Governing Adaptation and Coordination: A Generalization of Multi-agent Delegation

Tan Gan, Ju Hu, and Xi Weng*

June 3, 2020

Abstract

This paper investigates the optimal delegation mechanism in a two-division organization. The uninformed principal wants decisions to be both adapted to local conditions and coordinated with each other, while the privately informed agents only care about adaptation. We explicitly characterize the essentially unique optimal delegation mechanism with a weak continuity restriction in this framework. An agent’s decision in the optimal mechanism falls into three cases: i) full adaptation; ii) state-dependent unilateral coordination; and iii) state-independent joint coordination. This special structure makes the optimal mechanism group strategyproof, and allows us to conduct multiple comparative static analyses with respect to delegated discretion.

Key words: Adaptation, Coordination, Delegation, Dominant Strategy Mechanism Design

JEL: D23, D82, L23, M11

*Gan: Department of Economics, Yale University (email: tan.gan@yale.edu); Hu: National School of Development, Peking University (email: juhu@nsd.pku.edu.cn); Weng: Guanghua School of Management, Peking University (email: wengxi125@gsm.pku.edu.cn). We thank Dirk Bergemann, Tilman Börgers, Marina Halac, Johannes Hörner, Matthew Knudson, and Heng Liu for comments and suggestions that greatly improved the article. We thank the seminar participants at Peking University, Yale and YES in Columbia for helpful comments.
1 Introduction

This paper investigates the optimal delegation mechanism in a two-division organization, which must resolve a trade-off between adaptation and coordination: the more closely activities are synchronized across these two divisions, the less they can be adapted to the local conditions of each division. Such a trade-off poses an organizational challenge due to the existence of information asymmetry and incentive misalignment: the agents (i.e., the division managers) who are privately informed about local conditions care more about adaptation while the uninformed principal (i.e., the boss) cares more about coordination.\(^1\)

Following the seminar work of Holmstrom (1984), many studies have analyzed the optimal delegation mechanism between an uninformed principal and an informed agent with correlated but conflicting interest (e.g., Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2007), and Alonso and Matouschek (2008)). The principal is normally assumed to have perfect commitment power over the decision rule but cannot use monetary transfers. The optimal mechanism in this setting can be implemented by "delegation with constraint". That is, the agent is granted full freedom to make any decision subject to some prescribed constraint specified by the principal. Such implementation has a neat structure that helps to reduce information transmission costs\(^2\) and is believed to demand less commitment power\(^3\), which may explain its popularity among both realistic organizations and academic studies.

However, to incorporate coordination among different divisions into the model, we have to move beyond the one-divisional case. Although various studies have characterized strategic communication equilibria under some specific organizational structures, little is known about the general optimal delegation mechanism in multi-divisional cases. What will the optimal mechanism look like? Will it have some connection with the one-divisional case, and thus enjoy similar facility in implementation? To give a first pass on these questions, this paper studies the optimal dominant strategy incentive compatible (DSIC) mechanism in a setting similar to Alonso, Dessein and Matouschek (2008), with the following features:\(^4\)

- Decision making by the principal involves a trade-off between coordination and adaptation. In particular, two decisions have to be made and the principal must balance the benefit of setting the decisions close to each other against that of setting each decision

---

\(^1\)Examples of this trade-off can be found in Garicano and Rayo (2016).


\(^4\)Difference in settings between these two papers will be discussed in the model setting section.
close to its local condition or “state”.

- Information about the states is dispersed and held by agents who are self-interested, in the sense that they only wish that the specific state of their own division is matched with the corresponding decision, and do not care about the coordination of the whole organization.

- The principal has perfect commitment power over decision rules, but cannot provide monetary transfers. The limited ability to rely on transfers for providing incentives under this context is fully reasoned in the related literature. Compared to Bayesian mechanisms, dominant strategy mechanisms are much more robust informationally. This guarantees the robustness of our results under various information structures, in which cases Bayesian mechanisms are even impossible to define. For instance, both principal and agents may not know others’ beliefs about the unknown states.

Since the agents care only about adaptation loss in their own division, any DSIC direct mechanism could be alternatively interpreted as a contingent delegation mechanism in which the principal delegates a delegation set to each agent contingent only on the other agent’s report. Agents then could choose their most preferred action within the delegation sets. This mechanism can be implemented by “delegation with autonomous constraints”, which is a generation of “delegation with constraints” in the single-agent case. To carry out “delegation with autonomous constraint”, the principal simply delegates the right of making decisions as well as the right of imposing constraints to the agents, i.e., each agent is required to put some prescribed constraint on his counterpart’s decision set, conditioned on the realized state privately observed by himself. Throughout this paper, we will restrict our attention to a decision rule with weak continuity such that each agent’s decision rule is continuous in his reported state. As in the single-agent model, this assumption immediately implies that the delegation set is an interval.

Our major contribution is in revealing the elegant structure of the optimal mechanism, which is entirely shaped by unilateral coordinated boundary function. The unilateral coordinated boundary refers to the optimal contingent delegation scheme where only one agent coordinates unilaterally. It is a contingent delegation interval satisfying that both the lower and upper bounds are increasing functions of the report of the other agent (who does no coordination). In the optimal mechanism, the contingent delegation bounds granted to each agent will be either the unilateral coordinated bounds contingent on the other agent’s report or state-invariant joint coordinated bounds. More concretely, when the reported state from agent B is
very high/low (in the sense that the state is beyond the joint coordinated bound), agent A will receive the constant joint coordinated lower/upper bound. Otherwise, the agent will receive the contingent unilateral coordinated bounds.

The intuition of why the design of the optimal mechanism can be decomposed into two single agent unilateral coordination problem will be fully revealed alongside the proof and discussion which are presented in Sections 4 and 5. Briefly, when State B is moderate, Agent B’s most preferred action will not be limited by the contingent delegation bounds induced by the state reported from Agent A, and hence agent B is free to take his mostly preferred action. Consequently, when State B is moderate, the optimal mechanism will seek coordination by solely putting constraints on Agent A’s action. As a result, Agent A’s contingent delegation interval is determined by the unilateral coordinated contingent bounds. On the contrary, when State B is extremely high/low, Agent B’s action will be constrained by the contingent delegation bounds induced by State A. Then, when State A is moderate, agent A is free to take his most preferred action as well. However, when State A is at extreme opposites to State B (i.e., State A is extremely high while State B is extremely low or vice versa), the optimal mechanism will require both agents to coordinate jointly by choosing the state-invariant joint coordinated bounds, which generalize the state-invariant decision caps/floors in the single-agent delegation model.

Although we cannot prove that the restricted optimal mechanism is global maximizer in the unrestricted problem, we will show that the above mechanism is indeed locally optimal with respect to more general DSIC decision rules. That is, given the other agent’s weakly continuous decision rule characterized by the above mechanism, one agent’s decision rule is optimal among all of his DSIC decision rules. Moreover, our mechanism has the nice property of group strategyproofness. The two agents cannot benefit from misreporting their states jointly under the mechanism.

Based on the structure of the optimal mechanism, multiple comparative analysis results will extend our understanding of how the delegated discretion is affected by the modeling parameters. In single-agent delegation models, the key comparative static analysis result is that the agent will enjoy more discretion when his information is more valuable, or when the conflict of interests is smaller (Koessler and Martimort (2012)). This principle is also inherited in our multi-agent setting. In fact, when the importance of project A relative to coordination decreases with project B’s relative importance unchanged, agent A will receive less delegated discretion: he will face more restrictive contingent delegation bounds and take his most preferred action less frequently. In contrary, he will receive more discretion as the
relative importance of his counterpart’s project decreases.

The delegated discretion is also influenced by the principal’s belief of state distributions. For example, if the distribution of state A has a fatter lower/upper tail, agent A will enjoy more discretion as the lower/upper delegation bound becomes more slack. Agent B on the contrary, will suffer from tighter upper/lower delegation bound and hence less delegated discretion.

**Related Literature.** The branch of literature analyzing delegation via the mechanism design approach stems from the seminal paper of Holmstrom (1984) who limits most of his discussions to interval delegation. The relaxation of such a limitation is discussed in Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2007), and Alonso and Matouschek (2008) who give various conditions guaranteeing the optimality of interval delegation. Unfortunately, due to the technical issue which will be discussed in Section 6, we cannot borrow their approaches to defend the continuity restriction needed in this paper. Among other outstanding research, two existing papers, Martimort and Semenov (2008) and Koessler and Martimort (2012), discuss some multi-dimensional extensions. Koessler and Martimort (2012) add the number of decision from one to two in their model. Their major goal is to highlight the possibility to utilize distortion of decisions for better information revelation. Thus the utility which is assumed to be simple summation of gains from each decision, has no connection with coordination. Methodologically, the extra decision, which can be somehow viewed as a tool similar to monetary transfer, helps the authors to solve the mechanism using calculus of variations after an appropriate substitution. Our approach is entirely different as the mechanism in our model is not sufficiently smooth. In another paper, Martimort and Semenov (2008) analyze a model with two agents, two states and one decision, under a political background where different interest groups are trying to influence the policy making process. Since only one decision is included, the story is different from ours. The methodology between the two papers is also different, except that we both restrict our attentions to dominant strategy mechanisms with weak continuity.

Another related literature studies the strategic communication between the principal and the agents under different organizational structures such as centralization and decentralization (e.g., Alonso, Dessein and Matouschek (2008), Rantakari (2008), and Li and Weng (2017)) to compare the organizational performances of these specific structures. This paper departs from the previous literature by investigating the optimal dominant strategy incentive compatible communication mechanism in this framework.\(^5\)

\(^5\)The concept of dominant strategy mechanism coincides with ex-post mechanism in our setting since it is a
An important paper from this literary is Alonso, Dessein and Matouschek (2008) which motivates the framework of our model setting. The most significant difference between the two papers is that they analyze the communication problem by comparing two specific mechanisms, while we use mechanism design approach under a simplified setting. On the other hand, our result, showing that “delegation with autonomous delegation” as a decentralized mechanism can robustly implement the optimal mechanism, is a supplement of their conclusion that decentralization may dominate centralization.

2 Model Setting

There are three players: two agents ($A_1$, $A_2$) and a principal ($P$). As in Alonso, Dessein and Matouschek (2008), we interpret the principal as the headquarters of an organization while the two agents as the division managers.

Information. Agent $A_i$ has private information about his ideal point $s_i$ (or state) of his project. The uninformed principal subjectively believes that $s_1$ and $s_2$ are independent random variables distributed on the interval $[0, 1]$, with strictly positive continuous density functions $f_1(s_1)$ and $f_2(s_2)$. Since we focus on dominant strategy mechanism design, we do not need any more assumptions, such as agents’ knowledge of each other’s state distribution or common prior on state distributions.

Decisions and Preferences. After observing the ideal points, the agents send reports to the principal simultaneously. The principal will make two decisions $a_1, a_2 \in [0, 1]$ after receiving the agents’ reports. Agent $A_i$ only cares the adaptation loss in his division, and has single peaked preference over decision $a_i$ with his ideal point to be $s_i$. Equivalently, if $U_A^i(a_i, s_i)$ represents his preference, then for a fixed $s_i$, $U_A^i$ is maximized at $a_i = s_i$, and for $(a_i - s_i)(a'_i - s_i) \geq 0$, $U_A^i(a_i, s_i) \geq U_A^i(a'_i, s_i)$ if and only if $|a_i - s_i| \leq |a'_i - s_i|$.

The principal, unlike the agents, cares both adaptation loss and coordination loss. Following the literature, we assume the principal’s utility function $U_P(a_1, a_2, s_1, s_2)$ to be quadratic:

$$U_P(a_1, a_2, s_1, s_2) = -\lambda_1(a_1 - s_1)^2 - \lambda_2(a_2 - s_2)^2 - (a_1 - a_2)^2,$$

where $(a_i - s_i)^2$ reflects the adaptation loss and $(a_1 - a_2)^2$ reflects the coordination loss. Notice that in this utility function, the weight of coordination loss is normalized to 1, and $\lambda_i$ is a static private-value model.
non-negative parameter describing the relative importance of the adaptation loss in division \( i \) to the principal.

In the models of Alonso, Dessein and Matouschek (2008), Rantakari (2008), and Li and Weng (2017), agents also care about the coordination loss but with a weight different from the principal’s. Different from these models, we simply assume that agents only care about adaptation loss. This simplifies our analysis, and also allows us to make more general assumptions on the agents’ preferences instead of the quadratic utility functions used in Alonso, Dessein and Matouschek (2008), Rantakari (2008), and Li and Weng (2017).

**Mechanisms.** The principal can commit to a deterministic decision rule without monetary transfers. Invoking the revelation principle, we can represent these mechanisms as two bivariate functions \( a_1(s_1, s_2) \) and \( a_2(s_1, s_2) \) from the state space to the action space. Basically, the agents simultaneously report their states to the principal, and based on the two reports, the principal chooses decision \( a_i(s_1, s_2) \) for division \( i \). We restrict our attention to a dominant strategy incentive compatible (DSIC) and hence can write down the optimization problem as:

\[
\max_{a_1(s_1, s_2), a_2(s_1, s_2)} - \int_0^1 \int_0^1 \left[ \lambda_1 (a_1 - s_1)^2 + \lambda_2 (a_2 - s_2)^2 + (a_1 - a_2)^2 \right] dF_1(s_1) dF_2(s_2)
\]

\[
\text{s.t. } U_i^A(a_i(s_i, s_{-i}), s_i) \geq U_i^A(a_i(\hat{s}_i, s_{-i}), s_i) \quad \forall s_i, s_{-i}, \hat{s}_i \quad (IC)
\]

### 3 Preliminary Analysis

This section gives a preliminary analysis of the optimal mechanism. First, we show that the optimal mechanism takes the form of “delegation with autonomous constraint”, where each agent is granted a delegation interval contingent on the other’s report. Hence, the characterization of the optimal mechanism is equivalent to solving the delegation boundaries. Second, we establish the existence of the optimal dominant strategy mechanism with a weak continuity restriction based on the concept of “one-sided optimizer”.

#### 3.1 Delegation with Autonomous Constraints

**Contingent Delegation Interpretation of Direct Mechanism.** Recall that in the single-agent delegation problem, the DSIC decision rules can be reinterpreted as if the principal gave the agent a set of actions (commonly referred to as delegation set), from which to choose.\(^6\)

\(^6\)See Holmstrom (1984), for instance.
Under the self-interested assumption that

\[ U_A^i(a, s) = U_A^i(a_i, s_i) \]  

we could similarly describe the DSIC decision rule in the direct mechanism using the language of the delegation set, instead of specifying it as a two-dimensional function. This interpretation helps to provide a nicer characterization of DSIC decision rules and to shape the optimal mechanism in the following analysis. For a given DSIC mechanism \((a_1, a_2)\), denote

\[ D_i(s_{-i}) = \{a_i(s_i, s_{-i}) | \forall s_i \in [0, 1] \subseteq [0, 1] \}. \]

According to \((IC)\) condition, \(a_i(s_i, s_{-i}) \in \arg\max_{s \in D(s_{-i})} U_A^i(a, s) = \arg\max_{s \in D(s_{-i})} U_A^i(a_i, s_i).\)

Using this formula, we could recover \(a_i(s_i, s_{-i})\) from \(D_i(s_{-i})\). Similar to the single-agent delegation problem, we could interpret the set \(D_i(s_{-i})\) as the delegation set delegated to agent \(i\). However, differently from that problem, the delegation set \(D_i(s_{-i})\) is contingent on the other agent’s report \(s_{-i}\). In other words, the principal determines the delegate set \(D_i\) to each agent \(i\) only after receiving the report \(s_{-i}\), and then the agents are allowed to choose their most preferred action \(a_i \in D_i(s_{-i})\). Below we provide an example of contingent delegation interval, which is the focus of our subsequent analysis.

**Example 3.1.** Consider the case where \(D_i(s_{-i}) = [\phi_i(s_{-i}), \psi_i(s_{-i})]\). After the reports \((s_1, s_2)\), agent \(i\) receives delegation set \([\phi_i(s_{-i}), \psi_i(s_{-i})]\), from which he could choose his most preferred action. Thus the corresponding decision rules are:

\[
    a_i(s_i, s_{-i}) = \begin{cases} 
    \phi_i(s_{-i}), & \text{if } s_i \in [0, \phi_i(s_{-i})]; \\
    s_i, & \text{if } s_i \in [\phi_i(s_{-i}), \psi_i(s_{-i})]; \\
    \psi_i(s_{-i}), & \text{if } s_i \in [\psi_i(s_{-i}), 1]. 
    \end{cases} \]  

The above mechanism is illustrated by Figure 1. The solid curves represent functions \(\phi_i(s_{-i})\) and \(\psi_i(s_{-i})\), while each dotted line represents a contingent delegation interval. Each agent \(A_i\) will receive a contingent delegation interval for each different report \(s_{-i}\) from the other agent, and thus we can easily find out the corresponding decision rules for \(A_1\) and \(A_2\).

The following proposition shows that we can additionally require \(D_i(s_{-i})\) to be a closed set.

**Proposition 3.2.** A decision rule \((a_1, a_2)\) is DSIC if and only if there exist two collections of...
sets \( \{D_1(s_2)\}, \{D_2(s_1)\} \), such that:

- For each \( s_{-i} \), \( D_i(s_{-i}) \subset [0, 1] \) is closed, this is the contingent delegation set received by agent \( i \), if agent \( \rightarrow i \) report \( s_{-i} \).

- \( a_i(s_i, s_{-i}) \in \arg\max_{a \in D(s_{-i})} U_A(a, s_i) \), i.e. agent chooses his best action within the received contingent delegation set.

The intuition of the above result is quite straightforward. DSIC requires that reporting the true state is the optimal choice when the agent knows the other state and therefore knows the decision rule induced by his report perfectly. Since the agent only cares about the action of his own project and does not care the other project, choosing the optimal report is the same as choosing the optimal action within the possible range imposed by the agent, and the principal is free to impose any constraint she likes on the other agent.

As a direct corollary of the delegation structure, the agent’s action will either reach his most preferred point or be constrained at the boundary of the delegation set. Consequently, the coordination mostly will either be deficient or excessive. The design problem is to seek a balance between these two situations.

**Delegation with Autonomous Constraints.** The above mechanism has a nice property of “autonomous constraints” in the sense that the principal can implement the mechanism by requiring the agents to impose constraints on each other’s action space, and ask each agent to choose within the constraints imposed by the other agent. In doing so, the principal delegates the right of making decisions as well as the right of imposing constraints to the agents, and agents act autonomously.
As an extension of simple delegation, delegation with autonomous constraints might inherit some good properties. First, such mechanism might require less commitment power. As Alonso, Dessein and Matouschek (2008) have argued, decision makers in realistic organizations are often unable to commit to make their decisions dependent on the information they receive. The only formal mechanism they can commit to is the ex ante allocation of rights, including the right of decision and the right of constraint. Second, the mechanism might have lower informational costs. Vertical information transmission through authority levels can sometimes be too costly. By limiting communication to horizontal divisions, delegation with autonomous constraints might alleviate this problem.

The trustful execution of autonomous constraints can be guaranteed by the self-interested assumption which make agents indifferent to each other’s actions. Such indifference in general is a knife-edged one, since any degree of conspiracy could potentially break it. However, as we will show later, the structure of the optimal mechanism gives rise to group strategyproofness, which is not a property of dominant incentive compatible mechanism per se. The principal can therefore protect the implementation of the optimal mechanism from conspiracy by banning all the joint action profiles outside the range of the mechanism.

3.2 The Existence of the Optimal Mechanism

To simplify the analysis, we restrict our attention to mechanisms \((a_1, a_2)\), where \(a_i\) is continuous with respect to \(s_i\). Equivalently, we are restricting our attention to delegation schemes in which all contingent delegation sets are intervals. Similar assumptions are common in the single-agent mechanism design literature such as Holmstrom (1984) and Armstrong (1995), as well as in the multi-agent delegation mechanism design literature such as Martimort and Semenov (2008).

Under the above restriction on feasible mechanisms, for any DSIC decision rule \((a_1, a_2)\), there exist functions \(\phi_i(s_{-i})\) and \(\psi_i(s_{-i})\) such that for all \(s_{-i} \in [0, 1]\),

\[
0 \leq \phi_i(s_{-i}) \leq \psi_i(s_{-i}) \leq 1
\]  

(3)

and the decision rules \((a_1, a_2)\) take the form in Equation (2). In other words, the boundary functions \(\phi_i\) and \(\psi_i\) as functions of \(s_{-i} \in [0, 1]\) represent the bounds of contingent delegation interval granted to agent \(i\). We will use boundary function vectors \((\phi_1, \psi_1, \phi_2, \psi_2)\) to represent \((a_1, a_2)\) when convenient. So the optimal design problem is to select the best boundary functions from any boundary functions satisfying Condition (3).
Monotone Boundary Function We will first show that, to solve the optimal mechanism, it is without loss of generality to focus on increasing boundary functions:

**Theorem 3.3.** For any feasible (incentive compatible) mechanism \((\phi_1, \psi_1, \phi_2, \psi_2)\), there is a mechanism \((\bar{\phi}_1, \bar{\psi}_1, \bar{\phi}_2, \bar{\psi}_2)\) whose boundary functions are all non-decreasing, and yields weakly higher ex ante payoff to the principal.

The intuition of the above theorem is straightforward. As \(s_i\) increases, agent \(A_i\) would like to choose a higher action \(a_i\) as he only cares about adaptation loss. To reduce the coordination loss, the principal would naturally respond by increasing the delegation boundaries \((\phi_i, \psi_i)\) such that \(a_i\) increases as well for every \(s_i\).

The proof of the theorem, however, is quite involved. To prove the theorem, we first introduce the following definition of “one-sided optimizer”, which is an important tool throughout the paper.

**Definition 3.4.** Denote \(y_{s_i}(p)\) as an arbitrary function from \([0, 1]\) to \([0, 1]\). The pair \((c^*, d^*)\), indicating the boundaries of a delegation interval, is called a one-sided optimizer of function \(y_{s_i}(p)\) for \(i\), if:

\[
(c^*, d^*) \in \arg\min_{\{c, d: 0 \leq c \leq d \leq 1\}} \int_0^1 \left[ (\bar{a}_i - y_{s_i}(s_i))^2 + \lambda_i (\bar{a}_i - s_i)^2 \right] dF_i(s_i)
\]

where

\[
\bar{a}_i = \bar{a}_i(c, d, s_i) = \begin{cases} 
  c, & \text{if } s_i \in [0, c] \\
  s_i, & \text{if } s_i \in [c, d] \\
  d, & \text{if } s_i \in [d, 1]
\end{cases}
\]

is called as the action function with delegation interval \([c, d]\).

In the above definition, \(F_i\) refers to the distribution of agent \(A_i\)’s ideal point and \(\lambda_i\) refers to the relative importance of agent \(A_i\)’s division. Notice that Equation (4) immediately implies that a one-sided optimizer \((c^*, d^*)\) maximizes

\[
-\int_0^1 \left[ (\bar{a}_i - y_{s_i}(s_i))^2 + \lambda_{-i}(y_{s_i}(s_i) - s_{-i})^2 + \lambda_i (\bar{a}_i - s_i)^2 \right] dF_i(s_i),
\]

which is the principal’s expected payoff for a fixed \(s_{-i}\) and fixed \(a_{-i} = y_{s_i}\).

**Remark 3.4.1.** Since \(\{(c, d) | 0 \leq c \leq d \leq 1\}\) is a compact set in Euclidian space, and the mapping from \((c, d)\) to \(\int_0^1 (\bar{a}_i - a_{-i})^2 + \lambda_i (\bar{a}_i - s_i)^2 dF_i(s_i)\) is obviously continuous, the arg min in definition 3.4 is well defined but may not be unique.
**Definition 3.5.** The pair of boundary functions \((\phi_i, \psi_i)\) is a one-sided optimizer of \((\phi_{-i}, \psi_{-i})\) for \(i\), if for any fixed \(s_{-i} \in [0,1]\), the pair \((\phi_i(s_{-i}), \psi_i(s_{-i}))\) is a one-sided optimizer of \(a_{-i}(s_i, s_{-i})\) (which is a function of \(s_i\)) for \(i\), where \(a_{-i}(s_i, s_{-i})\) is defined by Equation (2) using \((\phi_{-i}, \psi_{-i})\).

Literally, the pair of boundary functions \((\phi_i, \psi_i)\) is a one-sided optimizer of \((\phi_{-i}, \psi_{-i})\) for \(i\) if it maximizes the principal’s expected payoff given \((\phi_{-i}, \psi_{-i})\). This is because

\[
\min_{\phi_i \leq \psi_i} \int_0^1 \int_0^1 [(a_i - a_{-i})^2 + \lambda_i(a_i - s_i)^2 + \lambda_{-i}(a_{-i} - s_{-i})^2] dF_i(s_i) dF_{-i}(s_{-i})
\]

\[
= \mathbb{E} \lambda_{-i}(a_{-i} - s_{-i})^2 + \min_{\phi_i \leq \psi_i} \int_0^1 \int_0^1 [(a_i - a_{-i})^2 + \lambda_i(a_i - s_i)^2] dF_i(s_i) dF_{-i}(s_{-i})
\]

\[
= \mathbb{E} \lambda_{-i}(a_{-i} - s_{-i})^2 + \int_0^1 \left\{ \min_{\phi_i \leq \psi_i} \int_0^1 [(a_i - a_{-i})^2 + \lambda_i(a_i - s_i)^2] dF_i(s_i) \right\} dF_{-i}(s_{-i})
\]

Therefore, as long as the pair of boundary functions \((\phi_i, \psi_i)\) is a one-sided optimizer of \((\phi_{-i}, \psi_{-i})\) for every \(i\) and every \(s_{-i} \in [0,1]\), the whole contingent delegation scheme is optimal.

Theorem 3.3 is proved based on two propositions. The first proposition says that although the one-sided optimizer of a given function \(y(s_i)\) might not be unique, the set of one-sided optimizers luckily has a nice lattice structure.

**Proposition 3.6.** The set of one-sided optimizers of a given function \(y(s_i)\) consists of a compact sub-lattice of \(\mathbb{R}^2\), thus it is complete. In particular, the minimal and maximal one-sided optimizer exist.

The proof of Proposition 3.6 is based on the swapping property of one-sided optimizers. That is, suppose \((c_1, d_1)\) and \((c_2, d_2)\) are one-sided optimizers of a given function \(y(s_i)\), and \([c_1, d_1] \cap [c_2, d_2] \neq \emptyset\). Then both \([c_1, d_2]\) and \([c_2, d_1]\) are one-sided optimizers as well.

The second proposition establishes a monotone property of the one-sided optimizers. Recall that when calculating the one-sided optimizer of an arbitrary function \(y(s_i)\), we are searching for the optimal bounds of agent \(A_i\)’s delegation interval, in the hope of matching \(A_i\)’s action to both \(s_i\) and \(y(s_i)\). It is therefore natural to conjecture that if the target function \(y(s_i)\) increases, the corresponding optimal delegation interval shall shift upwards as well. To prove this comparative result, we follow the approach of Milgrom and Shannon (1994) by verifying that the integration \(\int_0^1 [(\bar{a}_i - y(s_i))^2 + \lambda_i(\bar{a}_i - s_i)^2] dF_i(s_i)\), as a function of \((c, d, y)\), satisfies the single crossing property and is also quasi-supermodular in \((c, d)\). Therefore, we can directly apply Theorem 4 in Milgrom and Shannon (1994) and obtain:

**Proposition 3.7 (Monotone Property).** The set of one-sided optimizers of \(y\) is increasing in the strong set order in \(y\).
Formally, suppose \( y(s_i) \leq \bar{y}(s_i) \) are two functions from \([0, 1]\) to \([0, 1]\). \((c, d)\) and \((\bar{c}, \bar{d})\) are one-sided optimizers of \( y \) and \( \bar{y} \), respectively. Then \((c, d) \wedge (\bar{c}, \bar{d})\) is a one-sided optimizer of \( y \), and \((c, d) \vee (\bar{c}, \bar{d})\) is a one-sided optimizer of \( \bar{y} \).

As a direct corollary, we could get a “real” monotone property once we focus on the maximal optimizer or minimal optimizer:

**Corollary 3.8.** The maximal (minimal) one-sided optimizer of \( y \) is increasing in \( y \).

Now the proof of Theorem 3.3 is very straightforward.

**Proof of Theorem 3.3**

**Proof.** For any DSIC mechanism \((\phi_1, \psi_1, \phi_2, \psi_2)\), let \((\bar{\phi}_1, \bar{\psi}_1)\) be the maximal one-sided optimizer of \((\phi_2, \psi_2)\) for agent \( A_1 \), and let \((\bar{\phi}_2, \bar{\psi}_2)\) be the maximal one-sided optimizer of \((\bar{\phi}_1, \bar{\psi}_1)\) for agent \( A_2 \). Then clearly the principal’s ex ante payoff from mechanism \((\bar{\phi}_1, \bar{\psi}_1, \bar{\phi}_2, \bar{\psi}_2)\) will be no less than \((\phi_1, \psi_1, \phi_2, \psi_2)\).

From incentive compatibility we know that \( a_2(s_1, s_2) \) is increasing with respect to \( s_2 \), therefore for any \( \bar{s}_2 \geq s_2 \), \( a_2(s_1, \bar{s}_2) \geq a_2(s_1, s_2) \ \forall s_1 \). By definition \((\bar{\phi}_1(\bar{s}_2), \bar{\psi}_1(\bar{s}_2))\) and \((\bar{\phi}_2(\bar{s}_2), \bar{\psi}_2(\bar{s}_2))\) are maximal one-sided optimizers of one dimensional function \( a_2(s_1, s_2) \) and \( a_2(s_1, \bar{s}_2) \) respectively. Then according to corollary 3.8, we know \( \bar{\phi}_1(\bar{s}_2) \geq \bar{\phi}_1(s_2) \) and \( \bar{\psi}_1(\bar{s}_2) \geq \bar{\psi}_1(s_2) \), i.e. \( \bar{\phi}_i \) and \( \bar{\psi}_i \) are non-decreasing.

**The Existence of the Optimal Mechanism** With the help of Theorem 3.3, we could focus on the set of mechanisms with monotone boundary functions, \( M \equiv \{(\phi_1, \psi_1, \phi_2, \psi_2) | \phi_i, \psi_i \text{ are non-decreasing, and } 0 \leq \phi_i \leq \psi_i \leq 1\} \). Using some lengthy but standard technique, we can establish the existence of the optimal mechanism:

**Proposition 3.9.** There exists an optimal mechanism \((\bar{\phi}^*_1, \bar{\psi}^*_1, \bar{\phi}^*_2, \bar{\psi}^*_2)\) in \( M \).

**4 Solving the Optimal Mechanism**

Although Proposition 3.9 establishes the existence, it is in general difficult to explicitly solve for the optimal mechanism. To obtain a further characterization, we henceforth introduce the following two assumptions on the distributions:

**Assumption 4.1.** Both \( \log F_i \) and \( \log(1 - F_i) \) are strictly concave. Equivalently, \( f_i / F_i \) is strictly decreasing with respect to \( s_i \), while \( f_i / (1 - F_i) \) is strictly increasing.
Assumption 4.2. \( f' \geq 0 \) and \( f'/f \) is decreasing.

Obviously, Assumption 4.2 is a strengthening of Assumption 4.1, and density functions with the form of \( x^k \) satisfy both assumptions. This class of density functions includes the uniform distribution, which is widely used in the related literature.

Based on Assumption 4.1, we can explicitly characterize the essentially unique optimal mechanism.\(^7\) This section proceeds as follows. First, we derive properties satisfied by the optimal mechanism under the restriction of contingent delegation intervals. Second, using the derived properties, we fully characterize the unique “restricted” optimal mechanism. Finally, we show that the “restricted” optimal mechanism is indeed locally optimal with respect to more general DSIC decision rules if Assumption 4.2 is satisfied. Readers who are not interested in technical details may jump directly to our main result, Theorem 4.6 and the intuitive discussions in Section 4.3.

4.1 Analysis of the Optimal Mechanism

For every optimal mechanism \( (\phi_1, \psi_1, \phi_2, \psi_2) \) in \( M \), the previous section shows that for almost every \( s_{-i}, (\phi_i(s_{-i}), \psi(s_{-i})) \) is an one-sided optimizer of \( a_{-i}(s_i, s_{-i}) \), as a function of \( s_i \) induced by \( (\phi_{-i}, \psi_{-i}) \). In this subsection, we will figure out more properties satisfied by the essentially unique mechanism based on this property. To begin with, we formally introduce a new concept, unilateral coordinated boundary function, which eventually shapes the optimal mechanism. The unilateral coordinated boundary function is the optimal contingent delegation bounds for one agent when the other agent is free to act with no coordination concern, i.e., \( a_{-i}(s_i, s_{-i}) = s_{-i} \) holds for every \( s_i \).

Lemma 4.3. Assume Assumption 4.1 holds. For every \( s_{-i} \), the one-sided optimizer \( (c^i(s_{-i}), d^i(s_{-i})) \) on a constant function \( s_{-i} \) for \( i \) is the unique pair satisfying the following equations:

\[
\begin{align*}
    c^i(s_{-i}) &= s_{-i} - \lambda_i \int_0^{c^i(s_{-i})} \frac{F_i(s)}{F'_i(c^i(s_{-i}))} ds_i; \\
    d^i(s_{-i}) &= s_{-i} + \lambda_i \int_{d^i(s_{-i})}^1 \frac{1 - F_i(s)}{1 - F'_i(d^i(s_{-i}))} ds_i;
\end{align*}
\]

\(^7\)Of course, any modification of the optimal mechanism on a zero-measure set will induce a new optimal mechanism, but its boundary functions need not be monotone.
or equivalently,

\[(\lambda_i + 1)c_i^*(s_{-i}) = s_{-i} + \lambda_i \mathbb{E}(S_i|S_i \leq c_i^*(s_{-i}));\]  

(11)

\[(\lambda_i + 1)d_i^*(s_{-i}) = s_{-i} + \lambda_i \mathbb{E}(S_i|S_i \geq d_i^*(s_{-i})).\]  

(12)

Moreover, both functions \(c_i^*(s_{-i})\) and \(d_i^*(s_{-i})\) are strictly increasing and Lipschitz continuous with parameter 1. We call these two functions the unilateral coordinated boundary functions of \(i\) given a constant \(s_{-i}\).\(^8\)

The logic behind the construction of the unilateral coordinated delegation interval is that the other agent reaches full adaptation. Of course, in the optimal mechanism, the other agent never reaches full adaptation. Yet still each agent will frequently receive the unilateral coordinated delegation interval as if the other reached full adaptation. The reason for this result comes from the following local determining property:

**Proposition 4.4** (Local-determining Property). Assume \(y_1(s_i)\) and \(y_2(s_i)\) are two functions mapping from \([0, 1]\) to \([0, 1]\). Denote \(D_1\) and \(D_2\) as the sets of one-sided optimizers of \(y_1(s_i)\) and \(y_2(s_i)\), respectively.

If \(\exists c_0\) s.t. \(y_1(s_i) = y_2(s_i) \forall s_i \leq c_0\), and \(\forall (c, d) \in D_1 \cup D_2\), we have \(c \leq c_0, d \geq c_0\). Then \(\{c|(c, d) \in D_1\} = \{c|(c, d) \in D_2\}\).

Similarly, if \(\exists d_0\) s.t. \(y_1(s_i) = y_2(s_i) \forall s_i \geq d_0\), and \(\forall (c, d) \in D_1 \cup D_2\), we have \(c \leq d_0, d \geq d_0\). Then \(\{d|(c, d) \in D_1\} = \{d|(c, d) \in D_2\}\).

The local determining property states that, if two functions are the same on the “relevant” interval, then the lower (upper) bound of their one-sided optimizers shall be the same, no matter how different these two functions are outside the interval. In the optimal mechanism, even though full adaptation is not reached and hence \(a_i\) is different from \(s_i\), the optimal delegation intervals often have the same lower (upper) bound as the unilateral coordinated interval as long as \(a_i\) is the same as \(s_i\) on the mostly “relevant” interval.

With the help of the local determining property and the monotone property, we can give the first estimation of the optimal mechanism.

**Proposition 4.5** (Diagonal-separation Property). Let \((\phi_1^*, \psi_1^*, \phi_2^*, \psi_2^*)\) be an optimal mechanism in \(M\). Then \(\phi_1^*(s_{-i}) \leq s_{-i}\), and \(\psi_1^*(s_{-i}) \geq s_{-i}\).

The intuition behind the above property is also clear. Recall that the conflict of the interests between the agents and the principal will gradually vanish as state pairs approach the

\(^{8}\)From now on we will use \(c^*\) and \(d^*\) to denote the unilateral coordinated boundary functions.
diagonal line, i.e., $s_1 = s_2$. Thus it is natural to conjecture that the optimal mechanism would only push the joint actions $a_1$ and $a_2$ towards the diagonal line, but would never exceed it.

The diagonal-separation property formalizes the intuition mentioned in the introduction that “when state B is moderately low/high, agent B’s most preferred action will not be ruled out by the contingent delegation bound induced by a low/high state reported from agent A”. Then using the local determining property, we know that “consequently, when state B is moderately low/high, the optimal mechanism will seek coordination by putting contingent unilateral coordinated lower/upper bound solely on agent A”. Finally, we use the monotone property to extend the range of “moderate states” to be large enough to shape the whole mechanism.

4.2 Main Results

With the preparations of the above subsection, we can derive the optimal mechanism presented below:

**Theorem 4.6.** Assume Assumption 4.1 holds. The unique optimal mechanism $(\phi^*_1, \psi^*_1, \phi^*_2, \psi^*_2)$ in $M$ takes the following form:

$$
\phi^*_i(s_{-i}) = \begin{cases} 
    c^*_i(s_{-i}) & \text{if } s_{-i} \leq \bar{x}_{-i}; \\
    c^*_i(\bar{x}_{-i}) & \text{if } s_{-i} \geq \bar{x}_{-i},
\end{cases}
$$

$$
\psi^*_i(s_{-i}) = \begin{cases} 
    d^*_i(s_{-i}) & \text{if } s_{-i} \leq \bar{x}_{-i}; \\
    d^*_i(\bar{x}_{-i}) & \text{if } s_{-i} \geq \bar{x}_{-i},
\end{cases}
$$

where $(\bar{x}_{-i}, x_i)$ is the unique pair satisfying:

$$
x_i = c^*_i(\bar{x}_{-i}); \quad \bar{x}_{-i} = d^*_i(x_i). \quad (13)
$$

To get a better understanding about the calculation process, the whole structure, and the intuition for the optimal mechanism, we will present a simple numerical example where both states are uniformly distributed, and $\lambda_i = 1/\lambda$.

We use the following algorithm to explicitly solve the optimal mechanism:

1. Calculate unilateral coordinated boundary functions with respect to the distributions according to Equation (9). The results of this example are:

$$
c^*_i(s_{-i}) = \frac{2\lambda}{2\lambda + 1} s_{-i} \quad ; \quad d^*_i(s_{-i}) = \frac{2\lambda}{2\lambda + 1} s_{-i} + \frac{1}{2\lambda + 1}.
$$
2. Find the intersection points of the unilateral coordinated boundary functions according to Equation (13). The corresponding results of this example are:

\[(\bar{x}_1, \bar{x}_2) = (\bar{x}_2, \bar{x}_1) = \left( \frac{2\lambda + 1}{4\lambda + 1}, \frac{2\lambda}{4\lambda + 1} \right).\]

3. Based on \(x_i\) and \(\bar{x}_i\), determine the optimal mechanism according to Theorem 4.6.

Theorem 4.6 fully characterizes the unique optimal mechanism in \(M\), i.e., both agents receive a contingent delegation interval. However, it is unclear whether it is indeed optimal for the principal to use contingent delegation intervals instead of more general contingent delegation sets. The next proposition partially addresses this concern by showing that if we strengthen Assumption 4.1 by Assumption 4.2, then the optimal mechanism characterized in Theorem 4.6 is indeed a local maximizer with respect to more general DSIC decision rules. The proof of this proposition follows from an existing result from Amador and Bagwell (2013).

**Proposition 4.7.** Assume Assumption 4.2 holds. Denote \((a_1, a_2)\) to be the optimal mechanism induced by \((\phi^*_1, \psi^*_1, \phi^*_2, \psi^*_2)\) in Theorem 4.6. Then, among all of DSIC decision rules, \(a_i\) maximizes the principal’s payoff given \(a_{-i}\).

### 4.3 Structure and Intuition of the Optimal Mechanism

The optimal mechanism is illustrated in Figures 2 and 3. We call \(x_i\) and \(\bar{x}_i\) as the joint coordinated lower and upper bound of the optimal mechanism, which are the largest lower bound and smallest upper bound of agent \(A_i\), respectively. They divide the state space into three parts: low states \([0, \bar{x}_i]\), middle states \([x_i, \bar{x}_i]\) and high states \([\bar{x}_i, 1]\).

- If \(s_i \in [0, \bar{x}_i]\), \(a_i \geq s_i\). When state is low, \(A_i\)’s action will only be coordinated above \(s_i\).
- If \(s_i \in [x_i, \bar{x}_i]\), \(a_i = s_i\). When state is in the middle, full adaptation is reached for \(A_i\).
- If \(s_i \in [\bar{x}_i, 1]\), \(a_i \leq s_i\). When state is high, \(A_i\) action will only be coordinated below \(s_i\).

The joint decision rule is characterized by dividing the state space into four areas as in the right panel of Figure 3.

- Area I includes all state pairs where two states are within the joint coordinated bound, and so no agent coordinates.
Figure 2: The contingent delegation interval of each agent.

Figure 3: The determining of optimal mechanism and different areas of decision rules.

- Area II includes all state pairs where two states are not close enough, and $s_1$ is more extreme than $s_2$ in the sense that either $s_1 \leq \bar{x}_1$ but $s_2 \leq \bar{x}_2$ or $s_1 \geq \bar{x}_1$ but $s_2 \geq \bar{x}_2$. In this case, only $A_1$ coordinates while $A_2$ reaches full adaptation. Consequently, $A_1$ will receive a unilateral delegation bound in this area. Area III is opposite to Area II.

- Area IV includes state pairs where both states are far away and both are very extreme such that whenever $s_i \leq \bar{x}_i$, $\bar{x}_{-i} \leq s_{-i}$. Therefore, both $A_1$ and $A_2$ receive the joint coordinated bound to coordinate together. Notice that the joint coordination is state-independent: $(a_1, a_2)$ is constant for every state in either the upper left or the lower right corner.

We can compare the above joint decision rule with the first best decision rule where the
principal can observe the states $s_1$ and $s_2$. The first best rule minimizes

$$\lambda_1(a_1 - s_1)^2 + \lambda_2(a_2 - s_2)^2 + (a_1 - a_2)^2$$

and hence is a convex combination of both states:

$$a_{i}^{FB}(s_1, s_2) = \frac{\lambda_i(\lambda_{-i} + 1)s_i + \lambda_{-i}s_{-i}}{\lambda_i(\lambda_{-i} + 1) + \lambda_{-i}};$$

$$|a_{i}^{FB}(s_1, s_2) - s_i| = \frac{\lambda_{-i}|s_{-i} - s_i|}{\lambda_i(\lambda_{-i} + 1) + \lambda_{-i}}.$$

It is straightforward to verify that coordination is always moderate in both the optimal mechanism and the first best decision rule:

**Proposition 4.8.** For $\forall s_i \leq s_{-i}$, we have:

$$s_i \leq a_i(s_1, s_2) \leq a_{-i}(s_1, s_2) \leq s_{-i} \quad \text{and} \quad s_i \leq a_{i}^{FB}(s_1, s_2) \leq a_{-i}^{FB}(s_1, s_2) \leq s_{-i}.$$

As we have seen, the design of the optimal contingent delegation scheme is governed by both the state-dependent unilateral coordination bounds and the state-independent joint coordination bounds. Moreover, whenever agent $A_i$’s action reaches his joint coordinated bound, the other agent’s action will also be forced away from his most preferred point. On the contrary, whenever agent $A_i$’s action reaches unilateral coordinated bounds, the other agent always reaches full adaptation. So in the optimal mechanism, we have three cases: i) no coordination where both agents reach full adaptation; ii) unilateral coordination where one agent coordinates while the other one reaches full adaptation; and iii) joint coordination where both agents choose a pair of state-independent coordination actions. In contrast, under the first best decision rule, there is always joint coordination with state-dependent coordination actions.

To solve the optimal mechanism, we only need to solve two single agent problems (unilateral coordinated boundaries) and then modify the solution according to Theorem 4.6. To understand the intuition of this result, or put it in another way, to understand why $(\phi_i, \psi_i)$ as the one sided optimizer of $a_{-i}$ will almost coincide with the unilateral coordinated bounds, even though $a_{-i} \neq s_{-i}$, we consider a special case where state 2 satisfies $s_0^2 \leq s_2$. The key is to explain why $\phi_1(s_0^2)$ coincides with $c_1^a(s_0^2)$ and $\psi_1(s_0^2)$ coincides with $d_1^a(x_2)$, using the shape of $a_2$ and the fact that $(\phi_1, \psi_1)$ is one-sided optimizer of $a_2$. 
When $s_1 \leq s_2^0$, then we immediately obtain:

$$c_2^s(s_1) \leq s_1 \leq s_2^0 \leq \bar{x}_2 < \bar{x}_2,$$

which implies $s_2^0 \in [\phi_2^s(s_1), \psi_2^s(s_1)]$. As a result, when $s_1 \leq s_2^0$, agent $A_2$ is free to choose his ideal point $a_2 = s_2^0$. Notice that $(\phi_1^s(s_2^0), \psi_1(s_2^0))$ is the one-sided optimizer of $a_2(s_1, s_2^0)$ as a function of $s_1$, and that $[0, s_2^0]$ is the critical neighbour of $s_1$ which determines $\phi_1(s_2^0)$, since we know $\phi_1(s_2^0) \leq s_2^0$ from diagonal separation. Now we can apply the local-determining property in Proposition 4.4 and the fact that $a_2(s_1, s_2^0)$ coincides with $s_2^0$ when $s_1 \in [0, s_2^0]$ to see that $\phi_1(s_2^0)$ coincides with $c_2^s(s_2^0)$. The argument for the other side is more involved, but intuitively since $a_2(s_1, s_2^0)$ coincides with $\bar{x}_2$ in the critical neighborhood $[\bar{x}_1, 1]$, $\psi_1(s_2^0)$ coincides with $d_1^s(x_2)$.

5 Properties of the Optimal Mechanism

This section analyzes the properties of the optimal mechanism characterized in Theorem 4.6. We show that the optimal mechanism is group strategyproof, and conduct comparative static analysis to illustrate how relative importance and state distribution affect the optimal mechanism. Throughout this section, we assume Assumption 4.1 holds.

State-independent Joint Coordination & Group Strategyproofness One interesting property of the optimal mechanism is that the joint coordination is state-independent. As long as both agents coordinate together, their action will always be the same. This structure brings about the group strategyproofness property, which is not a property of DSIC mechanism per se.

Proposition 5.1 (Group Strategyproofness). Denote $a_i^s$ to be the optimal decision rule. For any given $(s_1, s_2)$, any $(s'_1, s'_2)$ satisfying $U_A^1(a_i^s(s'_1, s'_2), s_1) > U_A^1(a_i^s(s_1, s_2), s_1)$ must lead to $U_A^2(a_2^s(s'_1, s'_2), s_2) < U_A^2(a_2^s(s_1, s_2), s_2)$.

As a matter of fact, at least one agent reaches his ideal point in Areas I, II and III, thus any profitable group deviation must report his state truthfully, then incentive compatibility will guarantee that no profitable group deviation exists. So we only need to verify the group strategyproofness in area IV, where joint coordination happens. Since any group deviation could only lead to joint actions in Area I (including the boundaries). Compared with the joint actions located at the intersection point $(\bar{x}, \bar{x})$ or $(\bar{x}, \bar{x})$, these actions will always make at least one agent worse off when the state pairs are in Area IV.
The state-independent joint coordination in the optimal mechanism plays an important role in the above group strategyproofness result. To demonstrate this point, let us consider another mechanism that always delegates the unilateral coordinated bounds to both agents. This new mechanism is illustrated in Figure 4, and similarly we can define the four areas. The joint actions when the state pair is located at point B in Figure 4 are located inside Area I (point C), and two agents can benefit from misreporting their states to be at the intersection point $(\bar{x}, \bar{x})$ (point A in Figure 4).

![Figure 4: Example of mechanism without group strategyproofness.](image)

The group strategyproofness we have just shown is a property of direct mechanism. To prevent group deviation in the implementation (delegation with autonomous constraint), the principal could simply prohibit any joint action outside the support of the optimal mechanism. He could do so by an ex post check and by imposing a sufficiently large punishment for any violation.

**Optimal Discretion and Relative Importance** Having established a whole picture of the optimal mechanism in mind, we now turn to the comparative static analysis. We are particularly interested in the optimal discretion delegated to the agents, which is the central topic in the single-agent delegation literature. Since the optimal mechanism is governed by the unilateral coordination, our next observation points out the impact of changing the unilateral coordinated boundary functions on the optimal delegated discretion:

**Observation 5.2.** When agent $A_i$’s unilateral coordinated lower/upper boundary shifts upward/downward, this agent will receive less delegated discretion in terms of higher probability.
of unilateral coordination and tighter delegation bounds. His counterpart, on the contrary, will enjoy higher probability of joint coordination.\footnote{For agent $A_i$, the joint coordinated bound will be looser than the unilateral coordinated bound for the same $s_{-i}$.}

The above observation is illustrated in Figure 5, and its proof is hence omitted. Figure 5 depicts the case where agent $A_1$’s unilateral coordinated lower boundary shifts upward. From Equation (13), we can easily see a change in the intersection point: both $\bar{x}_1$ and $\bar{x}_2$ become larger. This directly implies that the lower bound of agent $A_1$’s contingent delegation interval $\phi_1^*$ shifts upward as shown by the middle panel of Figure 5, and the upper bound of agent $A_2$’s contingent delegation interval $\psi_2^*$ shifts upward as shown by the right panel of Figure 5. So agent $A_1$ receives less delegated discretion while agent $A_2$ receives more delegated discretion.

Observation 5.2 implies that we only need to explore the impact of changing parameters on the unilateral coordinated boundary functions. The next proposition shows how the change in relative importance parameters affects the unilateral coordinated boundary functions.

**Proposition 5.3.** For a fixed distribution $F_i$, denote $(c_i^{\lambda_i}, d_i^{\lambda_i})$ to be the unilateral coordinated boundary functions with relative importance parameter $\lambda_i$.\footnote{Note that to determine $c_i^*$ and $d_i^*$, we need no information about $F_{-i}$ and $\lambda_{-i}$.} Then:

$$c_i^{\tilde{\lambda}_i} \leq c_i^{\lambda_i}, \quad d_i^{\tilde{\lambda}_i} \geq d_i^{\lambda_i}, \quad \forall \tilde{\lambda}_i \geq \lambda_i.$$  

Moreover,

$$\lim_{\lambda_i \to \infty} c_i^{\lambda_i} = 0, \quad \lim_{\lambda_i \to \infty} d_i^{\lambda_i} = 1;$$

$$\lim_{\lambda_i \to 0} c_i^{\lambda_i}(s_{-i}) = s_{-i} = \lim_{\lambda_i \to 0} d_i^{\lambda_i}(s_{-i}).$$
Combing Proposition 5.3 and Observation 5.2, we can rebuild the classical conclusion that, with less important adaptation or more interest conflict, the agent will suffer from less delegated discretion. In particular, if the state distributions are the same such that $F_1 = F_2$, the agent with less important adaptation will receive less delegated discretion than the agent with more important adaptation. Moreover, different from the classical model (e.g., Dessein (2002) and Alonso, Dessein and Matouschek (2008)) in which the principal can only choose between delegation and non-delegation, the optimal mechanism in general features partial delegation for interior value of relative importance.

However, in settings where both $\lambda_1$ and $\lambda_2$ change in the same direction, it is usually not easy to tell how an agent’s delegated discretion would change because the counter effect from the other project may mitigate or even overturn the claim in Proposition 5.3. Intuitively, such an overturn happens because the relative importance of adaptation losses between these two projects may also change significantly, leading to an additional transition of delegated discretion. Thus it is reasonable to expect that when the relative importance of adaptation losses between the two projects remains stable or alternatively when there is a pure increase in the relative importance of coordination (recall we normalize it to 1), the discretion received by each agent will decrease as adaptation becomes less important, because the power transition between the two agents is not significant. This is indeed true as shown by the following proposition:

**Proposition 5.4.** For fixed distribution function $F_i$ and $\kappa_i$, denote $(\phi_1^\lambda, \psi_1^\lambda, \phi_2^\lambda, \psi_2^\lambda)$ as the optimal boundary functions with relative importance parameters $\lambda_i = \lambda \kappa_i$. Then both agents will receive tighter delegation interval if $\lambda$ decreases. Formally, we will have:

$$
\bar{x}_i^\lambda \leq \bar{x}_i^\lambda, \quad \bar{x}_i^\lambda \geq \bar{x}_i^\lambda; \quad \phi_i^\lambda \leq \phi_i^\lambda, \quad \psi_i^\lambda \geq \psi_i^\lambda \quad \forall \lambda \geq \lambda.
$$

The above proposition establishes another monotonic relationship: as coordination becomes less important, not only does the total discretion received by both agents increase, but also each agent receives more delegated discretion.

**Discretion and State distribution** Apart from the relative importance, the principal’s subjective belief about the state distributions will also influence the delegated discretion. For further characterization, we introduce the notion of tail expectation dominance.

**Definition 5.5.** Consider two random variables $X$ and $Y$. We say that random variable $X$
dominates $Y$ in lower tail expectation, if and only if:

$$E(p_X|X \leq x) \leq E(p_Y|Y \leq x), \quad \forall x \in \mathbb{R}.$$ 

We say that random variable $X$ dominates $Y$ in upper tail expectation, if and only if:

$$E(p_X|X \geq x) \geq E(p_Y|Y \geq x), \quad \forall x \in \mathbb{R}.$$ 

If a random variable $X$ dominates another random variable $Y$ in either lower or upper tail expectations, this means that the distribution of $X$ has a fatter tail than the distribution of $Y$. The following example provides a class of density functions satisfying dominance in lower/upper tail expectations:

**Example 5.6.** Consider a class of r.v. $X_k$ with density function $f_k$ defined as:

$$f_k(x) = \begin{cases} 10x \leq 1(k + 1)x^k & \text{if } k \in \mathbb{Z}^+ \\ 10x \leq 1(1 - k)(1 - x)^{-k} & \text{if } k \in \mathbb{Z}^- \end{cases}$$

Then $X_k$ dominates $X_l$ in upper tail expectation and $X_l$ dominates $X_k$ in lower tail expectation for any $k \geq l$.

Intuitively, given the other agent’s state distribution, if one agent’s state distribution has a fatter tail, the constrained delegation will induce more adaptation loss for this agent. Therefore, it is natural that as the tail grows fatter, the corresponding agent should be granted more delegated discretion. In addition, for a better coordination with such extra discretion, his counterpart will face less discretion. This intuition is formalized by the following proposition:

**Proposition 5.7.** For fixed $\lambda_i$, denote $(c_i^a, d_i^a)$ and $(\bar{c}_i^a, \bar{d}_i^a)$ as the unilateral coordinated boundary functions with respect to distribution $F_i$ and $\bar{F}_i$.

If $\bar{F}_i$ dominates $F_i$ in lower tail expectation, we will have $\bar{c}_i^a \leq c_i^a$; and similarly if $\bar{F}_i$ dominates $F_i$ in upper tail expectation, we will have $\bar{d}_i^a \leq d_i^a$.

For instance, let us recall Example 5.6. If one of the states is distributed according to $f_k$ and the other is fixed, then combining Proposition 5.7 and Observation 5.2, we immediately know that the corresponding optimal contingent delegation interval for both agents will shift upwards as $k$ increases.
6 Concluding Remarks

This paper is the first study of the optimal dominant strategy incentive compatible mechanism in the framework of multi-divisional firms facing the trade-off between adaptation and coordination. Under certain assumptions on the distributions, we fully characterize the unique optimal mechanism in the class of weakly continuous mechanisms, i.e., mechanisms that can be represented by contingent delegation intervals. We show that this optimal mechanism shares some similarity with the optimal delegation mechanism in the single-agent problem through the joint coordinated delegation bounds. Moreover, this optimal mechanism is group strategyproof and has nice comparative static properties.

In future work, our model can be extended in the following dimensions. First of all, it is still unclear whether it is indeed optimal for the principal to use contingent delegation intervals instead of more general contingent delegation sets. Proposition 4.7 provides a partial justification by showing that the optimal mechanism characterized in our model is a local maximizer with respect to more general DSIC decision rules when slightly strengthening the assumptions on the distributions. However, it is difficult to arrive at a complete justification in our multi-divisional setting. The main difficulty comes from the fact that when deciding the optimal delegation scheme for agent $A_i$, the other agent’s decision rule $a_{-i}$ will enter into the principal’s objective function. However, even very nicely behaved $a_{-i}$ cannot guarantee the optimality of interval delegation for agent $A_i$. For example, we could construct examples where $a_{-i}$ is continuous and monotonic, but it is suboptimal for agent $A_i$ to use interval delegation. There are several papers deriving the conditions that guarantee the optimality of continuous mechanisms in the single-agent mechanism design setting such as Melumad and Shibano (1991), Martimort and Semenov (2006), Alonso and Matouschek (2007), Alonso and Matouschek (2008), and Amador, Bagwell and Frankel (2018). It is a challenge for future work to derive such conditions in multi-agent mechanism design problems.

Second, we focus on dominant strategy incentive compatible mechanisms so that our problem has a nice connection to the optimal delegation problems. Yet in our setting, it is possible that Bayesian mechanisms can reach a higher expected payoff to the principal than the dominant strategy mechanisms. However, it is very challenging to solve the optimal Bayesian mechanism since we do not have the nice characterization of incentive compatible mechanisms. Moreover, it would also make for interesting future work to investigate whether introducing stochastic mechanisms à la Kovac and Mylovanov (2009) or transfers à la Krishna and Morgan

\footnote{The equivalence result in Gershkov et al. (2013) does not apply due to our non-linear setting.}
(2008) can improve the principal’s expected payoff.

Appendix

A Proofs in Section 3

Proof of Proposition 3.2

Proof. For a given DSIC mechanism \((a_1, a_2)\), let

\[
B_i(s_{-i}) = \{a_i(s_i, s_{-i}) | \forall s_i \in [0, 1]\} \subseteq [0, 1], \quad \text{and} \quad D_i(s_{-i}) = \overline{B_i(s_{-i})}.
\]

Then \(D_i(s_{-i})\) is closed by definition. Besides, according to the \((IC)\) condition 2, we have

\[
a_i(s_i, s_{-i}) \in \arg \max_{a \in B(s_{-i})} U_A^i(a, s_i).
\]

Since \(U_A^i\) is continuous, \(\sup_{a \in B(s_{-i})} U_A^i(a, s_i) = \sup_{a \in D(s_{-i})} U_A^i(a, s_i)\). Therefore \(a_i(s_i, s_{-i}) \in \arg \max_{a \in D(s_{-i})} U_A^i(a, s_i)\).

Conversely, for any decision rule \((a_1, a_2)\) and two collections of sets \((D_1(s_2), D_2(s_1)\) satisfying the two properties in the proposition, we know \(a_i(s_i, s_{-i}) \in D(s_{-i})\), and thus it is straightforward to verify that the \((IC)\) conditions are all satisfied.

Proof of Proposition 3.6

Proof. The proof is based on the following lemma:

Lemma A.1 (Swapping Property). Suppose \((c_1, d_1)\) and \((c_2, d_2)\) are both one-sided optimizers of a given function \(y(s_i)\), and \([c_1, d_1] \cap [c_2, d_2] \neq \emptyset\). Then both \((c_1, d_2)\) and \((c_2, d_1)\) are one-sided optimizers of \(y(s_i)\) as well.

First, we show that the set of one-sided optimizers is a sub-lattice. Denote \((c_1, d_1)\) and \((c_2, d_2)\) are two one-sided optimizer of \(y\) for \(i\), and \(c_1 \leq c_2\). If \(c_1 \leq c_2\) and \(d_1 \leq d_2\), then it is trivial. Otherwise \([c_1, d_1] \cap [c_2, d_2] \neq \emptyset\), thus according to Lemma A.1, both \((c_1, d_2)\) and \((c_2, d_1)\) are one-sided optimizers.

Second, since the objective function is continuous with respect to the pair \((c, d)\), the set of one-sided optimizers is clearly closed, and thus compact. Then we can apply a topology result of Birkhoff (1967) which says any nonempty compact sub-lattice of \(R^n\) is complete.

26
Proof of Lemma A.1

Proof. Denote \( b \in [c_1, d_1] \cap [c_2, d_2] \). Denote \( a_{1i}(s_i) \) as the action function with delegation interval \((c_1, d_1)\) and \( a_{2i}(s_i) \) as the action function with delegation interval \((c_2, d_2)\) as defined by Equation (5). Then we claim that:

\[
\begin{align*}
\int_0^b (a_{1i} - y)^2 + \lambda_i(a_{1i} - s_i)^2 dF_i(s_i) &= \int_0^b (a_{2i} - y)^2 + \lambda_i(a_{2i} - s_i)^2 dF_i(s_i); \\
\int_0^1 (a_{1i} - y)^2 + \lambda_i(a_{1i} - s_i)^2 dF_i(s_i) &= \int_0^1 (a_{2i} - y)^2 + \lambda_i(a_{2i} - s_i)^2 dF_i(s_i).
\end{align*}
\]

Otherwise, assuming in the first equation \(-\) is replaced with \(<\) for example, we can construct a new action function \( a_i \) with delegation interval \((c_1, d_2)\). In fact, we will have:

\[
\begin{align*}
\int_0^1 H(c_1, d_2, y) dF_i &\overset{\Delta}= \int_0^1 (a_i - y)^2 + \lambda_i(a_i - s_i)^2 dF_i(s_i) \\
= \int_0^b H(c_1, d_2, y) dF_i + \int_b^1 H(c_1, d_2, y) dF_i \\
= \int_0^b H(c_1, d_1, y) dF_i + \int_b^1 H(c_2, d_2, y) dF_i \\
< \int_0^b H(c_2, d_2, y) dF_i + \int_b^1 H(c_2, d_2, y) dF_i
\end{align*}
\]

where Equation (18) holds because \( d \in [c_1, d_1] \cap [c_2, d_2], \) and the inequality comes from the assumption that \( a_i \) brings strictly less loss than \( a_{2i}, \) which is a contradiction.

Now that Equations (14) and (15) both hold, then delegation intervals \([c_1, d_2]\) and \([c_2, d_1]\) are clearly equally as good as the optimal delegation interval. \( \square \)

Proof of Proposition 3.7

Proof. Denote the integral in the definition of one-sided optimization as a function of delegation interval and \( y \):

\[
\pi(c, d, y) = \int_0^1 - (a_i(c, d, s_i) - y(s_i))^2 - \lambda_i(a_i(c, d, s_i) - s_i)^2 dF_i(s_i),
\]

where \( a(c, d, s_i) \) is the action function with delegation interval \((c, d)\). Now we will prove that \( f \) is quasi-supermodular in \((c, d)\) and satisfies the single crossing property in \((c, d, y)\). Then a direct application of Milgrom and Shannon (1994) will complete the proof.

First we prove \( f \) satisfies the single crossing property. Suppose \( y \leq \bar{y} \) and \((c, d) \leq (\bar{c}, \bar{d}). \)
Then \( \bar{a}_i = a_i(\bar{c}, \bar{d}, s_i) \geq a_i(c, d, s_i) = a_i \). Therefore, the difference

\[
\pi(\bar{c}, \bar{d}, y) - \pi(c, d, y) = \int_0^1 (a_i - y)^2 + \lambda_i(a_i - s_i)^2 - (\bar{a}_i - y)^2 - \lambda(\bar{a}_i - s_i)^2 dF_i(s_i)
\]

is clearly weakly increasing in \( y \), thus \( f \) satisfies the single crossing property. Now we prove the quasi-supermodular property. Consider \((c_1, d_1)\) and \((c_2, d_2)\). Suppose without loss of generality that \( d_2 \geq d_1 \). Then if \( c_2 \geq c_1 \), we have nothing to verify. If \( c_2 < c_1 \), we need to prove that \( \pi(c_1, d_1, y) \geq \pi(c_2, d_1, y) \), then \( \pi(c_1, d_2) \geq \pi(c_2, d_2) \). We will use again the cutting and pasting trick: denote \( b \in (c_1, d_1) \cap (c_2, d_2) \neq \emptyset \). Since:

\[
\pi(c_1, d_1, y) = \int_0^b \left[ a_i(c_1, d_1, s_i) - y(s_i) \right]^2 + \lambda_i(a_i(c_1, d_1, s_i) - s_i)^2 dF_i(s_i)
\]

Therefore we have:

\[
\int_0^b H(c_1, d_1, y) dF_i \geq \pi(c_2, d_1, y)
\]

\[
\pi(c_1, d_2, y) = \int_0^b H(c_1, d_2, y) dF_i \int_0^1 H(c_1, d_2, y) dF_i
\]

\[
= \int_0^b H(c_1, d_2, y) dF_i \int_0^1 H(c_1, d_2, y) dF_i
\]

\[
\geq \int_0^b H(c_2, d_2, y) dF_i \int_0^1 H(c_2, d_2, y) dF_i
\]

\[
= \pi(c_2, d_2, y).
\]

Proof of Proposition 3.9  Let \( X = \left\{ (h_1, h_2, h_3, h_4) \big| h_i \text{ is function defined on } [0, 1] \text{ with finite } L_1 \text{ norm} \right\} \). It is easy to verify that:

Observation A.2. \( X \) is a normed linear space\(^{12}\) under standard addition, scalar multiplication

\(^{12}\)Here, we regard two vectors \((h_1, h_2, h_3, h_4)\) and \((g_1, g_2, g_3, g_4)\) satisfying \( h_i - g_i \) a.e. as one equivalent class.
and the norm\(^{13}\) defined as:

\[|(h_1, h_2, h_3, h_4)| = \max_i \int_0^1 |h_i(x)| dx.\] (20)

Now we will complete the first step of proof:

**Lemma A.3.** The set \(M\) in normed linear space \(X\) is a self sequentially compact set.

**Proof of Lemma A.3.** In order to prove \(M\) is a self sequentially compact set, we only need to verify that for any vector sequence \(\{ (\phi_1^n, \psi_1^n, \phi_2^n, \psi_2^n) \}\) \(\subseteq M\), there exist a subsequence \(\{n_k\}\) and a vector \((\phi_1^*, \psi_1^*, \phi_2^*, \psi_2^*) \in M\) such that \(\{ (\phi_1^{n_k}, \psi_1^{n_k}, \phi_2^{n_k}, \psi_2^{n_k}) \}\) converge to \((\phi_1^*, \psi_1^*, \phi_2^*, \psi_2^*)\).

First we use famous Diagonal method to select a subsequence \(\{n_j\}\) such that \(\phi_1^{n_j}\) converge in all rational number in \([0, 1]\) (which is a dense subset of \([0, 1]\)). Puzzled readers may refer to the proof of the Arzela-Ascoli theorem.

Then, because \(\{\phi_1^{n_j}\}\) is pointwisely bounded, we can define two real-value functions:

\[
\bar{\phi}_1(x) = \limsup_{j \to \infty} \phi_1^{n_j}(x) \forall x \in [0, 1]; \\
\bar{\phi}_1(x) = \liminf_{j \to \infty} \phi_1^{n_j}(x) \forall x \in [0, 1].
\]

Note that \(\phi_1^{n_j}\) are all non-decreasing. \(\bar{\phi}_1(x)\) and \(\bar{\phi}_1(x)\) are therefore both non-decreasing, and thus both continuous a.e. in \([0, 1]\). Denote \(C \overset{\Delta}{=} \{x \in [0, 1] \mid \bar{\phi}_1, \bar{\phi}_1\text{ are continuous at } x\}\), for \(\forall x \in C\), we can find two sequence of rational number \(\{r_1^n\}\) and \(\{r_2^n\}\), such that \(r_1^n < x < r_2^n\) and \(\lim r_1^n = \lim r_2^n = x\). Thus the following equations will hold:

\[
\lim_{n \to \infty} \bar{\phi}_1(r_1^n) \leq \bar{\phi}_1(x) \leq \lim_{n \to \infty} \bar{\phi}_1(r_2^n) \\
\lim_{n \to \infty} \bar{\phi}_1(r_1^n) = \lim_{n \to \infty} \bar{\phi}_1(r_2^n) = \bar{\phi}_1(x) = \lim_{n \to \infty} \bar{\phi}_1(r_2^n) = \lim_{n \to \infty} \bar{\phi}_1(r_2^n),
\]

which yields \(\bar{\phi}_1(x) = \bar{\phi}_1(x)\).

Now we can define:

\[
\phi_1^*(x_0) = \sup_{x \in C, x \leq x_0} \bar{\phi}_1(x_0)
\]

It is easy to see \(\phi_1^*\) is a non-decreasing function in \([0, 1]\), and for \(x \in C\),

\[
\phi_1^*(x) = \bar{\phi}_1(x) = \bar{\phi}_1(x) = \lim_{j \to \infty} \phi_1^{n_j}(x).
\]

\(^{13}\)We will call it the vector norm in the following analysis.
That is to say, $\phi_1^{n_1}$ converge to $\phi_1^*$ a.e., according to bounded convergent theorem, we can further know $\phi_1^{n_j}$ converge to $\phi_1^*$ a.e. in the $L_1$ norm.

By taking subsequence iteratively, we can finally define the desired $(\phi_1^*, \psi_1^*, \phi_2^*, \psi_2^*) \in M$, and subsequence $\{n_k\}$ such that $\phi_1^{n_k}$ and $\psi_1^{n_k}$ converge in the $L_1$ norm to $\phi_1^*$ and $\psi_1^*$ respectively. It is easy to see the convergence in the vector norm is also satisfied.

Since $X$ is a normed linear space, and thus a metric space, to prove continuity requirement in the second step, it is sufficient to prove:

**Lemma A.4.** For any sequence $(\phi_1^n, \psi_1^n, \phi_2^n, \psi_2^n)$ that converges to $(\phi_1, \psi_1, \phi_2, \psi_2)$ in the vector norm. the value of its corresponding ex ante payoff $U_n$ also converges to $U$.

**Proof of Lemma A.4.** In order to prove $U_n$ converges to $U$, we will prove that for any subsequence $U_{n_k}$ of $U_n$, there is a sub-subsequence $U_{n_{k_j}}$ that converges to $U$.

Assume $U_{n_k}$ is a subsequence. From the definition of the vector norm, $\phi_i^n, \psi_i^n$ converge to $\phi_i, \psi_i$ in the $L_1$ norm respectively. Thus $\phi_i^n, \psi_i^n$ converge to $\phi_i, \psi_i$ in measure, and therefore $\phi_i^{n_k}$ converges to $\phi_i$ in measure as well. According to the Riesz theorem, there exists a sub-subsequence $\phi_i^{n_{k_j}}$ which converge to $\phi_i$ a.e. in $[0, 1]$. By taking subsequence iteratively, we can finally get a sub-subsequence $\{n_{k_j}\}$, such that, $\phi_i^{n_{k_j}}$ and $\psi_i^{n_{k_j}}$ converge to $\phi_i^{n_{k_j}}$ and $\psi_i^{n_{k_j}}$ a.e. in $[0, 1]$ respectively.

Denote $(a_1^{n_{k_j}}, a_2^{n_{k_j}})$ as the corresponding mechanism of $(\phi_i^{n_{k_j}}, \psi_i^{n_{k_j}}, \phi_2^{n_{k_j}}, \psi_2^{n_{k_j}})$. Denote the set where $a_i^{n_{k_j}}$ does not converge to $a_i$, the set where $\phi_i^{n_{k_j}}$ does not converge to $\phi_i$, the set corresponding to $\psi_i^{n_{k_j}}$ as $A_i^1$; the set corresponding to $\psi_i^{n_{k_j}}$ as $A_i^2$. Notice that $A_i^1$ and $A_i^2$ are zero measure sets in $R^1$, and $A_i \subseteq (A_i^1 \cup A_i^2) \times [0, 1]$. Thus $A_i$ is a zero measure set in $R^2$, that is, $(a_1^{n_{k_j}}, a_2^{n_{k_j}})$ converge to $(a_1, a_2)$ a.e. in $[0, 1] \times [0, 1]$.

Then using bounded dominant theorem we will know $U_{n_{k_j}}$ converges to $U$, which completes the proof.

Now the remaining part of proof is merely standard routine:

**Proof of proposition 3.9.** Since $M$ is a self sequentially compact set in metric space $X$, and the payoff mapping is continuous, there is an equivalent class in $M$ who has the highest ex ante payoff. Since all the vectors in one equivalent class share the same ex ante payoff, we can select a vector $(\phi_1^*, \psi_1^*, \phi_2^*, \psi_2^*)$ that is really in $M$ (not in a.e. sense), which is the desired mechanism.

---

**Note:** This is a degenerate form of Lemma A.4.
B Proofs in Section 4

Proof of Lemma 4.3

Proof. Again we denote \( \pi(c,d,y = s_{-i}) \) \( \triangleq \int_{0}^{1} -(\bar{a}_{i} - s_{-i})^2 - \lambda_{i}(\bar{a}_{i} - s_{i})^2 dF_{i}(s_{i}) \). Then:

\[
\frac{\partial \pi}{\partial c} = -2\lambda_{i}cF_{i}(c) + 2\lambda_{i} \int_{0}^{c} s_{i}f_{i}(s_{i})ds_{i} - 2F(c)(c - s_{-i}) = -2\int_{0}^{c} \lambda_{i}F_{i}(s_{i})ds_{i} - 2F_{i}(c)(c - s_{-i}) .
\]

Note that for \( c > s_{-i} \), \( \partial \pi/\partial c < 0 \), so \( c(s_{-i}) \leq s_{-i} \). Moreover, when \( c = s_{-i} > 0 \), \( \partial \pi/\partial c < 0 \),\(^{15}\) so \( c^{*}(s_{-i}) < s_{-i} \). We then define:

\[
g(c) = \frac{\partial \pi/\partial c}{2F_{i}} = - \int_{0}^{c} \lambda_{i} \frac{F_{i}(s_{i})}{F_{i}(c)}ds_{i} - (c - s_{-i})
\]

\[
\frac{\partial g}{\partial c} = -\lambda_{i}(1 - \int_{0}^{c} \frac{F_{i}(s_{i})f_{i}(c)}{F_{i}^{2}(c)}) - 1 < -\lambda_{i}(1 - \int_{0}^{c} \frac{F_{i}(s_{i})f_{i}(c)}{F_{i}^{2}(c)}) .
\]

Recall the log concavity of \( F_{i} \), we know \( f_{i}/F_{i} \) is decreasing w.r.t \( s_{i} \), thus:

\[
\frac{F_{i}(c)}{\int_{0}^{c} F_{i}(s_{i})ds_{i}} = \frac{\sum_{k} f_{i}(s_{i})ds_{i}}{\int_{0}^{c} F_{i}(s_{i})ds_{i}} = \lim_{n} \frac{\sum_{k} f_{i}(ck/n)/n}{\lim_{n} \sum_{k} F_{i}(ck/n)/n} = \frac{\sum_{k} f_{i}(ck/n)}{\sum_{k} F_{i}(ck/n)} > \frac{f_{i}(c)}{F_{i}(c)} ;
\]

therefore \( g \) is decreasing in \([0,s_{-i}]\). Note that when \( s_{-i} = 0 \), \( c^{*}(s_{-i}) = 0 \) is the unique optimal solution. While when \( s_{-i} > 0 \), we have \( g(0) > 0 \) and \( g(s_{-i}) < 0 \), so the unique zero point of \( g \) (which is also the unique zero point of \( \partial \pi/\partial c \)) is indeed the unique solution maximizing \( \pi \). \( g(c^{*}(s_{-i})) = 0 \) is exactly equivalent to Equation 9, and by applying integration by parts, we can immediately get Equation 11.

Now that since the uniqueness of the one-sided optimizer has been proved, monotone Lemma B.1 immediately tells us \( 0 \leq c^{*}_{1}(s_{-i}) - c^{*}_{1}(s'_{-i}) \). And from the first order condition, it is easy to see the above inequality strictly holds. Furthermore, we have:\(^{16}\)

\[
(s_{-i} - s'_{-i} - c^{*}_{1}(s_{-i}) + c^{*}_{1}(s'_{-i})) = \lambda_{i} \int_{0}^{c^{*}_{1}(s_{-i})} \frac{F_{i}(s_{i})}{F_{i}(c^{*}_{1}(s_{-i}))}ds_{i} - \lambda_{i} \int_{0}^{c^{*}_{1}(s'_{-i})} \frac{F_{i}(s_{i})}{F_{i}(c^{*}_{1}(s'_{-i}))}ds_{i}
\]

\[
= \lambda_{i}(1 - \int_{0}^{c} \frac{f_{i}(c)F_{i}(s_{i})}{F_{i}^{2}(c)}ds_{i})(c^{*}_{1}(s_{-i}) - c^{*}_{1}(s'_{-i})) > 0 ,
\]

which is the desired results. The proof for \( d \) is exactly the same, using the condition that \( f_{i}/(1 - F_{i}) \) is increasing. 

\(^{15}\)Here we use the assumption that \( f_{i} \) is strictly positive.

\(^{16}\)First differential mean value formula is used in the above equation, and \( c \in [c^{*}_{1}(s_{-i}),c^{*}_{1}(s'_{-i})] \)
Proof of Proposition 4.4

Proof. We prove the first part of Proposition 4.4 by contradiction. If instead, $\exists c^* \in \{c| (c, d) \in D_1\} \backslash \{c| (c, d) \in D_2\}$, then $\exists d^* s.t. (c^*, d^*) \in D_1$, and $\exists (c, d) \in D_2$. Then since $(c^*, d) \notin D_2$, we have:

$$\int_{0}^{1} (a_i - y_2)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c^*, d)} > \int_{0}^{1} (a_i - y_2)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c, d)}.$$  

With condition $c, c^* \leq c_0 \leq d, d^*$, we obtain:

$$\int_{0}^{c_0} (a_i - y_2)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c^*, d)} = \int_{0}^{1} H(c^*, d, y_2) dF_i - \int_{c_0}^{1} H(c, d, y_2) dF_i$$

$$> \int_{0}^{1} H(c, d, y_2) dF_i - \int_{c_0}^{1} H(c, d, y_2) dF_i = \int_{0}^{c_0} (a_i - y_2)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c, d)}.$$  

Moreover, because $y_1(s_i) = y_2(s_i) \forall s_i \leq c_0$, we have:

$$\int_{0}^{c_0} (a_i - y_1)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c^*, d)} > \int_{0}^{c_0} (a_i - y_1)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c, d)}.$$  

And thus:

$$\int_{0}^{1} (a_i - y_1)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c^*, d^*)} = \int_{0}^{c_0} H(c^*, d, y_1) dF_i + \int_{c_0}^{1} H(c, d^*, y_1) dF_i$$

$$> \int_{0}^{c_0} H(c, d, y_1) dF_i + \int_{c_0}^{1} H(c, d^*, y_1) dF_i = \int_{0}^{1} (a_i - y_1)^2 + \lambda_i(a_i - s_i)^2 dF_i \bigg|_{(c, d^*)}.$$  

This contradicts with the optimality of $(c^*, d^*)$ with respect to $y_1$.

The second part of this proposition can be proved similarly, and hence is omitted. □

Proof of Proposition 4.5

Proof. To prove this proposition, we first establish a useful corollary from the monotone property of one-sided optimizers.

Lemma B.1. Assume $y(s_i) \leq \bar{y}(s_i)$ are two functions mapping from $[0, 1]$ to $[0, 1]$ and $(c^*, d^*)$ (resp. $(\bar{c}^*, \bar{d}^*)$) is a one-sided optimizer of $y(s_i)$ (resp. $\bar{y}(s_i)$) for $i$. Moreover, one of the one-sided optimizers is unique. Then, we must have: $c^* \leq \bar{c}^*$ and $d^* \leq \bar{d}^*$.

The proof of the above lemma is obvious and hence is omitted.

Before formal proof, we need to calculate another class of one-sided optimizers of $y(s_i)$ and $\bar{y}(s_i)$, which consist of a piece of a constant and a piece of the diagonal line.
Lemma B.2. Denote $\bar{y}_x(s_i)$ and $y_x(s_i)$ as follows:

$$
y_x(s_i) = \begin{cases} x & \text{if } s_i \leq x; \\
 s_i & \text{if } s_i > x; \\
\end{cases} \quad \bar{y}_x(s_i) = \begin{cases} s_i & \text{if } s_i \leq x; \\
x & \text{if } s_i > x. \end{cases}$$

Then the one-sided optimizers of $y_x(s_i)$ and $\bar{y}_x(s_i)$ are unique, and the corresponding optimizers are $(c^*_1(x), 1)$ and $(0, d^*_2(x))$, respectively.

Proof of Lemma B.2. The proof of the lemma is easy, so we just give a brief description. Take $y_x(s_i)$ for example. Via discussion of the derivative, it is easy to verify the optimal “upper bound” should be 1, and the optimal “lower bound” should be less than $x$. Intuitively, when $(s_1, s_2)$ lies on the diagonal line, it is a win-win situation for the agents and the principal, as we will discuss later. Then comparing $y_x$ with a constant function $x$, we know that the unique optimal lower bound should coincide with $c^*_1(x)$ from Proposition 4.4 and Lemma 4.3.

Recalling $(\phi^*_1, \psi^*_1, \phi^*_2, \psi^*_2)$ is a one-sided optimizer of itself, we can consequently apply Lemma B.2 to compare it with $y_x$ and $\bar{y}_x$, and bound itself using the one-sided optimizer of $y_x$ and $\bar{y}_x$. The “Lebesgue method”, a continuous version of induction method, will be used to strengthen our bound gradually.

We prove $\phi^*_i(s_{-i}) \leq s_{-i}$, and the other part is the same. Denote $S = \{s \in [0, 1] | \phi^*_i(s') \leq s', \forall s', \forall i = 1, 2\}$. It is obvious that $1 \in S$, thus $S$ is not empty set, which guarantees that $s = \inf S$ is well defined. Moreover, according to the definition, if $s' \in S$, $s \in S$, $\forall s' \geq s$, so we just need to prove $s = 0 \in S$.

First we prove $s \in S$. Take a decreasing sequence of $s^n \in S$, s.t. $s^n \rightarrow s$. According to the monotonicity of $\phi^*_i$, we have:

$$\phi^*_i(s) \leq \lim \phi^*_i(s^n) \leq \lim s^n = s.$$

Second, we prove that $s = 0$ by contradiction. Suppose $s > 0$, define $y_s$ as above. Denote $a_i$ as the decision rule of the optimal mechanism. Since we know $a_2$ is either bounded by the lower delegation bound $\phi^*_2(s_i)$ or equals $s_2$ exactly, that is to say $a^*_2(s_i, s_2) \leq \max\{s_2, \phi^*_2(s_i)\}$. Moreover $\phi^*_2(s_1) \leq s_1, \forall s_1 \geq s$, and $\phi^*_2(s_1) \leq \phi^*_2(s) \leq s, \forall s_1 \leq s$. Therefore, for any fixed $s_2 \leq s$, $a^*_2(s_1, s_2) \leq y_s(s_1), \forall s_1$.

According to Lemma B.2, the unique one-sided optimizer of $y_s(s_1)$ w.r.t to $F_1$ is $(c^*_i(s), 1)$. Now that the $(\phi_i, \psi_i)$ is almost everywhere a one-sided optimizer of $(\phi_{-i}, \psi_{-i})$, applying Lemma B.1,
we know that $\phi^*_i(s_2) \leq c^*_i(s)$ a.e. $s_2 \leq s$. By monotonicity $\phi^*_i(s_2) \leq c^*_i(s) \forall s_2 \leq s$. Similarly, $\phi^*_2(s_1) \leq c^*_2(s) \forall s_1 \leq s$.

However, according to Lemma 4.3, $c = \max\{c^*_1(s), c^*_2(s)\} < s$, we get $\phi^*_{-i}(s_i) \leq c \leq s_i$, $\forall c \leq s_i < s$, so $c \in S$, which contradicts with the minimum property of $s$.

In conclusion, we have shown that $S = [0, 1]$, which completes the proof.  

\textbf{Proof of Theorem 4.6}  In order to avoid a lengthy proof, we first analyze the one-sided optimizer of the function appearing in the optimal mechanism. It is the combination of the constant function and the unilateral coordinated boundary function.

\textbf{Lemma B.3.} Let $0 < x < 1$, and denote $\bar{y}(s_i)$ and $y(s_i)$ as follow:

$$y(s_i) = \begin{cases} d^*_i(x) & \text{if } s_i \leq x; \\ d^*_i(s_i) & \text{if } s_i \geq x; \end{cases} \quad \bar{y}(s_i) = \begin{cases} c^*_i(s_i) & \text{if } s_i \leq x; \\ c^*_i(x) & \text{if } s_i \geq x. \end{cases}$$

Denote $(c, d)$ and $(\bar{c}, \bar{d})$ as one-sided optimizers of $y$ and $\bar{y}$ w.r.t $F_i$ respectively. Then:

$$c = c^*_i(d^*_i(x)), \quad \text{if } x_i \leq x;$$

$$\bar{d} = d^*_i(c^*_i(x)), \quad \text{if } \bar{x}_i \geq x.$$

\textbf{Proof of Lemma B.3.} As usual, we will prove the first half of the lemma. By definition $c^*_i(d^*_i(x_i)) = x_i$ we will prove later in the theorem that $x_i$ is the unique fixed point of $c^*_i \circ d^*_i$. Assuming this, since $c^*_i(d^*_i(1)) < 1$ and $x \geq x_i$, we know that $c^*_i(d^*_i(x)) \leq x$. Now denote $\pi(c, d, \bar{y}) \overset{\Delta}{=} \int_0^1 (\bar{y}(c, d, s_i) - \bar{y}(s_i))^2 + \lambda_i(\bar{y}(c, d, s_i) - s_i)^2 dF_i(s_i)\mid_{(c, d)}$. Note that $y \geq d^*_i$, then with the help of Lemma 4.3 and B.1, we know that $\bar{d} \geq d^*_i(d^*_i(x)) > d^*_i(x)$. Thus $\partial \pi / \partial c$ has two forms conditioned on the relation of $c$ and $x$. It is easy to calculate that:

$$\frac{\partial \pi}{\partial c} \bigg|_{c \leq x} = 2\lambda_i \int_0^x F_i(s_i)ds_i + 2F_i(c - d^*_i(x))$$

$$\frac{\partial \pi}{\partial c} \bigg|_{c \geq x} = 2F(x)(c - d^*_i(x)) + 2 \int_c^x (c - d^*_i(s_i))f_i(s_i)ds_i + 2\lambda_i \int_0^c (c - s_i)f_i(s_i)ds_i.$$

When $c \leq x$, $\partial \pi / \partial c$ coincides with the derivatives in the proof of Lemma 4.3, thus it is negative on $[0, c^*_i(d^*_i(x)))$, and positive on $(c^*_i(d^*_i(x)), x]$. Therefore, to complete the proof, we only
need to show it remains positive on \((x, 1]\). To do so, we establish the following inequality:

\[
(\lambda_i + 1) \frac{F_i(x_i)}{f_i(x_i)} - d_{\text{ini}}^*(x_i) + x_i > \lambda_i \int_{0}^{x_i} \frac{F_i(s_i)}{F_i(x_i)} ds_i - d_{\text{ini}}^*(x_i) + x_i \\
= \lambda_i \int_{0}^{c_i^*(d_{\text{ini}}^*(x_i))} \frac{F_i(s_i)}{F_i(c_i^*(d_{\text{ini}}^*(x_i)))} ds_i - d_{\text{ini}}^*(x_i) + c_i^*(d_{\text{ini}}^*(x_i)) \\
= 0
\]

where the first inequality comes from the log concavity of the distribution function, and the second equality comes from the definition of \(x_{\text{ini}}\) (Equation 13). Since the LHS of the above inequality is increasing w.r.t \(x_i\), the above inequality holds for \(\forall c > x_i\):

\[
(\lambda_i + 1) \frac{F_i(c)}{f_i(c)} - d_{\text{ini}}^*(c) + c > 0
\]

It is easy to calculate that:

\[
\frac{1}{2f_i(c)} \frac{\partial^2 \pi}{\partial c^2} |_{c=x} = (\lambda_i + 1) \frac{F_i(c)}{f_i(c)} - (d_{\text{ini}}^*(c) - c)
\]

Therefore \(\partial^2 \pi / \partial c^2 > 0\) for \(c > x\), thus \(\frac{\partial \pi}{\partial c} > 0\) for \(\forall c > x\), which completes the proof.

Finally, we will get to the core result of this paper.

**Proof of Theorem 4.6.** First we need to verify the existence and uniqueness of pair \((\bar{x}_i, x_i)\). From Lemma 4.3, we know \(c_i^*\) and \(d_{\text{ini}}^*\) is strictly increasing and has Lipschitz condition with parameter 1. Thus \(d_{\text{ini}}^{s_i-1}\) exists, and the required \((\bar{x}_i, x_i)\) will be the cross point of function \(d_{\text{ini}}^{s_i-1}\) and \(c_i^*\). The existence of such a cross point is obvious since \(d_{\text{ini}}^{s_i-1}(1) = 1\) and \(d_{\text{ini}}^{s_i-1}(s) = 0\) for some \(s \in [0, 1]\), while \(c_i^*(1) < 1\) and \(c_i^*(s) > 0\). Moreover, the uniqueness is true because \(|d_{\text{ini}}^{s_i-1}(s) - d_{\text{ini}}^{s_i-1}(s')| > |s - s'| > |c_i^*(s) - c_i^*(s')|\).

Take any optimal mechanism \((\phi_i^*, \psi_i^*, \phi_2^*, \psi_2^* ) \in M\). \((\phi_i, \psi_i)\) is monotone and is almost everywhere a one-sided optimizer of \((\phi_{-i}, \psi_{-i})\). According to Proposition 4.5, \(\phi_i^*(0) = 0 = c_i^*(0), \psi_i^*(1) = 1 = d_i^*(1)\). We denote \(\phi_i^*(0) = \bar{y}_i, \phi_i^*(1) = y_i\).

By construction, for almost any given \(s_i\), the pair \((\phi_{\text{ini}}^*(s_i), \psi_{\text{ini}}^*(s_i))\) should be the one-sided optimizer of \(a_i(s_i, s_{-i})\) as a function of \(s_{-i}\). With the help of the local property we will prove, for any given \(s_i \in [0, \bar{y}_i], \phi_{\text{ini}}^*(s_i) = c_{\text{ini}}^*(s_i)\).

To use the local property in Proposition 4.4, we verify the two conditions as follows. First we point out a common bound on \(\phi_{\text{ini}}^*(s_i)\) and \(c_{\text{ini}}^*(s_i)\). With the help of Diagonal-separation Proposition 4.5 and Lemma 4.3, we know that \(\phi_{\text{ini}}^*(s_i) \leq s_i\) and \(c_{\text{ini}}^*(s_i) \leq s_i\). Second, we
show that \( a_i(s_i, s_{-i}) \) coincides with the constant function \( s_i \) within this bound. Since \( \psi^*_i \) is non-decreasing, \( \psi^*_i(s_{-i}) \geq \psi^*_i(0) = \bar{y}_i \). Using Proposition 4.5, we obtain \( \phi_i^*(s_{-i}) \leq s_{-i} \). Thus, for \( s_{-i} \leq s_i \leq \bar{y}_i \), \( \phi_i^*(s_{-i}) \leq s_i \leq \bar{y}_i \leq \psi^*_i(s_{-i}) \). Equation 2 then yields:

\[
a_i^*(s_i, s_{-i}) = s_i, \quad \forall s_i, s_{-i} \text{ s.t. } s_{-i} \leq s_i \leq \bar{y}_i
\]

That is to say, \( a_i(s_i, s_{-i}) \) coincides with the constant function \( s_i \) when \( s_{-i} \leq s_i \leq \bar{y}_i \). Then the local property in Proposition 4.4 indicates that \( \phi^*_i(s) = c^*_i(s) \) for almost every \( s_i \leq \bar{y}_i \) and from monotonicity, \( \phi^*_i(s_i) = c^*_i(s_i) \forall s_i \leq \bar{y}_i \). Similar arguments hold for the other three equations. In conclusion, we get:

\[
\phi_i^*(s_{-i}) = c_i^*(s_{-i}) \text{ if } s_{-i} \leq \bar{y}_{-i}; \quad \psi_i^*(s_{-i}) = d_i^*(s_{-i}) \text{ if } s_{-i} \geq y_{-i}.
\]

(21)

Now the shape of the mechanism is emerging. The final step is to “push” \( \bar{y}_i \) and \( y_i \) to the desired location. Actually, we will show \( \bar{y}_i \geq \bar{x}_i \) and \( y_i \leq x_i \) by contradiction. Suppose \( y_{-i} > x_{-i} \) for example. From monotonicity of \( \psi^*_i \), for any \( s_{-i} \leq y_{-i} \) we have \( \psi^*_i(s_{-i}) \leq \psi^*_i(y_{-i}) \). Equation 2 indicates that \( a_i(1, s_{-i}) = \psi^*_i(s_{-i}) \). Using Equation 21, we have:

\[
a_i(1, s_{-i}) \leq g(s_{-i}) = \begin{cases} 
   d_i^*(y_{-i}) & \text{if } s_{-i} \leq y_{-i}; \\
   d_i^*(s_{-i}) & \text{if } s_{-i} \geq y_{-i}.
\end{cases}
\]

the form of function \( g(s_{-i}) \) is exactly the form in Lemma B.3. This enables us to pin down \( c^*_i(y_{-i}) \) as the unique lower bound of the one-sided optimizers of \( g \) by Lemma B.3. Then the monotone property established in Lemma B.1 yields \( y_{-i} = \phi^*_i(1) \leq c^*_i(y_{-i}) < y_{-i} \). This is a contradiction, so \( y_{-i} \leq x_{-i} \).

Finally, since \( \phi^*, \psi^*, c^* \) and \( d^* \) are non-decreasing, we have:

\[
\phi_i^*(\bar{y}_{-i}) \leq \phi_i^*(1) = y_i \leq x_i = c_i^*(x_{-i}) \leq c_i^*(\bar{y}_{-i}) = \phi_i^*(\bar{y}_{-i}).
\]

Thus all the inequalities above are equalities. Similarly from the other three equation systems

\footnote{We can assume w.l.o.g that \( \phi^*_i(1) \) is the lower bound of the one-sided optimizer, otherwise just take \( s_i \) sufficiently close to 1.}
we can conclude:

\[ \phi_i^\#(\bar{y}_{-i}) = \phi_i^\#(1), \quad \psi_i^\#(y_{-i}) = \psi_i^\#(0) \]
\[ y_i = \bar{x}_i, \quad \bar{y}_i = \bar{x}_i. \]

These guarantee the optimal mechanism to be the desired one. \( \square \)

**Proof of Proposition 4.7.** We will prove the proposition by showing \( a_i(s_i, s_{-i}) \) is a one-sided optimizer of \( a_{-i} \) point wisely in the general sense:

\[ \phi(s_{-i}), \psi(s_{-i}) \in \arg \max_{D_i} \int_0^1 -(\bar{a}_i - a_{-i}(s_i, s_{-i}))^2 - \lambda_i(\bar{a}_i - s_i)^2 dF_i(s_i) \]
\[ \bar{a}_i(s_i) \in \arg \max_{a \in D_i} U^*_A(a, s_i). \]

Since we have already proven that \( [\phi_{-i}, \psi_{-i}] \) is the optimal interval delegation of the above maximization problem, all we need to show is the optimality of interval (potentially degenerate) delegation. We will apply the Proposition 1 in Amador and Bagwell (2013) to complete our proof. To help better mapping between our setting and theirs, we will rewrite primitives as: 18

\[ w(s_i, \bar{a}_i) = -(\bar{a}_i - a_{-i}(s_i))^2 - \lambda_i(\bar{a}_i - s_i)^2; \]
\[ U^*_A(s_i, \bar{a}_i) = s_i \bar{a}_i - \frac{1}{2} \bar{a}_i^2 = s_i \bar{a}_i + b(\bar{a}_i); \]
\[ \kappa = \frac{w_{aa}(s_i, \bar{a}_i)}{b'(\bar{a}_i)} = 2(\lambda_i + 1); \]
\[ \pi_f(s_i) = s_i. \]

Now we will verify the three conditions (c1), (c2), and (c3) in Amador and Bagwell (2013). This will guarantee the optimality of interval delegation. First we want to prove

\[ (c1) \quad g(s_i) = \kappa f_i(s_i) - w_a(s_i, \pi_f(s_i)) f_i(s_i) \text{ is non-decreasing for all } s_i \in [\phi_i(s_{-i}), \psi_i(s_{-i})], \]

with

\[ \frac{\partial g}{\partial s_i} = 2(\lambda_i + 1) f_i(s_i) + 2(s_i - a_{-i}(s_i)) f'_i(s_i) + 2(1 - \frac{\partial}{\partial a_{-i}} a_{-i}(s_i)) f(s_i) \]
\[ \geq 2(\lambda_i + 1) f_i(s_i) + 2(s_i - a_{-i}(s_i)) f'_i(s_i). \]

---

18 Here we choose a parametrized representation of agent’s preference, but any preference in our setting is strategically the same as the one we choose.
The inequality comes from the fact that the optimal boundary functions are Lipschitz continuous with parameter 1. When \( f_i' = 0 \), the inequality holds for sure. To verify (c1) we need to show:

\[
\frac{f_i(s_i)}{f_i'(s_i)}(\lambda_i + 1) \geq a_{-i}(s_i, s_{-i}) - s_i, \quad \forall s_i \in [\phi_i(s_{-i}), \psi_i(s_{-i})].
\]

(22)

The left hand side is increasing while the right hand side is decreasing. Note

\[
\frac{f_i(x_i)}{f_i'(x_i)}(\lambda_i + 1) \geq F_i(x_i)(\lambda_i + 1) > a_{-i}(\bar{x}_i) - \bar{x}_i \geq a_{-i}(x_i, s_{-i}) - x_i,
\]

where the second inequality has been proved in the proof of Lemma B.3. The verification of (c1) is thus completed since we have \( \phi_i(s_{-i}) \geq \bar{x}_i \). To verify (c2), we need:

\[
(\text{c2}) (s_i - \psi_i(s_{-i})) \kappa \geq \int_{s_i}^{1} w_a(s, \pi_f(\psi_i(s_{-i}))) \frac{f_i(s)}{1 - F_i(s_i)} ds, \quad \forall s_i \in [\psi_i(s_{-i}), 1].
\]

Equivalently, \( (1 - F_i(s_i))(\lambda_i + 1)(s_i - \psi_i(s_{-i})) \geq \int_{s_i}^{1} (a_{-i}(s, s_{-i}) - \psi_i(s_{-i}) + \lambda_i(s - \psi_i(s_{-i}))) f_i(s) ds \)

or \( h(s_i) = \int_{s_i}^{1} ((\lambda_i + 1)s_i - a_{-i}(s, s_{-i}) - \lambda_i s_i) f_i(s) ds \geq 0 \quad \forall s_i \in [\psi_i(s_{-i}), 1]. \)

This is true since \( \frac{\partial h}{\partial s_i} = -(s_i - a_{-i}(s_i, s_{-i})) f_i(s_i) + (\lambda_i + 1)(1 - F_i(s_i)) \)

\[
= f_i(s_i)((\lambda_i + 1)\frac{1 - F_i(s_i)}{f_i(s_i)} - (s_i - a_{-i}(s_i, s_{-i}))).
\]

Note that inside the bracket the expression is decreasing with respect to \( s_i \). Therefore, either \( \partial h/\partial s_i \) is always negative, or it is first positive and then negative. Since \( [\phi_i(s_{-i}), \psi_i(s_{-i})] \) is the optimal among all interval delegation, we know from Lemma 1 in Amador and Bagwell (2013) that \( h(\psi_i(s_{-i})) = 0 \). Plus, \( h(1) = 0 \) by definition and we know \( h(s_i) \geq 0 \quad \forall s_i \in [\psi_i(s_{-i}), 1]. \) The proof for (c3) is similar.

\[ \square \]

C Proofs in Section 5

Proof of Proposition 5.3. We prove it by contradiction. Suppose \( \exists s_0 \text{ s.t. } c_i^{*\lambda_i}(s_0) < c_i^{*\hat{\lambda}_i}(s_0) \), according to Equation 9, we have:

\[
c_i^{*\lambda_i}(s_0) + \lambda_i \int_{0}^{c_i^{*\lambda_i}(s_0)} \frac{F_i(s)}{F_i(c_i^{*\lambda_i}(s_0))} ds = c_i^{*\hat{\lambda}_i}(s_0) + \hat{\lambda}_i \int_{0}^{c_i^{*\hat{\lambda}_i}(s_0)} \frac{F_i(s)}{F_i(c_i^{*\hat{\lambda}_i}(s_0))} ds;
\]

38
Therefore:

\[
\int_0^{c_i^{x,s}(s_0)} \frac{F_i(s)}{F_i(c_i^{x,s}(s_0))} \, ds > \frac{\bar{\lambda}_i}{\lambda_i} \int_0^{c_i^{x,s}(s_0)} \frac{F_i(s)}{F_i(c_i^{x,s}(s_0))} \, ds \geq \int_0^{c_i^{x,s}(s_0)} \frac{F_i(s)}{F_i(c_i^{x,s}(s_0))} \, ds.
\]

However,

\[
\frac{d}{dx} \left( \int_0^x \frac{F_i(s)}{F_i(x)} \, ds \right) = \frac{F_i^2(x) - f_i(x) \int_0^x F_i(s) \, ds}{F_i^2(x)} > 0.
\]

Thus we get \( c_i^{x,s}(s_0) > c_i^{x,s}(s_0) \), which is a contradiction. The other side of the proof is similar.

**Proof of Proposition 5.4.** Denote \((\bar{x}_{-i}, \bar{x}_i)\) and \((\bar{x}_-i, \bar{x}_i)\) as the joint coordinated bound of the optimal mechanism under relative importance parameters \(\lambda \kappa_i\) and \(\nu \kappa_i\), respectively. We first claim that if \(\bar{x}_{-i} \leq \bar{x}_-i\), then \(\bar{x}_i \geq \bar{x}_i\). In fact, from Equation 13, for any \(\lambda\) we have:

\[
x_i^\lambda = c_i^{x_i}(\bar{x}_i) = \bar{x}_i - \lambda \kappa_i \int_0^{\bar{x}_i} \frac{F_i(s)}{F_i(x_i^\lambda)} \, ds;
\]

\[
\bar{x}_i = d_i^{x_i}(\bar{x}_i) = x_i + \lambda \kappa_i \int_{\bar{x}_i}^{1} \frac{1 - F_i(s)}{1 - F_i(\bar{x}_i)} \, ds.
\]

Adding the two above equations yields:

\[
\kappa_i \int_0^{\bar{x}_i} \frac{F_i(s)}{F_i(x_i^\lambda)} \, ds - \kappa_i \int_{\bar{x}_i}^{1} \frac{1 - F_i(s)}{1 - F_i(\bar{x}_i)} \, ds = g(\bar{x}_i) - h(\bar{x}_i) = 0.
\]

And according to Equation 23, \(g(x)\) is strictly increasing while \(h(x)\) is strictly decreasing, which leads to our claim.

For \((\bar{x}_-i, \bar{x}_i)\) and \((\bar{x}_-i, \bar{x}_i)\), according to our claim, either \(\bar{x}_-i \geq \bar{x}_-i\) and \(\bar{x}_i \leq \bar{x}_i\), or \(\bar{x}_-i \leq \bar{x}_-i\) and \(\bar{x}_i \geq \bar{x}_i\). However, since Proposition 5.3 guarantees that \(c_i^{x} \geq c_i^{x}\) and \(d_i^{x} \leq d_i^{x}\), the only possible situation is \(\bar{x}_-i \leq \bar{x}_-i\) and \(\bar{x}_i \geq \bar{x}_i\), which together with \(c_i^{x} \geq c_i^{x}\) and \(d_i^{x} \leq d_i^{x}\), leads to the desired result.

**Proof of Proposition 5.7.** We prove the first part of the proposition by contradiction. Assume \(\bar{F}_i\) dominates \(F_i\) in lower tail expectation and suppose there is a \(s_0\) such that \(c_i^{x}(s_0) > c_i^{x}(s_0)\).
Denote:
\[ G(c) \triangleq \mathbb{E}(S_i|S_i \leq c) = c - \frac{\int_0^c F_i(s)ds}{F_i(c)}; \]
satisfying \( \frac{\partial G}{\partial c}(c) = \frac{f_i(c) \int_0^c F_i(s)ds}{F_i(c)^2} < 1. \)

Using Equation 11, we have:

\[
(\lambda_i + 1)c_i^*(s_0) - \lambda_i \mathbb{E}(S_i|S_i \leq c_i^*(s_0)) = (\lambda_i + 1)c_i^*(s_0) - \lambda_i \mathbb{E}(\bar{S}_i|\bar{S}_i \leq \bar{c}_i^*(s_0));
\]

\[
(\lambda_i + 1)(c_i^*(s_0) - \bar{c}_i^*(s_0)) - \lambda_i(\mathbb{E}(S_i|S_i \leq c_i^*(s_0)) - \mathbb{E}(\bar{S}_i|\bar{S}_i \leq \bar{c}_i^*(s_0))) \leq 0;
\]

\[
(\lambda_i + 1)(c_i^*(s_0) - \bar{c}_i^*(s_0)) \leq \lambda_i(\mathbb{E}(S_i|S_i \leq c_i^*(s_0)) - \mathbb{E}(\bar{S}_i|\bar{S}_i \leq \bar{c}_i^*(s_0)));
\]

\[
(\lambda_i + 1)(c_i^*(s_0) - \bar{c}_i^*(s_0)) < \lambda_i(c_i^*(s_0) - \bar{c}_i^*(s_0));
\]

which contradicts with \( c_i^*(s_0) > \bar{c}_i^*(s_0). \)
References


Amador, Manuel, Kyle Bagwell, and Alex Frankel. 2018. “A Note on Interval Delegation.”


