $\theta_I^*(s_I, s_O)$ to denote the solution to (10) under (s_I, s_O) . Then, the following expressions must hold:

$$EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O)) + EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O)) = 0. \quad (22)$$

We denote the second-best mechanism under (s_I, s_O) as $\langle c^{SB}(\theta_I, s_I, s_O), t, s_I, s_O \rangle$. Now we construct an alternative mechanism denoted as $\langle c(\theta_I, s_I, \underline{s_O}), t', s_I, \underline{s_O} \rangle$. We define the allocation rule $c(\theta_I, s_I, \underline{s_O})$ as:

$$c(\theta_I, s_I, \underline{s_O}) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^*(s_I, s_O) \\ TH(s_I, \underline{s_O}) & \text{for } \theta_I \geq \theta_I^*(s_I, s_O) \end{cases} \quad (23)$$

Both t and t' are uniquely determined by other elements of mechanism. Since they are not directly relevant to the proof, we do not specify them here.

First, it can be established that this alternative mechanism must be implementable. Note that the implementability constraints for the constructed mechanism can be represented by

$$EP_{ER}(s_I, \underline{s_O}, \theta_I^*(s_I, s_O)) + EP_{LR}(s_I, \underline{s_O}, \theta_I^*(s_I, s_O)) \geq 0. \quad (24)$$

The left-hand side of this implementability constraint (24) falls to the left-hand side of (10) if replacing $\underline{s_O}$ with s_O . Therefore, we can conclude which mechanism's implementability constraint is easier to satisfy by looking at how the left-hand side value of (24) changes with the s_O . Since both $EP_{ER}(s_O, s_I, \theta_I^*(s_I, s_O))$ and $EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O))$ are decreasing

functions of s_O (under positive interdependence), the constraint for the constructed mechanism (24) must be strictly easier than the constraint of the second-best mechanism under (s_I, s_O) . Since second-best mechanism's constraint binds by definition (reflected by 10), the constructed mechanism must be implementable.

It can be further shown that this constructed mechanism also yields a higher expected welfare. For the second-best allocation under (s_I, s_O) , the corresponding welfare is

$$\int_0^{\theta_I^*(s_I, s_O)} \theta_O + \eta(x) dF(x) + \int_{\theta_I^*(s_I, s_O)}^1 TH(s_I, s_O) \cdot (x + \eta(\theta_O)) + (1 - TH(s_I, s_O)) \cdot (\theta_O + \eta(x)) dF(x).$$

If we instead apply the constructed rule, the corresponding welfare is

$$\int_0^{\theta_I^*(s_I, \underline{s}_O)} \theta_O + \eta(x) dF(x) + \int_{\theta_I^*(s_I, \underline{s}_O)}^1 TH(s_I, \underline{s}_O) \cdot (x + \eta(\theta_O)) + (1 - TH(s_I, \underline{s}_O)) \cdot (\theta_O + \eta(x)) dF(x),$$

which is obviously higher. The reason is that it allocates control to the efficient controller O with a higher probability (which decreases the welfare loss) while keeping the efficient allocation region the same.

Note that the second-best mechanism under $(s_I, \underline{s}_O)$ must outperform the constructed mechanism by definition, since the implementability constraint for the second-best mechanism must be binding. By combining the following points: (1) the constructed mechanism under $(s_I, \underline{s}_O)$ is implementable; (2) the constructed mechanism under $(s_I, \underline{s}_O)$ yields a higher welfare than the second-best mechanism under (s_I, s_O) for any $s_O > \underline{s}_O$; (3) the second-best mechanism under $(s_I, \underline{s}_O)$ must yield a higher welfare than the constructed mechanism, we can conclude that $s_O^{SB} = \underline{s}_O$. We can apply the similar methodology and show that $s_I^{SB} = 1$.

Proof of Proposition 5

For the general case, we define the expressions $EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O))$ and $EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O))$ as:

$$\begin{aligned} & EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O)) \\ &= \int_0^{\theta_I^*(s_I, s_O)} \left[s_O v_O(\theta_O, x) - (1 - r) v_I(x, \theta_O) - r v_I(\theta_I^*(s_I, s_O), \theta_O) \right. \\ & \quad \left. + (1 - s_O) v_O(\theta_O, \theta_I^*(s_I, s_O)) \right] dF_I(x) \\ & \quad + \int_{\theta_I^{**}(s_I, s_O)}^1 \left[(r - s_I) v_I(x, \theta_O) + (s_I - r) v_I(\theta_I^{**}(s_I, s_O), \theta_O) \right] dF_I(x) \end{aligned} \tag{25}$$

$$\begin{aligned}
& EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O)) \\
&= \int_{\theta_I^*(s_I, s_O)}^{\theta_I^{**}(s_I, s_O)} (1 - TH(s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x)
\end{aligned} \tag{26}$$

The left-hand side of the equation (7) consists of parts relevant to efficient region (expression 25) and lottery region (expression 26). Since the second-best mechanism under (s_I, s_O) must have a binding constraint (as reflected by (7)), the following holds:

$$EP_{ER}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O)) + EP_{LR}(s_I, s_O, \theta_I^*(s_I, s_O), \theta_I^{**}(s_I, s_O)) = 0 \tag{27}$$

We use a similar process as in the proof for Proposition 4 under assumption that O is the preferred controller (i.e., $PP(s_I, s_O) < 0$ for all $(s_I, s_O) \in S^F$). We denote the second-best mechanism under (s_I, s_O) as $\langle c^{SB}(\theta_I, s_I, s_O), t, s_I, s_O \rangle$. We then construct an alternative mechanism $\langle c(\theta_I, s_I, \underline{s_O}), t', s_I, \underline{s_O} \rangle$. The allocation rule $c(\theta_I, s_I, \underline{s_O})$ is specified as:

$$c(\theta_I, s_I, \underline{s_O}) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^*(s_I, s_O) \\ TH(s_I, \underline{s_O}) & \text{for } \theta_I^*(s_I, s_O) \leq \theta_I \leq \theta_I^{**}(s_I, s_O) \\ 1 & \text{for } \theta_I > \theta_I^{**}(s_I, s_O) \end{cases} \tag{28}$$

Again, it can be shown that the constructed mechanism is implementable, because both expression 25 and 26 are decreasing functions of s_O (which implies constructed mechanism's compatibility constraint is easier to be satisfied). In addition, under this constructed rule, the *ex ante* expected welfare is higher. We then can conclude that the second-best mechanism under $(s_I, \underline{s_O})$ must outperform the second-best under (s_I, s_O) for any s_I and any $s_O > \underline{s_O}$. Using a similar approach, it can be shown that $s_I^{SB} = \underline{s_I}$ if I is the preferred controller (i.e., $PP(s_I, s_O) > 0$ for all $(s_I, s_O) \in S^F$.)

Online Appendix

6.1 Extension: two-sided private information

We consider a more general setup: O 's type is given by $\theta_O = \epsilon\omega + (1 - \epsilon)a$. The type θ_O is consist of two components: a is a constant term commonly known by both partners; ω is privately known by partner O and follows the distribution $F_O(\cdot)$. The parameter ϵ captures the weight of these two components in determining O 's type. It can be interpreted as the importance of O 's private information, as it determines the ranges of I 's types where efficient reallocation outcome is indeterminate without O 's private information. When $\epsilon = 0$, it falls back to our main setup: O does not hold any relevant private information. When $\epsilon = 1$, O 's private information is always pivotal, regardless of I 's private information. When $0 < \epsilon < 1$, O 's private information matters when $\theta_I \in [\underline{\theta}_O, \overline{\theta}_O] = [(1 - \epsilon)a, \epsilon + (1 - \epsilon)a] \subset [0, 1]$.

For convenience, we represent θ_O as following distribution $G(\cdot)$ on $[\underline{\theta}_O, \overline{\theta}_O] = [(1 - \epsilon)a, \epsilon + (1 - \epsilon)a]$, where $G(\cdot)$ is an affine transformation of $F_O(\cdot)$. Throughout this extension, we use $\mathbb{E}[\cdot]$ to denote the expectation over the distribution of both I 's and O 's types, and $\mathbb{E}_i[\cdot]$ to denote the expectation over only the distribution of θ_i .

As in Section 3, we begin by characterizing the possibility condition for efficient reallocation under (s_I, s_O) . Similar to the approach of Ferreira et al. (2015), we construct an efficient direct mechanism where transfers depends on both the identity of the assigned controller and whether O 's private information is relevant given report $\tilde{\theta}_I$. The mechanism proceeds as below:

1. Each partner $i \in \{I, O\}$ pays a participation fee k_i .
2. Each partner i announces his or her type $\tilde{\theta}_i$.
3. Depending on I 's report of $\tilde{\theta}_I$:
 - (a) If $\tilde{\theta}_I \leq \underline{\theta}_O$, O is allocated with control, and is required to transfer $T = (s_I + s_O - 1)\mathbb{E}[v_O(\theta_O, \theta_O)]$ directly to I .
 - (b) If $\tilde{\theta}_I \in (\underline{\theta}_O, \overline{\theta}_O)$, partner i is assigned control if $\tilde{\theta}_i > \tilde{\theta}_j$. There is no direct monetary transfer between two partners. The partner j receives a transfer $t_j = (s_I + s_O - 1)v_i(\tilde{\theta}_i, \tilde{\theta}_i)$ from the mechanism.
 - (c) If $\tilde{\theta}_I \geq \overline{\theta}_O$, I retains the firm and no monetary transfer happens.
4. The partner assigned with control keeps the ownership s_i , and the other partner keeps ownership $1 - s_i$.

The following lemma justifies that we can focus on this constructed mechanism without loss of generality.

Lemma 3 *Given (s_I, s_O) , any efficient mechanism that is incentive compatible should be payoff-equivalent to the mechanism with a pair of (k_I, k_O) .*

Previous literature (Williams (1999); Fieseler et al. (2003); Ferreira et al. (2015)) has established the following revenue equivalence result: for any mechanism that is both incentive compatible and efficient, each partner's interim expected utility is the same across all such mechanisms, differing only by an additive constant. The constructed mechanism is efficient by its design. And it can be shown that both partners are incentive compatible in the mechanism. Consequently, for any other mechanism that is both incentive compatible and efficient, there must exist a pair of (k_I, k_O) making the partners' payoffs in the constructed mechanism equivalent to the payoffs in that mechanism. Therefore, we can focus on the constructed mechanism without loss of generality.

To characterize the condition required by individual rationalities, we need to identify the worst-off type for each partner. We use $\widehat{NE}_I(\theta_I, s_I, s_O)$ to denote the net expected payoff of partner I minus the participation fee k_I in the mechanism. With $\epsilon \in (0, 1)$, though I no longer holds complete information, she can still distinguish three cases based on her private information θ_I : (i) O 's private information is irrelevant to efficient outcome and I loses control; (ii) O 's private information is relevant, and control allocation depends on both partners' report; (iii) O 's private information is irrelevant and I retains control. Then, we have three expressions of $\widehat{NE}_I(\theta_i, s_I, s_O)$, corresponding to each case:

(i) When $\theta_I \leq \underline{\theta}_O$,

$$\widehat{NE}_I(\theta_I, s_I, s_O) = s_I \mathbb{E}_O[v_O(\theta_O, \theta_O)] + (1 - s_O) \mathbb{E}_O[\eta(\theta_I) - \eta(\theta_O)] - r \mathbb{E}_O[v_I(\theta_I, \theta_O)].$$

(ii) When $\theta_I \in (\underline{\theta}_O, \overline{\theta}_O)$,

$$\begin{aligned} \widehat{NE}_I(\theta_I, s_I, s_O) &= s_I \mathbb{E}_O[v_I(\max\{\theta_I, \theta_O\}, \theta_O)] + (1 - s_O) \mathbb{E}_O[\eta(\theta_I) - \eta(\theta_O) \mid \theta_O > \theta_I] \cdot \Pr(\theta_O > \theta_I) \\ &\quad - r \mathbb{E}_O[v_I(\theta_I, \theta_O)]. \end{aligned}$$

(iii) When $\theta_I \geq \overline{\theta}_O$,

$$\widehat{NE}_I(\theta_I, s_I, s_O) = s_I \mathbb{E}_O[v_I(\theta_I, \theta_O)] - r \mathbb{E}_O[v_I(\theta_I, \theta_O)].$$

Given the expressions, the worst-off type for I , represented as θ_I^* , must be within $[\underline{\theta}_O, \bar{\theta}_O]$, and is characterized by:

$$\theta_I^* = G^{-1}\left(\frac{r - (1 - s_O)\eta'(\theta_I^*)}{s_I - (1 - s_O)\eta'(\theta_I^*)}\right) \quad (29)$$

One observation is that once $\epsilon > 0$, the worst-off type varies by (s_I, s_O) , differing from the one-sided private information case. This is because for one-sided private information (i.e., $\epsilon = 0$), I know exactly when she buys or sells shares by efficiency, thus her worst-off type is uniquely determined by the situation where buying or selling is inconsequential at all. In such case, the worst-off type $\theta_I^* = \theta_O = a$, which is not a function of (s_I, s_O) . However, when $\epsilon > 0$, I can not tell exactly the identity of seller (or buyer) when her type is in the region $[\underline{\theta}_O, \bar{\theta}_O]$, thus the worst-off type becomes the type where on average, she is indifferent between being a seller or a buyer. The restructured ownership (s_I, s_O) affects the expected payoff for being a seller or buyer, thereby affecting the worst-off type.

Similarly, we denote the net expected interim payoff of O minus constant fee k_O as $\widehat{NE}_O(\theta_O, s_I, s_O)$. Since O cannot distinguish if his private information is relevant to efficient reallocation outcome, $\widehat{NE}_O(\theta_O, s_I, s_O)$ is captured by the following expression:

$$\begin{aligned} & \widehat{NE}_O(\theta_O, s_I, s_O) \\ &= (s_O \mathbb{E}_I[v_O(\theta_O, \theta_I) \mid \theta_I < \underline{\theta}_O] - T) Pr(\theta_I < \underline{\theta}_O) \\ &+ s_O \mathbb{E}_I \left[v_I(\max\{\theta_O, \theta_I\}, \theta_I) \mid \theta_I \in (\underline{\theta}_O, \bar{\theta}_O) \right] Pr(\theta_I \in (\underline{\theta}_O, \bar{\theta}_O)) \\ &+ (1 - s_I) \mathbb{E}_I[\eta(\theta_O) - \eta(\theta_I) \mid \theta_I \in (\theta_O, \bar{\theta}_O)] Pr(\theta_I \in (\theta_O, \bar{\theta}_O)) \\ &+ (1 - s_I) \mathbb{E}_I[v_I(\theta_I, \theta_O) \mid \theta_I > \bar{\theta}_O] Pr(\theta_I > \bar{\theta}_O) - (1 - r) \mathbb{E}_I[v_I(\theta_I, \theta_O)]. \end{aligned}$$

Thus, the worst-off type for O depends on the interdependence relationship and is captured by:

$$\theta_O^* = \begin{cases} \underline{\theta}_O & \text{if } \eta'(\cdot) \leq 0 \\ \min\{\max\{F^{-1}(\frac{(s_I - r)\eta'(\theta_O^*)}{s_O - (1 - s_I)\eta'(\theta_O^*)}), \underline{\theta}_O\}, \bar{\theta}_O\} & \text{if } \eta'(\cdot) > 0 \end{cases} \quad (30)$$

Similar to I , O 's worst-off type can depend on (s_I, s_O) . However, it is not always the case for two main reasons: (1) the status quo (I retaining control) makes O more likely to regard himself as a buyer; (2) O 's type (including the worst-off type) is bounded by $[\underline{\theta}_O, \bar{\theta}_O]$. Under negative interdependence or independence ($\eta'(\cdot) \leq 0$), O always regards himself more as a buyer (reason (1)), as a result, his worst-off is always $\theta_O^* = \underline{\theta}_O$. Under

positive interdependence ($\eta'(\cdot) > 0$), O 's worst-off type can be either at the boundary or interior.

Efficient reallocation is possible under this more general case if and only if there exists an efficient mechanism satisfying incentive compatibilities, individual rationalities, and the *ex ante* budget balance constraints.²¹ The following proposition characterizes the possibility condition:

Proposition 7 *Given (s_I, s_O) , efficient reallocation is possible if and only if*

$$\begin{aligned} & \widehat{NE}_I(\theta_I^*, s_I, s_O) + \widehat{NE}_O(\theta_O^*, s_I, s_O) \\ & \geq (s_I + s_O - 1)\mathbb{E}[\max\{v_I(\theta_I, \theta_I), v_I(\theta_O, \theta_O)\} \mid \theta_I \in (\underline{\theta}_O, \overline{\theta}_O)]Pr(\theta_I \in (\underline{\theta}_O, \overline{\theta}_O)) \end{aligned} \quad (31)$$

Proof.

For the constructed mechanism to satisfy individual rationalities, each partner must be willing to participate the mechanism compared to status quo at the worst-off type. Since $\widehat{NE}_i(\theta_i^*, s_I, s_O)$ is the net expected payoff at i 's worst-off type minus the participation fee k_i , individual rationalities of both partners would require:

$$\widehat{NE}_I(\theta_I^*, s_I, s_O) + \widehat{NE}_O(\theta_O^*, s_I, s_O) - k_I - k_O \geq 0 \quad (32)$$

The remaining condition is *ex ante* budget balance condition. The mechanism charges k_I and k_O as participation fee. Also, when $\theta_I \in [\theta_O, \overline{\theta}_O]$ (IC is satisfied, so $\tilde{\theta}_I = \theta_I$), it transfers $t_j = (s_I + s_O - 1)v_i(\tilde{\theta}_i, \tilde{\theta}_i)$ to the partner not assigned control. Thus, the budget balance requires:

$$k_I + k_O \geq (s_I + s_O - 1)\mathbb{E}[\max\{v_I(\theta_I, \theta_I), v_I(\theta_O, \theta_O)\} \mid \theta_I \in (\underline{\theta}_O, \overline{\theta}_O)]Pr(\theta_I \in (\underline{\theta}_O, \overline{\theta}_O)) \quad (33)$$

The necessity part is proved once we combine (32) and (33). We can prove sufficiency part by showing that once the inequality holds, it is always possible to find k_I and k_O making all constraints satisfied. ■

We denote the following expression as $EL(s_I, s_O)$:

$$\begin{aligned} & EL(s_I, s_O) \\ & = (s_I + s_O - 1)\mathbb{E}[\max\{v_I(\theta_I, \theta_I), v_I(\theta_O, \theta_O)\} \mid \theta_I \in (\underline{\theta}_O, \overline{\theta}_O)]Pr(\theta_I \in (\underline{\theta}_O, \overline{\theta}_O)) \\ & \quad - \widehat{NE}_I(\theta_I^*, s_I, s_O) - \widehat{NE}_O(\theta_O^*, s_I, s_O) \end{aligned} \quad (34)$$

²¹When *ex ante* budget balance is satisfied, it is always possible to construct transfers that satisfy *ex post* budget balance (see d'Aspremont and Gérard-Varet (1979))

When $EL(s_I, s_O) \leq 0$, efficient reallocation is possible. When $EL(s_I, s_O) > 0$, it captures the subsidy required by efficiency. Again, the optimal (s_I^E, s_O^E) for the first-best analysis is defined as the pair minimizing $EL(s_I, s_O)$.

Lemma 4 $(s_I^E, s_O^E) = \arg \min_{(s_I, s_O) \in S^F} EI_I(s_I, s_O) + EI_O(s_I, s_O)$, where $EI_i(s_I, s_O)$ represents i 's *ex ante* information rent.

This lemma can be proved by showing that $EL(s_I, s_O)$ equals to the sum of two partners' *ex ante* information rents minus net surplus from the efficient reallocation (compared to status quo). Since the net surplus is not a function of (s_I, s_O) , (s_I^E, s_O^E) must be the pair minimizing the sum of information rents by definition. This lemma helps to build up the intuition for the following proposition.

Proposition 8 As $\epsilon \rightarrow 0$, $(s_I^E, s_O^E) \rightarrow (\underline{s}_I, \underline{s}_O)$ when $\eta'(\cdot) > 0$, and $(s_I^E, s_O^E) \rightarrow (\underline{s}_I, 1)$ when $\eta'(\cdot) < 0$.

This proposition can be proved by showing $EL(s_I, s_O)$ (34) converges to the one in Section 3 as $\epsilon \rightarrow 0$. It is also intuitive. When $\epsilon \rightarrow 0$, O 's information rents converge to zero, because O earns information rents only when $\theta_I \in [\underline{\theta}_O, \overline{\theta}_O]$, which is severely limited here. In addition, I 's information rents converge to the value under one-sided private information. This proposition establishes that when O 's private information has limited importance, our main results under one-sided private information still tend to hold.

However, when ϵ becomes larger, the results can diverge from the one-sided private information case in various way. That is because the optimal (s_I^E, s_O^E) have to account for O 's information rents besides I 's information rents, which react to (s_I, s_O) differently and tend be larger as ϵ increases. In addition, the worst-off type can vary with (s_I, s_O) , changing the way each partner's information rents react to (s_I, s_O) as well. As a result, the optimal shares s_i can be dramatically different. As one example, when $\eta(\cdot) \leq 0$, we have $s_O^E = 1$ when $\epsilon = 0$, while $s_O^E = \underline{s}_O$ when $\epsilon = 1$.

6.2 Second-best Mechanism with Nonlinear $\eta(\cdot)$

In Section 4, we restrict attention to valuation function $v_i(\theta_i, \theta_j) = \theta_i + \eta(\theta_j)$, assuming that $\eta(\cdot)$ is linear. We now allow $\eta(\cdot)$ to be convex or concave, and ask how the characterization of the second-best mechanism changes, and whether our results of (s_I^{SB}, s_O^{SB}) still hold given these changes.

When $\eta(\cdot)$ is not linear, one important difference is that the lottery probability corresponding to the worst-off type(s) now depends on the value of θ_I . In other words,

$TH(\theta_I, s_I, s_O)$ is a function of θ_I :

$$TH(\theta_I, s_I, s_O) = \frac{r - (1 - s_O)\eta'(\theta_I)}{s_I - (1 - s_O)\eta'(\theta_I)}$$

Except for this, the non-linearity does not affect the form of optimization problem and its constraint. Thus, the proposition 2 is the same except that $TH(s_I, s_O)$ is replaced by $TH(\theta_I, s_I, s_O)$. To ensure that the second-best mechanism still has a clear and tractable form, we impose the following two assumptions:

Assumption 1'. $(s_I - (1 - s_O)\eta'(\theta_I)) \frac{F_I(\theta_I)}{f_I(\theta_I)}$ is increasing in θ_I , and $(s_I - (1 - s_O)\eta'(\theta_I)) \frac{1 - F_I(\theta_I)}{f_I(\theta_I)}$ is decreasing in θ_I .

Assumption 2'. $\int_0^1 v_O(\theta_O, x) dF_I(x) \geq \int_0^1 v_I(x, \theta_O) dF_I(x)$.

Different from the Assumption 1 in Section 4, Assumption 1' is a modified version of regularity condition. This is required to ensure the monotonicity of the virtual valuations (or I 's interim information rents) when $\eta(\cdot)$ is not linear. The following proposition characterizes the second-best mechanism when $\eta(\cdot)$ is concave.

Proposition 9 *Given Assumptions 1'-2' and $\eta(\cdot)'' \leq 0$, the second-best mechanism follows the control reallocation rule:*

$$c(\theta_I, s_I, s_O) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^* \\ TH(\theta_I, s_I, s_O) & \text{for } \theta_I^* \leq \theta_I \leq \theta_I^{**} \\ 1 & \text{for } \theta_I > \theta_I^{**} \end{cases}$$

where $(\theta_I^*, \theta_I^{**})$ is the smallest interval s.t. $0 < \theta_I^* < \theta_O$ and $1 > \theta_I^{**} > \theta_O$ satisfy

$$\begin{aligned} & \int_0^{\theta_I^*} [s_O v_O(\theta_O, x) - (1 - r)v_I(x, \theta_O) - r v_I(\theta_I^*, \theta_O) + (1 - s_O)v_O(\theta_O, \theta_I^*)] dF_I(x) \\ & + \int_{\theta_I^*}^{\theta_I^{**}} (1 - TH(x, s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) dF_I(x) \\ & + \int_{\theta_I^{**}}^1 (r - s_I)(v_I(x, \theta_O) - v_I(\theta_I^{**}, \theta_O)) dF_I(x) = 0. \end{aligned} \tag{35}$$

and

$$\frac{v_O(\theta_O, \theta_I^*) - v_I(\theta_I^*, \theta_O)}{s_I - (1 - s_O)\eta'(\theta_I^*)} \cdot \frac{f_I(\theta_I^*)}{F_I(\theta_I^*)} = \frac{v_I(\theta_I^{**}, \theta_O) - v_O(\theta_O, \theta_I^{**})}{s_I - (1 - s_O)\eta'(\theta_I^{**})} \frac{f_I(\theta_I^{**})}{1 - F_I(\theta_I^{**})} \tag{36}$$

When $\eta(\cdot)$ is concave, instead of applying a lottery with constant probability for types within $[\theta_I^*, \theta_I^{**}]$, the mechanism applies the lottery rule $c(\theta_I) = TH(\theta_I, s_I, s_O)$ that differs with types, with higher type of I has a higher chance of being assigned with control. However, the two conditions capture the similar idea as in the linear case: the condition (35) ensure that information rents are reduced just enough for second-best mechanism to be implementable, and the condition (36) balances the two ways of reducing information rents—via increasing θ_I^{**} or decreasing θ_I^* , such that the marginal welfare costs for giving up efficiency are the same.

It can be established that the results regarding the optimal (s_I^{SB}, s_O^{SB}) in Section 4 still holds. Specifically, Proposition 4 and 5). That is because though the second-best mechanism is different in format, it still consists of two regions with similar characteristics: (1) efficient allocation region with positive information rents, and (2) lottery region with zero information rents. As a result, both the information rent channel and the lottery loss channel operate in the same direction as in the linear case.

The following proposition characterizes the second-best mechanism under (s_I, s_O) when $\eta(\cdot)$ is convex:

Proposition 10 *Given Assumptions 1'-2', and $\eta''(\cdot) > 0$, the second-best mechanism follows the control reallocation rule:*

$$c(\theta_I, s_I, s_O) = \begin{cases} 0 & \text{for } \theta_I < \theta_I^A \\ TH(\theta_I^*, s_I, s_O) & \text{for } \theta_I^A \leq \theta_I \leq \theta_I^B \\ 1 & \text{for } \theta_I > \theta_I^B \end{cases}$$

where θ_I^* is the unique worst-off type, and (θ_I^A, θ_I^B) is the smallest interval s.t. $0 < \theta_I^A < \theta_I^*$ and $1 > \theta_I^B > \theta_I^*$. The types θ_I^* , θ_I^A and θ_I^B must satisfy:

$$\begin{aligned}
& \int_0^{\theta_I^A} \left[s_O v_O(\theta_O, x) - (1-r)v_I(x, \theta_O) - r v_I(\theta_I^A, \theta_O) + (1-s_O)v_O(\theta_O, \theta_I^A) \right] dF_I(x) \\
& \int_{\theta_I^A}^{\theta_I^*} \left[(1 - TH(\theta_I^*, s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) \right. \\
& \quad \left. + (TH(\theta_I^*, s_I, s_O) - TH(x, s_I, s_O))(s_I - (1-s_O)\eta'(x)) \frac{F_I(x)}{f_I(x)} \right] dF_I(x) \\
& + \int_{\theta_I^*}^{\theta_I^B} \left[(1 - TH(\theta_I^*, s_I, s_O))(v_O(\theta_O, x) - v_I(x, \theta_O)) \right. \\
& \quad \left. + (TH(x, s_I, s_O) - TH(\theta_I^*, s_I, s_O))(s_I - (1-s_O)\eta'(x)) \frac{1 - F_I(x)}{f_I(x)} \right] dF_I(x) \\
& + \int_{\theta_I^B}^1 (s_I - r)(v_I(\theta_I^B, \theta_O) - v_I(x, \theta_O)) dF_I(x) = 0
\end{aligned} \tag{37}$$

and

$$\begin{aligned}
& \frac{v_O(\theta_O, \theta_I^A) - v_I(\theta_I^A, \theta_O)}{s_I - (1-s_O)\eta'(\theta_I^A)} \cdot \frac{f_I(\theta_I^A)}{F_I(\theta_I^A)} = \frac{v_I(\theta_I^B, \theta_O) - v_O(\theta_O, \theta_I^B)}{s_I - (1-s_O)\eta'(\theta_I^B)} \frac{f_I(\theta_I^B)}{1 - F_I(\theta_I^B)} \\
& \quad - \int_{\theta_I^A}^{\theta_I^B} (v_I(\theta_I, \theta_O) - v_O(\theta_O, \theta_I)) dF_I(\theta_I) \\
& = \frac{\int_{\theta_I^A}^{\theta_I^*} (s_I - (1-s_O)\eta'(\theta_I)) \frac{F_I(\theta_I)}{f_I(\theta_I)} dF_I(\theta_I) - \int_{\theta_I^*}^{\theta_I^B} (s_I - (1-s_O)\eta'(\theta_I)) \frac{1-F_I(\theta_I)}{f_I(\theta_I)} dF_I(\theta_I)}{1}
\end{aligned} \tag{38}$$

When $\eta(\cdot)$ is convex, $TH(\theta_I, s_I, s_O)$ is a decreasing function of θ_I . However, the allocation rule $c(\cdot)$ has to be weakly increasing in θ_I to satisfy *ICI*. As a result, it is not feasible to have a range of worst-off types earning zero information rents: the worst-off type has to be unique. To reduce information rents most effectively, the second-best mechanism applies the lottery rule $c(\theta_I) = TH(\theta_I^*, s_I, s_O)$ for a range of types $[\theta_I^A, \theta_I^B]$, where θ_I^* is the unique worst-off type. In this case, the lottery region unavoidably entails positive interim information rents.

As reflected in the above proposition, the second-best mechanism under (s_I, s_O) is characterized by three specific types of I : θ_I^A , θ_I^B and θ_I^* , where the first two are thresholds for lottery region and the last one is the worst-off type. The condition (37) ensures that information rents are reduced just enough to recover implementability. The condition (38) ensures the three specific types are adjusted properly so that both (1) the information rents in the lottery region is reduced in the most efficient way, and (2) the marginal welfare loss (after correction) should be equal for two thresholds θ_I^A and θ_I^B .

In this case, however, our results regarding the optimal (s_I^{SB}, s_O^{SB}) may not hold. The main reason is that all types in the lottery region (except for the worst-off type) earn some

information rents. As a result, the information rent channel not only operates in the efficient allocation region, but also in the lottery region. This makes the direction of information rent channel ambiguous without further specification. Therefore, when $\eta(\cdot)$ is convex, divergence from our results can happen.