Who Versus When: Designing Decision Processes in Organizations

Roi Orzach*

August 24, 2022

Abstract This paper analyzes concurrent versus sequential decision-making in a model where two agents first communicate and then make decisions, attempting to both adapt to their local conditions and coordinate with their partner. Sequential decision-making improves overall information sharing. However, first movers also have an incentive to over-adapt to their state, knowing second movers will conform to their decision. Sequential decision-making is optimal if and only if the two units' local conditions have sufficiently different volatilities or if their need to coordinate is sufficiently asymmetric. Finally, this paper discusses the empirical literature on the optimal timing of decisions within firms through the lens of the model. (JEL D21, D23, D82)

1 Introduction

The organizational economics literature often analyzes who should make decisions in an organization (Aghion and Tirole (1997), Baker et al. (2002), Grossman and Hart (1986), Hart and Moore (1990)). Less focus has been given to the order in which those decisions should be made. For example, consider a firm that produces complementary products, each of which are produced and marketed by a different division. Because of private information, it may be optimal for each of these divisions to set their own prices. When should the divisions set prices simultaneously and when should they set prices sequentially? Many

*Department of Economics, Massachusetts Institute of Technology; orzach@mit.edu
I am indebted to Robert Gibbons for comments and feedback. I am very thankful for helpful conversations with Charles Angelucci, Glenn Ellison, Heikki Rantakari, Kramer Quist, and seminar participants at the MIT Organizational Economics Lunch. All remaining errors are my own.

theoretical models cannot answer this question because they take the timing of decision-making as fixed.

This question of whether firms should use sequential or concurrent decision-making extends beyond the pricing example given above and is one of long standing and extensive study in the management literature. For example, Hayes and Wheelwright (1979) analyze when manufacturing firms should use concurrent or sequential production plants and suggest that firms with non-standardized products should use concurrent production. However, while Hayes and Wheelwright (1979) restrict attention to concurrent or sequential manufacturing, a similar question exists throughout the entire production cycle of a product. Takeuchi and Nonaka (1986) analyze this question and suggest that only firms involved in breakthrough innovations, such as chemical research firms, would benefit from developing their products sequentially. Section 4 analyzes how the predictions of Hayes and Wheelwright (1979) and Takeuchi and Nonaka (1986) can be interpreted within the coordinated adaptation model described below.

In this paper I address two questions: when should decisions be made sequentially, and in sequential cases, who should go first? I analyze these timings in a model of coordinated adaptation which consists of two self-interested units and a surplus-maximizing headquarters. A unit's profit depends on how well aligned their decision is to their privately-known local condition (adaptation) and to the other unit's decision (coordination). Before making their decisions, the units engage in strategic communication about their local conditions. Units may asymmetrically weight coordination and adaptation.

The organizational economics literature on multi-divisional firms has focused on who should make decisions rather than when the decisions should be made by a fixed set of decision-makers. For instance, Alonso et al. (2008) and Rantakari (2008) study concurrent decision-making using this model to determine the optimal allocation of decision rights across divisions and headquarters. These papers note that decentralization (i.e. units make decisions themselves), as opposed to centralization (i.e. headquarters makes all decisions), allows for better communication between units, but each unit makes their decision without internalizing its effect on their partner.

However, in keeping with the adage that actions speak louder than words, staggering the decisions under decentralization allows the unit that moves second to perfectly observe the decision of the first mover. This is in contrast to concurrent decision-making, where units communicate with only strategic communication. Hence, under concurrent decision-making, information is imperfect, which causes units to be unsure of the exact decision the other unit will make. Under sequential decision-making, by contrast, the leader has no such incentive to send a biased message since their decision is always fully revealed, but they do have an incentive to make a biased decision (e.g., excessively adapt to their local state). Analyzing when the additional informational gain due to revealing the leader's decision outweighs the loss from biased decisions creates an additional governance choice for firms.

Without this additional choice, previous analyses conclude centralization is optimal after comparing it to only concurrent decentralization and thus over-prescribe centralization as a solution to coordinated-adaptation problems. Instead, my model suggests decentralization with sequential decision-making can out-perform centralized decision-making.

An extensive literature exists on sequential versus simultaneous contributions to public goods and analyzes which yields more contributions. This literature is fundamentally related to the coordination versus adaptation tension within an organization, since each unit does not internalize the fact that their decision to coordinate positively affects the other unit. Admati and Perry (1991) show that, in a threshold public-goods game, some projects that would have been funded if donations had occurred concurrently do not get funded with sequential donations.[1] In fact, Varian (1994) shows that in a continuous public goods setting, weakly smaller contributions will occur when they happen sequentially. However, unlike the literature on organizations, this literature focuses on cases with public information about what the public good is. With public information about which state to coordinate on, coordination is a public good. However, within organizations each unit may not know which state their partner wishes to coordinate on. This uncertainty implies that the intuition from these public-goods papers cannot be blindly extended into the setting of an organization since these papers lack private information. With private information, information quality increases under a sequential structure because the follower can perfectly observe the decision of the leader, which their models do not capture.

This paper connects to the literature in Organizational Economics on coordinated-adaptation trade-offs such as Dessein and Santos (2006) and, especially, Alonso et al. (2008) and Rantakari (2008) on the choice of the optimal governance structure of a firm via alloca-

[1]If the donations are strategic compliments, as in the case of Akerlof and Holden (2019), then sequential donations can be optimal.

tion of decision rights.[2] However, none of these papers gives predictions as to when decisions should be made concurrently or sequentially. In an extension, Alonso et al. (2008) does compare sequential decision-making to centralization. However, since timing is not the focus of their paper, they do not compare concurrent decentralization to sequential decentralization. In contrast to their paper analyzing if centralization or decentralization is optimal given the different timings, I first analyze which timing is optimal, conditional on being decentralized. Next, I analyze how the comparison between centralization and decentralization changes with the use of the optimal version of decentralization. Additionally, since Alonso et al. (2008) consider symmetric units, they are unable to answer which unit should go first when decisions are sequential. Finally, Rantakari (2013) shows that decentralized firms do better in volatile environments. However, Rantakari (2013) does not give insights into whether the decentralization should be sequential or concurrent, which is the focus of this paper. This paper, in contrast, shows in volatile environments firms prefer sequential decision-making.

The paper which answers a question most similar to mine is Lewis and Mistree (1997), which looks at a game in which one unit chooses the aircraft length and the other chooses the wing length. The paper then analyzes when the aircraft is better designed: when the units decide concurrently or sequentially. The model in Lewis and Mistree (1997) does not contain any private information and thus reaches a similar conclusion to the public-goods literature that concurrent decision-making is always better. However, Lewis and Mistree (1997) get closest to the idea of decision-making orders within an organization.

Further, this paper connects to a literature in Industrial Organization on pricing behavior within firms. For instance, Hortaçsu et al. (2021) analyzes the interactions between various sub-units in the pricing decision of an airline showing inefficiencies due to sequential decision-making. Additionally, Ellison (2005) analyzes a model of "Add-On Pricing" in which prices for complementary products are revealed only after the primary product is sold. While Ellison (2005) rationalizes this through a competitive search market, this process can also be rationalized due to organizational rather than competitive forces. Further, Moorthy and Png (1992) argues that firms may release high quality products before low quality products (e.g., flagship phones are released before budget models) to prevent product cannibalization. However, the coordination-adaptation intuition offers an alternative

[2]The literature on delegation, such as Aghion and Tirole (1997), is also relevant. However, these papers traditionally have one decision right, and thus cannot study optimal decision-timing.

explanation: for example, the budget unit has a large gain from moving after the flagship unit as coordinating on as many features as possible drives down R&D costs.

Section 3 analyzes the model under three different information structures: public information, incomplete information without cheap talk, and incomplete information with cheap talk. I first show that the results in Varian (1994) and Lewis and Mistree (1997), which assume public information, generate the same prediction as this coordinated adaptation model: namely, the firm would always have units decide concurrently if there was public information. I next show that if there is private information and the units are unable to communicate with each other, the units should always decide sequentially. After going through these benchmarks, I solve the model under incomplete information with cheap talk and show that, unlike the benchmarks, the solution depends on the parameters of the organization. In this case, firms with asymmetric units are most likely to benefit from asymmetric (sequential) timings.

Section 4 shows that despite sequential decision-making allowing for more information transmission, if the units place high and similar weights on coordination and have similar local volatilities, then it is optimal to decide at the same time; otherwise, it is almost always optimal for the unit that cares more about coordination to decides second. I next return to the predictions from Hayes and Wheelwright (1979) and Takeuchi and Nonaka (1986) and show how to analyze them through the model's implications. Additionally, Section 4 shows that, by not considering sequential decentralization, previous analyses have prescribed centralization in instances in which sequential decentralization is, in fact, optimal. Finally, Section 5 states conclusions, and the Appendix contains extensions and proofs for all statements not proved in the text.

2 Model

There are three players: a headquarters and two units (1 and 2). Each unit, i, has a local condition denoted by θ_i, which is private information. In keeping with the literature, I assume θ_1 and θ_2 are independently distributed such that $\theta_i \sim U\left[-\bar{\theta}_i, \bar{\theta}_i\right]$.[3] I define the volatility of unit i to be the variance of its local condition, $V(\theta_i)$. Unit i wishes to match

[3] The analysis in Section 3.1 of public information and private information without cheap talk is done without the distributional assumption.

their decision, d_i, to both their local condition, θ_i, and the other unit's decision, d_{-i}. These preferences are described by the loss function:

$$\mathcal{L}_i(d_i, d_{-i}, \theta_i) = (1 - r_i)(\theta_i - d_i)^2 + r_i(d_{-i} - d_i)^2. \tag{1}$$

Here, $(\theta_i - d_i)^2$ and $(d_{-i} - d_i)^2$ represent the adaptation and coordination loss, respectively. Moreover, $r_i \in (0, 1)$ measures the weight unit i places on coordination as opposed to adaptation.[4] Notice that each unit might weight these two losses asymmetrically. The headquarters' loss is the sum of the losses for the units (hereafter referred to as surplus).[5] For now, the only decision the headquarters makes regards the timing of the decisions made by units 1 and 2. For most of the analysis, unit i chooses their own d_i. In Section 4, I describe how the analysis changes if headquarters is allowed to make the decisions.[6]

In the main analysis, the headquarters chooses between two different governance structures: sequential decision-making and concurrent decision-making. In both structures, each unit observes their θ_i and then the units engage in one round of simultaneous cheap talk communication by sending messages, m_i. However, the governance structures differ in the timing of their decisions:

- **Sequential Decision-Making:** Unit j chooses d_j. Upon observing d_j, unit i makes their decision d_i.

- **Concurrent Decision-Making:** Units 1 and 2 choose d_1 and d_2 simultaneously.

The solution concept employed is Perfect Bayesian Equilibrium. I focus on the most informative equilibrium of the cheap talk communication game, as is standard in the literature. In my setting, this is the Pareto dominant equilibrium. The timing is summarized in Figure 1.

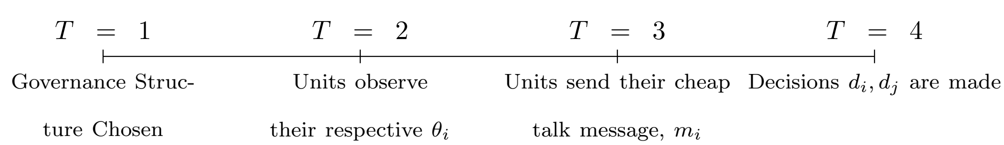


Figure 1: Timing of the Game

[4]This is simply a normalization and captures cases where the coordination weight is arbitrarily large.

[5]Alonso et al. (2008) and Rantakari (2008) allow for the headquarters to weight the units asymmetrically, however this does not change the analysis presented herein.

[6]Section 4 also allows for the headquarters to place asymmetric weights on the two units.

3 Organizational Performance

In this section, I compare the organizational performance of sequential and concurrent decision-making under three information structures. Section 3.1 analyzes the model first assuming the units' local conditions are public information, and then assuming the local conditions are private information and the units cannot communicate. Finally, Section 3.2 shows that when the units' local conditions are private information and the units can communicate, the optimal timing depends on the volatility and coordination weights of the units.

3.1 Benchmarks

I begin by analyzing the model under two benchmark information structures: public information and private information without the possibility to engage in communication. All computations can be found in Sections 6.3 and 6.4 in the Appendix.

Under public information, the headquarters always prefers concurrent decision-making. If instead the timing was sequential, the first mover would over-adapt to their local state, knowing the second mover would coordinate with this observed decision at the expense of adapting to their own local state. Thus, as a whole, with sequential decision-making the firm would over-adapt to the first movers local state.

Meanwhile, without knowledge of their partner's local state, a unit's best option for coordination is choosing a decision closer to 0. However, with sequential decision-making, as opposed to concurrent, perfect transmission of the decision of the first mover always occurs. This transmission allows units to coordinate on states that would yield lower adaptation loss as opposed to choosing a decision closer to 0. This effect is strongest when the follower has incentives to coordinate with the leader. Thus conditional on utilizing sequential decision-making, it is optimal to have the follower be the unit that values coordination more. Further, this ability to coordinate without sacrificing adaptation loss outweighs the first-mover effect as stated in the following remark:

Remark 1 *With public information, the optimal governance structure is to have the units decide concurrently. By contrast, with private information but without the ability to communicate, the optimal governance structure is to have the unit that cares more about coordination move second.*

The next subsection analyzes the main information structure of interest, in which local states are private information but the units can communicate with each other.

3.2 Cheap Talk Communication Analysis

I solve the game by backwards induction and calculate the decisions made by the units after having engaged in one round of simultaneous cheap talk communication. Under sequential decision-making, given the message sent and the decision made by the leader, the follower ignores the message, because only the leader's decision is payoff-relevant to the follower. It is without loss of generality to assume that, under sequential decision-making, unit 1 is the leader and unit 2 the follower. Hence assuming that $m_1 = \theta_1$ is without loss.[7] For notational simplicity, I define $r_1(1-r_2)^2 := \beta_1$. This is a measure of unit 1's need to coordinate with unit 2. Note that as r_1 increases, unit 1 prefers to coordinate more. However, when r_2 increases, unit 1 knows that unit 2 will coordinate more and thus unit 1's need to coordinate decreases. Given this notation, one can concisely write the decisions under both timings.

Lemma 1 *Equilibrium decisions under sequential decision-making are*

$$d_1 = \frac{1-r_1}{(1-r_1)+\beta_1}\theta_1 + \frac{\beta_1}{(1-r_1)+\beta_1}\mathbf{E}(\theta_2|m_2),$$
$$d_2 = (1-r_2)\theta_2 + r_2 d_1,$$

whereas under concurrent decision-making they are

$$d_i = (1-r_i)\theta_i + r_i\mathbf{E}(d_j|m_j).$$

Note that unit 2's decision under sequential decision-making does not involve expectations because they have already observed the decision of unit 1, and hence do not care about either θ_1 or m_1. However, with concurrent decision-making or for unit 1 under sequential decision-making, decisions are a function of posterior beliefs generated after receiving the message from the other unit.

Looking at the coefficient on $\mathbf{E}(\theta_2|m_2)$ in the decision of unit 1 with sequential decision-making, which will be denoted by $\gamma_1 := \frac{\beta_1}{(1-r_1)+\beta_1}$, one can see that the denominator is less

[7] It is equivalent to assume $m_1 = \emptyset$, but the assumption $m_1 = \theta_1$ allows a definition of information loss consistent with the literature. However, any cheap talk communication equilibrium yields the same uncertainty about θ_1 which is what the definition encompasses.

than 1. This is the same denominator on the coefficient of θ_1 in decision 1 with sequential decision-making. This inequality implies that the leader is adapting to their state more than just $(1 - r_1)$, which is how much they would adapt under concurrent decision-making. In other words, the leader is adapting to their local state more than if they were a follower or if decisions were made concurrently.

Neither unit fully coordinates with their partner under either governance structure. This, in turn, causes communication from the follower under sequential decision-making or from either unit under concurrent decision-making to be imperfect. Perfect communication cannot occur in equilibrium because if there were perfect communication, a unit would then exaggerate their type to get perfect coordination with no adaptation cost to them. In what follows, I look for a cheap talk partition in the Crawford and Sobel (1982) sense such that a unit would prefer to correctly signal the interval of their type.

Take sequential decision-making. Given the decision rule of the leader, I find an information partition for the follower that induces truthful reporting. Namely, I am looking for a partition $k_1, \ldots, k_n$ of $\left[-\bar{\theta_2}, \bar{\theta_2}\right]$ such that the unit would always prefer to truthfully report the partitional set containing their type. I then take the limit as the number of partitional elements becomes arbitrarily large.[8] Given this limit partition, I define information loss about state i to be $c_i = \mathbf{V}(\theta_i - \mathbf{E}(\theta_i|m_i))$, where $\mathbf{V}$ denotes the variance across all realizations of θ_i, and hence m_i. Information loss on state i is unit j's expected uncertainty about θ_i before unit j makes their decision.[9] As the informational partitions become finer, the range of potential states given a message decreases; hence, better information can be viewed as a finer partition. The analysis for concurrent decision-making is very similar and follows Rantakari (2008). The next lemma calculates the information loss for both governance structures.

Lemma 2 *Equilibrium information loss, c_i, about state i under decision timing t is $V(\theta_i)\frac{1}{3\phi_{i,t}+4}$ where $\phi_{i,concurrent} = \frac{r_j(1-r_j)}{(1-r_j)}$, $\phi_{2,\ sequential} = \frac{\beta_1}{(1-r_1)} = \frac{r_1(1-r_2)^2}{(1-r_1)}$, and $\phi_{1,\ sequential} = \infty$.*

The term $V(\theta_i)$ represents the variance of the state being communicated, and the next

[8] Here, the partition must satisfy the following difference equation $k_{n+1} - k_n = k_n - k_{n-1} + \frac{4k_n}{\phi}$. For formal details, I refer the reader to Alonso et al. (2008) or Rantakari (2008) and Appendix Section 6.1.

[9] As will be shown in Lemma 4, these terms directly enter the loss functions of the units. If the preferred interpretation of communication under sequential decision-making is $m_i = \emptyset$, then define information loss to be zero, since the (observed) decision of the leader perfectly reveals the state.

term is the percentage of the total information lost due to imperfect communication.[10] One can see that because $\phi \in (0, \infty)$, the percentage lost is bounded between 0 and $\frac{1}{4}$. It is bounded by $\frac{1}{4}$ because no matter how strong the incentives are to misrepresent, a unit can always correctly communicate whether their state is positive or negative. Achieving perfect information transmission is equivalent to $\phi = \infty$. Given the information loss, the next lemma states expected total loss for the leader and follower under sequential decision-making.[11]

Lemma 3 *Under sequential decision-making, equilibrium expected total loss for the leader is*

$$\mathcal{L}_{Leader} = (1 - r_1)\gamma_1 \Big(\mathbf{V}(\theta_2) + \mathbf{V}(\theta_1)\Big) + \gamma_1 \beta_1 c_2. \tag{2}$$

Similarly, equilibrium expected total loss for the follower is

$$\mathcal{L}_{Follower} = (1 - r_2) r_2 (1 - \gamma_1)^2 \Big(\mathbf{V}(\theta_2) + \mathbf{V}(\theta_1)\Big) - (1 - r_2) r_2 (\gamma_1^2 - 2\gamma_1) c_2. \tag{3}$$

The total loss for each unit has two components: the loss that would occur even with perfect information transmission and the additional loss stemming from information loss due to strategic communication.[12] Note that whereas under sequential decision-making there is only information loss about the follower's state, with concurrent decision-making there is information loss associated with both states. Combining the statements in Rantakari (2008) about the concurrent case with the notation defined throughout yields the following Lemma.

Lemma 4 *The expected total losses to unit 1 when both units decide concurrently can be summarized as follows:*

$$\begin{aligned}\mathcal{L}_{concurrent} = {} & (V(\theta_1) + V(\theta_2)) \frac{(1 - r_1)\beta_1}{(1 - r_1 r_2)^2} + V(\theta_1) r_1 (1 - r_1) \frac{(1 - r_1 r_2)^2 - (1 - r_1)^2}{(1 - r_1 r_2)^2} V\left(\phi_1^{sim}\right) \\ & + V(\theta_2) r_1 (1 - r_2)^2 \frac{(1 - r_1 r_2)^2 - (1 - r_1)}{(1 - r_1 r_2)^2} V\left(\phi_2^{sim}\right)\end{aligned}$$

where $V(\phi) = \dfrac{1}{4 + 3\phi}, \phi_i = \dfrac{\beta_1}{1 - r_1}, \gamma_1 = \dfrac{\beta_1}{1 - r_1 + \beta_1},$ *and* $\beta_1 = r_1(1 - r_2)^2$.

I define coordination loss as the expected loss due to miscoordination, $\mathbf{E}((d_1 - d_2)^2)$. There are two aspects to coordination: knowing which state to coordinate on and being

[10] Given the uniform assumption on the local states, $V(\theta_i) = \frac{\bar{\theta_i}^2}{3}$.

[11] Recall, the headquarters loss is the sum of the losses for the two units.

[12] The coefficients on information loss are determined by how sensitive each unit's decision is to the message sent in equilibrium. In Rantakari (2008) these coefficients may at times be negative; however, in this paper's setting the leader's coefficient is positive since $\gamma_1\beta_1 > 0$. The follower's is also positive since $\gamma_1 < 1$, and hence the term in front of c_2 is positive.

willing to do so. For this next proposition, I hold the first aspect fixed and compare the coordination losses across the decision-making timings fixing the information loss. Namely, I assume that under all decision timings the informational partitions are exogenously given, equal, and not subject to truth telling constraints. The proposition below compares this loss under the various decision timings.

Proposition 1 *For a fixed information loss the following are true:*

i Concurrent decision-making always yields less coordination loss than sequential decision-making irrespective of which unit moves first.

ii Coordination loss decreases when unit 1 moves first as opposed to unit 2 if and only if $r_1 < r_2$.

The intuition for the first statement parallels Remark 1: when a unit acts as a leader, they over-adapt to their local conditions which increases coordination loss because the follower is not willing to coordinate on this overly self-interested decision. The intuition for the second statement in the proposition is that because the leader has an incentive to not coordinate, the ideal leader is one who would not have coordinated much anyway.[13] However, coordination loss for a fixed information loss is only one determinant of the total losses. The information loss is itself an endogenous object and enters the loss functions for each unit. The next proposition states that sequential decision-making induces less aggregate information loss, defined as the sum of the information loss on state 1 and state 2. Moreover, the unit that has a higher need to adapt to local conditions or has more volatile local conditions should move first.

Proposition 2

i There exists a way to order the units such that sequential decision-making induces less aggregate information loss than concurrent decision-making despite worse information loss about the follower's state.

ii Given equal volatility for the two units, there is less aggregate information loss when unit 1 leads as opposed to unit 2 if and only if $r_1 > r_2$.

[13]This proposition does not rely on the distributional assumption of θ. Additionally, this proposition holds not just in expectation over θ but for any realization of $m_1(\theta_1)$ and $m_2(\theta_2)$ in equilibrium.

iii If the units have equal preferences over coordination, namely $r_1 = r_2$, there is less aggregate information loss when unit 1 leads as opposed to unit 2 if and only if $V(\theta_1) > V(\theta_2)$.

The intuition for the first statement is that sequential decision-making always yields perfect information transmission about the leader's state and hence there will be at least one way to order the two units to improve upon concurrent decision-making.[14] Meanwhile, despite having lower aggregate information loss, the information loss for the follower is higher because the leader over-adapts to their state relative to concurrent decision-making, causing the follower to want to exaggerate their message and yielding worse information on their state.

The second statement follows from the logic that when units do not care about coordination, they have less of an incentive to influence the decisions of the other unit. This causes information loss to decrease when units that do not care about coordination, i.e. have a low r_i, follow and thus are the ones communicating with cheap talk messages as opposed to their decision, d_i.

Finally, recall that by letting a unit lead there is no welfare loss stemming from uncertainty about their state. Hence, if the two units have the same incentive to send biased messages, letting the unit with more information needing to be transmitted communicate their decision perfectly via revelation of their decision will reduce aggregate information loss. The next section reports on whether the effects in Proposition 1 (i.e. that concurrent decision-making reduces coordination loss conditional on information loss) or 2 (i.e. that sequential decision-making reduces information loss) dominate as the parameters of the model vary.

4 Optimal Governance

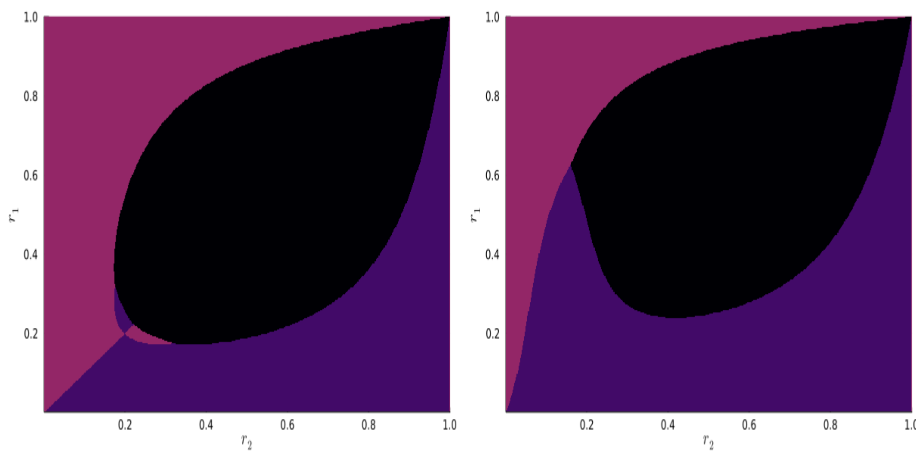
In this section I begin by characterizing the conditions under which it is optimal to have decisions be made sequentially or concurrently. The next subsection analyzes the empirical evidence on optimal decision-making timing in light of this model. The final subsection shows the results are robust to allowing decision rights to be centralized at the headquarters

[14] If the payoffs to a unit included a $(d_i - \theta_{-i})^2$ term, then perfect transmission about the leaders state yields an additional benefit. For parsimony with the literature, I do not include such a term.

level.

4.1 Relative Performance

To compute when the units should move sequentially or concurrently, one needs to check whether the effects in Proposition 1 or 2 dominate. To do so, I use numerical calculations to plot the optimal governance structure for varying parameter values.[15] In Figure 2, I compare both versions of sequential with concurrent decision-making. Figure 2a shows the optimal governance structure when both units have equal volatilities.[16] Figure 2b plots the optimal governance structure when unit 1 has twice the volatility of unit 2. In both plots, the color denotes the optimal timing with Purple corresponding to unit 1 leading, Pink unit 2 leading, and Black is concurrent.



(a) Equal volatility across units.

(b) Unit 1 has double unit 2's volatility.

Figure 2: The figure shows the governance structure that minimizes the sum of the units' losses as a function of the coordination motives of the two units, r_1 and r_2. The color denotes the optimal timing with Purple corresponding to unit 1 leading, Pink unit 2 leading, and Black is concurrent.

One should note that when r_1 and r_2 are of a similar magnitude and high the optimal

[15] As one can note from Lemmas 3 and 4 the closed form expressions for the losses do not lend for clean comparisons.

[16] In this setting the ratio of the volatilities, rather than the magnitude, determines the optimal governance.

timing is concurrent. Because first movers choose to coordinate less, as described in Proposition 1, this effect is largest when both (1) unit 1 would have coordinated if the decisions were made concurrently and (2) unit 2 attaches high value to coordination. Conditions (1) and (2) correspond to when both r_1 and r_2 are high, respectively. Hence, in this region concurrent decision-making is optimal. However, upon moving further away from the diagonal, one sees that sequential decision-making becomes optimal as now the effect of Proposition 1 is small in comparison to the informational effect described in Proposition 2.

Additionally, when r_1 and r_2 are sufficiently small, concurrent decision-making is never optimal. This result is driven by communication breaking down when both units' priority is predominantly on adaptation, and thus little information is being transmitted in equilibrium. The intuition behind communication weakening is that when coordination weights are small, unit j's decision puts little weight on unit i's message. This, in turn, gives unit i a greater incentive to exaggerate their state, which results in large informational losses. However, with sequential decision-making, no matter how uninformative communication is, having the follower see the payoff-relevant decision allows some information transmission. In fact, for sufficiently small r_1 and r_2, either sequential structure improves upon the concurrent one.[17]

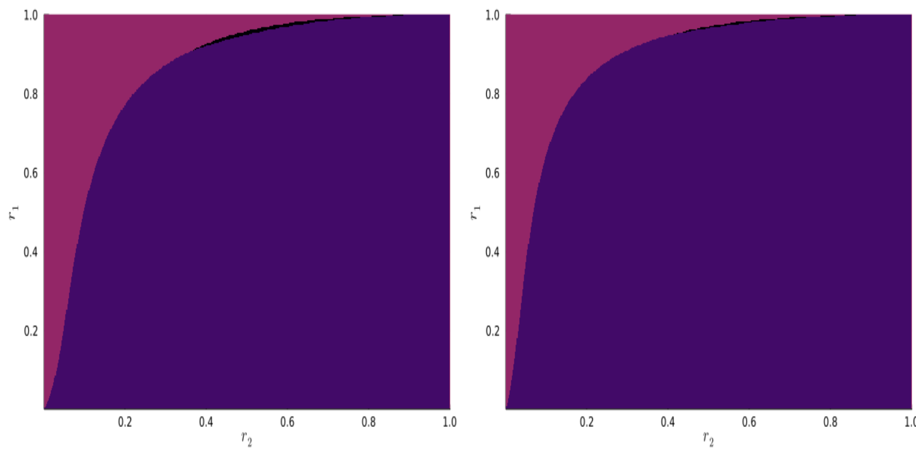
Finally, in comparing the results with equal and unequal variances one can see that when unit 1 has more volatility, the optimal governance shifts towards having unit 1 move first. This is because the coordination loss is scaled by the sum of the variances; meanwhile, information loss is scaled by only the variance of the state being communicated. Hence with unequal variances, the effect in Proposition 2 of increased information transmission becomes stronger than the effect in Proposition 1 of decreased coordination.

While Figure 2 analyzed how to minimize joint losses, Figure 3 analyzes the timing that minimizes the loss for a single unit. Without loss of generality let this be unit 1. Thus far the model has assumed that firm surplus is the sum of the two units' total losses, however some units may receive greater weights than others. For example, Apple may weight their phone division greater than the headphones division. As the weight the firm places on unit 1 increases, Figure 2 will continuously form into Figure 3.

When analyzing the total loss of a specific unit, a unit that cares sufficiently about coor-

[17] The rationale behind the small region where $r_i > r_j$ but unit i leads is discussed in Section 6.2 of the Appendix.

dination would in fact prefer to move second to ensure that they are able to coordinate with the decision of the leader. However, a unit that cares more about adaptation will prefer to move first so they can over-adapt knowing the follower will coordinate on their behalf. Additionally, as seen in the plot on the right, when unit 1's local state has more volatility, unit 1 would prefer to move first for a larger region of parameter values. This is because unit 1 now has a greater incentive to adapt to their local conditions and being a leader allows them to do so. For these reasons, with sufficiently asymmetric values of r_1 and r_2 or sufficiently asymmetric volatilities, moving to sequential decision-making is in fact a Pareto improvement upon concurrent decision-making.[18]



(a) Equal volatility across units. (b) Unit 1 has double unit 2's volatility.

Figure 3: The figure shows the governance structure that minimizes unit 1's loss as a function of the coordination motives of the units, r_1 and r_2. The color denotes the optimal timing with Purple corresponding to unit 1 leading, Pink unit 2 leading, and Black is concurrent.

However, as one can see comparing Figures 2 and 3, there exists a large region where the timing that maximizes firm surplus conflicts with what maximizes an individual unit's payoff. For instance, when both r_1 and r_2 are high and symmetric each unit wants to move first, but firm surplus is maximized when the units move concurrently. In these instances, as I describe in the conclusion, arguably one role for headquarters is to enforce the firm-optimal

[18]I.e., when $r_1 = .9$ and $r_2 = .1$, unit 1's preferred position is to lead and unit 2's is to follow. Hence, even without side payments, sequential decision-making can be a Pareto improvement.

timing of decisions.

4.2 Implications within Examples

Having derived the optimal governance structure conditional on the parameters of the model, I now return to the empirical evidence described in the introduction and show how to interpret it through this model.

Takeuchi and Nonaka (1986) suggest that research-focused firms benefit most from sequential decision-making. In terms of the model, one interpretation is that the research team is strongly motivated by the discovery of valuable working chemicals. While the sales unit has an adaptation loss, namely they may prefer to sell in less competitive markets or ones that they have an established presence, they will ultimately be more successful selling the most effective drug or chemical whichever this might be (and even if this drug is not their ideal drug to sell).[19] In the model, these situations correspond to $r_{\text{research}} \simeq 0$ and $r_{\text{sales}} \simeq 1$. As Figure 2 shows, this asymmetry is precisely when sequential development is optimal. Another natural interpretation is that a research unit faces a more uncertain environment compared to the sales unit because of the uncertainty inherent with R&D. As suggested by Figure 2, it will often be optimal for the research team to move first when $V(\theta_{\text{research}}) > V(\theta_{\text{sales}})$. Note that this finding is not driven by technological assumptions in the research industry and in some cases the sales unit moves first; for instance, for Covid vaccines the sales unit moved first to secure funding followed by the research divisions. The intuition from Takeuchi and Nonaka (1986) finds backing in Valle and Vázquez-Bustelo (2009) which uses survey evidence to show that firms involved in breakthrough innovation are more likely to utilize sequential decision-making.

Further, Hayes and Wheelwright (1979) argue that firms with non-standardized products (i.e. products tailored to each consumer) and multiple workers involved in the production process should use concurrent decision-making. An example of non-standardized products given in the article is a commercial printing company. Hayes and Wheelwright (1979) note that each printing job is unique and, therefore, if workers work sequentially first movers may make decisions that prove to be incompatible with what decision the last-mover needs to make. The predictions from Hayes and Wheelwright (1979) are tested in Safizadeh et al. (1996) which uses data from firms to show that firms with more product variants are

[19] For example, one can think of sales' local state as a user demographic where the firm has brand loyalty.

more likely to use concurrent manufacturing techniques. Turning to the model, concurrent decision-making is indeed optimal when workers' volatilities and need to coordinate are high and symmetric.

The most natural interpretation of this model is two units within a firm. However, one can also interpret each unit as a separate firm and the headquarters' choice of decision-timing as the outcome of an un-modeled bargaining game between the firms to reach the efficient timing.[20] Using this interpretation, I now turn to a case study about Chrysler and its suppliers and highlight how it relates to the mechanisms described in Propositions 1 and 2. Dyer (1996) describes how prior to the 1990s Chrysler frequently changed suppliers, leaving both types of parties hesitant to make relationship-specific investments.[21] The lack of these investments, in turn, caused neither party's profits to be too dependent on their partner's decision, i.e. r_{Chrysler} and r_{Supplier} were arguably low. At the same time, r_{Supplier} was likely greater than r_{Chrysler} because supplying for Chrysler was a large revenue stream for these suppliers, while the interaction with one specific supplier was only a small part of Chrysler's overall revenue. As one can see in Figure 2, these parameter configurations favor sequential development, with Chrysler designing the parts first and then the suppliers deciding how to make the part.

Dyer (1996) details how in the 1990s Chrysler changed their approach and decided to commit to their suppliers with the average contract length more than doubling from that prevailing in the 1980s. Given these longer contracts, Chrysler was now incentivized to "increase their investments in dedicated assets – plant equipment, systems, [and] processes" for their suppliers (Dyer:10). These investments, in turn, made miscoordination potentially more costly, as now Chrysler had become more dependent on their suppliers. Turning again to the model, this change corresponds to an increase in Chrysler's desire to coordinate ($r_{\text{chrysler}} \uparrow$) to the point where concurrent decision-making became optimal.

The new governance structure between Chrysler and their suppliers improved performance, which Dyer (1996) attributed, in part, to their improved communication. Whereas,

[20]One example of such bargaining would be Nash Bargaining (with equal weights) over the decision timing, with transfers contingent on the decision timing but not the decisions themselves.

[21]Naturally, some aspects of the relationship between Chrysler and its suppliers emphasized by Dyer (1996) are absent from the analysis (e.g., contract length and relationship-specific investments). Nevertheless, the model captures key insights about coordination and communication which are absent from models about contracts and property rights.

communication between Chrysler's suppliers and Chrysler had been minimal, the new governance structure was accompanied with numerous new communication channels which included joint meetings for manufacturers to share their opinions on proposed designs. This is consistent with Proposition 2 which states that manufacturers are able to engage in better information transmission about their needs because manufacturing now moves concurrently with design. Finally, another reason Dyer (1996) gives for the improved performance was that Chrysler and their suppliers were better able to coordinate on parts. This feature is consistent with Proposition 1 which states that units that interact concurrently have a greater incentive to coordinate.

Finally, as noted in Alonso et al. (2010), a natural interpretation of the payoffs is such that θ_i represents the deviation of the local demand state for a unit from their mean. Under this interpretation, d_i corresponds to a price decision, thus the adaptation loss measures the lost revenues for a product in its own market. Similarly, the coordination losses correspond to the fact that the goods may have inter-related costs or demands. In the context of airline pricing, Hortaçsu et al. (2021) shows information from units who choose their pricing first is not fully adapted to by units who choose pricing second. What this model suggests, as in Proposition 2, is switching to concurrent decision-making would, in fact, decrease the information transmitted. Meanwhile, within Ellison (2005), the price of the higher quality good is revealed only after search has taken place. Applying this idea to a firm, imagine the firm produces both a budget car, with low margins, and a luxury car, with high margins. By allowing the luxury division to choose prices after the budget division, the firm prevents the budget division from cannibalizing the high margins of the luxury division. Such a dynamic can be framed as the luxury division having a higher coordination weight and thus, as in Figure 2, moving second.

4.3 Alternative Governance Structures

Thus far in the analysis the only role of headquarters has been to decide the optimal timing of decision rights. These decision rights, however, were inalienable; allowing the headquarters to choose d_i and d_j (i.e. centralization) was not an option (i.e. decentralization was assumed). This approach was driven by the fact that in the motivating examples of Takeuchi and Nonaka (1986) and Hayes and Wheelwright (1979) there is no talk about the potential to centralize the decision-making let alone what centralization would mean from the R&D division in a biotechnology firm or on an assembly line. For instance, Aguilar and Bhambri (1983) describes Johnson and Johnson's difficulty in centralizing some of their division's functions.[22] Moreover, in the case of Dyer (1996), centralization may necessitate integration which may impose other costs on the firms. However, it is a natural question to ask whether the instances in which I find sequential decentralization to outperform concurrent decentralization are in fact the regions where the firm shouldn't be decentralized.

In this subsection, I will show, in fact, that the opposite tends to be true. Namely, as shown in Figure 2, when concurrent decision-making improves upon sequential decision-making is precisely the region where the firm should be centralized. This shows that sequential decision-making is the structure that should be used to make the centralization vs decentralization comparison.

To illustrate, the motivation behind centralization can be made most clear when considering r_1 close to 0 but r_2 close to 1 (i.e. unit 1 does not value coordination, but unit 2 does.). Under decentralization, unit 1 would not internalize the benefit unit 2 derives from coordination when choosing how well to coordinate with unit 2. This causes inefficiently low coordination between the units. In these instances, centralization corrects this externality. Extending the intuitions from Alonso et al. (2008) would then suggest that centralization dominates concurrent decentralization in these instances.[23] However centralization involves inefficiencies; decisions will often be improperly adapted to a unit's local condition due to strategic communication between the units and the headquarters.

To correct this lack of adaptation, Rantakari (2008) proposes yet another solution to this coordination issue, whose virtue is most easily seen when again considering r_1 close to 0 but r_2 close to 1. Instead of giving the headquarters both decision rights, give both decision rights to unit 1 (which Rantakari (2008) refers to as directional authority). This governance structure always results in $d_1 = d_2 = \theta_1$ in equilibrium. Since unit 2 is primarily interested in coordination, this gives sufficiently small losses to unit 2 and no losses to unit 1. Rantakari (2008) shows giving both decision rights to unit 1 improves upon centralization

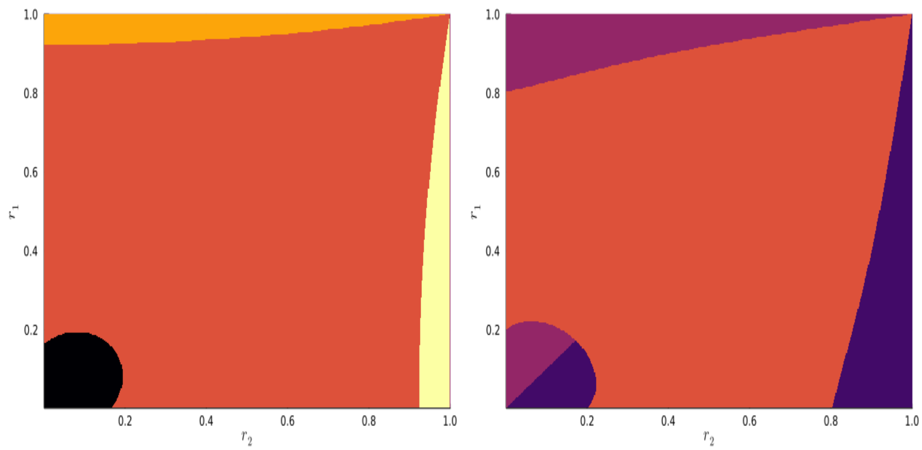
[22] Moreover, Aguilar and Bhambri (1983) also note that headquarters at Johnson and Johnson is composed of relatively few individuals. In such instances, it is hard to imagine easily tasking headquarters with more decisions. This is related to Aghion and Tirole (1997)'s canonical analysis in which it is assumed headquarters has a higher cost of making more decisions.

[23] This exact situation is not described in Alonso et al. (2008) since their analysis restricts attention to $r_1 = r_2$.

in these instances.

I show that by allowing sequential decision-making, in addition to concurrent decision-making, the region in which decentralization (with either timing) is optimal increases. In other words, sequential decision-making continues to be optimal even when allowing the alternative governance structures considered by Alonso et al. (2008) and Rantakari (2008). Returning to the thought experiment of r_1 close to 0 but r_2 close to 1: instead of moving decision rights, the firm could simply have unit 1 move first. Since unit 2 cares sufficiently about coordination, unit 2 is able to perform sufficiently well by knowing what decision unit 1 will make and then choosing their decision to be close to d_1. Having unit 1 move first accomplishes this goal and does so while decreasing the adaptation loss for unit 2. Figure 4 below shows the intuition above. In subfigure (a) the optimal governance choice not allowing for sequential decentralization is shown and subfigure (b) allows for sequential decentralization.[24]

[24]Subfigure (a) is the analysis done in Rantakari (2008).



(a) Not allowing sequential decision-making. (b) Allowing sequential decision-making.

Figure 4: The figure shows the governance structure that maximizes total surplus given equally weighted units with equal variances and allowing for additional governance structures. Subfigure (a) illustrates when decentralization is optimal not allowing for sequential decision-making while subfigure (b) allows for sequential decision-making. In both figures Light Orange corresponds to unit 2 makes both decisions, Yellow unit 1 makes both decisions, Dark Orange Centralization, Black concurrent decentralization, Purple unit 1 leads under sequential decentralization, and Pink Unit 2 leads under sequential decentralization.

A striking feature of Figure 4 is that concurrent decentralization is no longer optimal.[25] Comparing this with Figure 2 shows centralization removes the possibility that concurrent decentralization is optimal. Several features of centralization within this model, that may not always be met in practice, drive this stark result. First of all, as detailed in Marino et al. (2010), units may disobey orders from the headquarters. Additionally, having the headquarters make decisions d_1 and d_2 imposes no extra cost on the firm. In practice, headquarters may have higher decision-making costs as described in footnote 22. Further, cheap talk may be the incorrect informational assumption between a unit and their headquarters. Callander et al. (2008) argues a single cheap-talk message should not render the

[25] There may be alternative assumptions on the distribution of θ_i or various combinations of weights, volatilities, and coordination preferences that do yield concurrent decentralization to be optimal. Further, allowing for sequential messaging cannot improve sequential decentralization but may improve concurrent decentralization.

headquarters an expert. Finally, in some cases decentralization, and specifically concurrent decentralization, may allow for decisions to be made quicker, c.f. Van Zandt (1999).[26] To summarize, centralization may not always be feasible. Further, even in instances where it is feasible, this paper shows it may no longer be optimal.

5 Conclusion

The literature on firms' trade-off between coordination and adaptation has focused on who should make decisions rather than when the decisions should be made. I present a model of optimal timing of decisions within organizations which can determine when to have sequential or concurrent decision-making. In my setting, the fundamental trade off is that sequential decision-making allows for increased information transmission due to the revelation of a decision. Because unit j cares about only unit i's decision, not unit i's state, in my model sequential decision-making eliminates the need for communication by the first mover to the second, hence eliminating information loss when such communication is imperfect. However, sequential decision-making encourages the first mover to over-adapt to their state, causing the organization as a whole to over-adapt to the first-mover's local state. My model can compute when each of these effects dominates and shows that sequential is better when the preferences for coordination between two units are both sufficiently high and similar. Finally, comparing centralization to only concurrent decentralization may at times be misleading, because upon allowing for both centralization and sequential decentralization, concurrent decentralization is never optimal.

While the focus of this paper has been on the choice of the optimal decision timing given exogenous payoff parameters, it remains to be explored how the optimal choice of payoff parameters is shaped by optimal decision timing given the chosen parameters. A unit that knows they will move second has an incentive to increase their r_i, knowing that coordination will be relatively easier for them. This could come from a relationship specific investment that improves overall performance at the expense of being more sensitive to miscoordination as in Dyer (1996) and Rantakari (2013). These investments may additionally benefit the firm as when a unit values coordination more they have a reduced incentive to preempt the

[26] For instance, in the case of a biotechnology company producing a vaccine, the time savings may remove the optimality of centralization.

decision of their partner. Studying how the timing and payoff parameters move together is an interesting avenue for further research.

6 Appendix

6.1 Proofs

Unless otherwise stated, all proofs for the concurrent analysis can be found in Rantakari (2008).

Proof of Lemma 1. Due to the quadratic loss on both d_1 and θ_2, unit 2 chooses

$$d_2 = (1 - r_2)\theta_2 + r_2 d_1. \tag{4}$$

As a function of d_1 and θ_2 this gives unit 2 a loss of

$$U_2(\theta_2, d_1) = (1 - r_2) r_2 (d_1 - \theta_2)^2. \tag{5}$$

Knowing this is the behavior of unit 2, upon receiving message m_2 and generating interim beliefs $\mathbf{E}(\theta_2|m_i) := \bar{m}_i$, unit 1 minimizes

$$(1 - r_1)(d_1 - \theta_1)^2 + r_1(d_2 - d_1)^2.$$

Substituting in equation 4 yields

$$(1 - r_1)(d_1 - \theta_1)^2 + r_1(1 - r_2)^2(\theta_2 - d_1)^2 := 1 - r_1(d_1 - \theta_1)^2 + \beta_1(\theta_2 - d_1)^2.$$

Which is equivalent to minimizing

$$1 - r_1(d_1 - \theta_1)^2 + \beta_1(\bar{m}_i - d_1)^2.$$

Taking a first-order condition yields

$$d_1 = \frac{1 - r_1\theta_1 + \beta_1\bar{m}_i}{1 - r_1 + \beta_1} := (1 - \gamma_1)\theta_1 + \gamma_1\bar{m}_i.$$

■

Proof of Lemma 2. Given equation (5), this generates the following indifference condition:

$$\begin{aligned} & U_2(k_i,k_{i+1}) = U_2(k_i,k_i) \\ & \iff (d_1(k_{i+1}) - k_i)^2 = (d_1(k_i) - k_i)^2 \\ \text{now} & \iff k_i - \gamma_1 \bar{m}_i = \gamma_1 \bar{m_{i+1}} - k_i \\ \text{using uniform dist. assumption} & \iff k_i - \gamma_1 \frac{k_i + k_{i-1}}{2} = \gamma_1 \frac{k_i + k_{i+1}}{2} - k_i \\ & \iff k_{i+1} - k_i = k_i - k_{i-1} + \frac{4}{\frac{\gamma_1}{1-\gamma_1}} k_i \\ & \iff k_{i+1} - k_i = k_i - k_{i-1} + \frac{4}{\frac{\beta_1}{1-r_1}} k_i. \end{aligned}$$

Alonso et al. (2008) shows that when the difference equation is of the form $k_{i+1} - k_i = k_i - k_{i-1} + \frac{4}{\phi} k_i$ the communication loss, c_i, which is defined to be $\mathbf{V}(\theta_i - E(\theta_i|m_i)) = \frac{1}{3}\frac{1}{3\phi+4}$. Note that, in my case $\phi = \frac{\beta_1}{1-r_1}$

■

Proof of Lemma 3. I can calculate the loss of the leader given an interval i as follows

$$\begin{aligned} & \mathbb{E}\Big((1-r_1)(d_1-\theta_1)^2 + r_1((1-r_2)\theta_2 + r_2 d_1 - d_1)^2|i\Big) \\ & = (1-r_1)\mathbb{E}((d_1-\theta_1)^2|i) + r_1(1-r_2)^2\big(\mathbb{E}((\theta_2-d_1)^2|i)\big) \end{aligned}$$

recall $d_1 = (1-\gamma_1)\theta_1 + \gamma_1\bar{m}_i$

$$\begin{aligned} & = (1-r_1)\mathbb{E}((-\gamma_1\theta_1 + \gamma_1\bar{m}_i)^2|i) + r_1(1-r_2)^2\Big(\mathbb{E}((\theta_2 - (1-\gamma_1)\theta_1 - \gamma_1\bar{m}_i)^2|i)\Big) \\ & = (1-r_1)\gamma_1^2\mathbb{E}((\theta_1-\bar{m}_i)^2|i) + r_1(1-r_2)^2\Big((1-\gamma_1)^2V(\theta_1) + \mathbb{E}(\theta_2^2|i) + \gamma_1^2(\bar{m}_i)^2 - 2\gamma_1(\bar{m}_i)^2\Big) \qquad (6) \\ & = (1-r_1)\gamma_1^2(\mathbb{E}(\theta_1^2|i) + \bar{m_i}^2) + \beta_1\Big((1-\gamma_1)^2V(\theta_1) + \mathbb{E}(\theta_2^2|i) + (\gamma_1^2 - 2\gamma)(\bar{m}_i)^2\Big). \end{aligned}$$

Taking the expectation across intervals yields

$$\begin{aligned} & \sum_{i=1}^{N} P(i)\Bigg((1-r_1)\gamma_1^2(\mathbb{E}(\theta_1^2|i) + \bar{m_i}^2) + \beta_1\Big((1-\gamma_1)^2V(\theta_1) + \mathbb{E}(\theta_2^2|i) + (\gamma_1^2 - 2\gamma)(\bar{m}_i)^2\Big)\Bigg) \\ & = (1-r_1)\gamma_1^2(\mathbf{v}(\theta_1) + \mathbb{E}(\bar{m_i^2})) + \beta_1\Big((1-\gamma_1)^2V(\theta_1) + \mathbf{v}(\theta_2) + (\gamma_1^2 - 2\gamma)\mathbb{E}(\bar{m_i}^2)\Big). \end{aligned}$$

I can re-arrange this in terms of loss from miscoordination and loss from imperfect communication as follows where $c_i := V(\theta_i) - \mathbb{E}(\bar{m}_i^2)$ is the loss from communication:

$$
\begin{aligned}
&= (1-r_1)\gamma_1^2(\mathbf{v}(\theta_1) + V(\theta_2) - c_2) + \beta_1\Big((1-\gamma_1)^2 V(\theta_1) + \mathbf{v}(\theta_2) + (\gamma_1^2 - 2\gamma)(\mathbf{v}(\theta_2) - c_2)\Big)\\
&= \Big(\mathbf{v}(\theta_2) + \mathbf{v}(\theta_1)\Big)\Big((1-r_1)\gamma_1^2 + \beta_1(1-\gamma_1)^2\Big) - c_2\Big((1-r_1)\gamma_1^2 + \beta_1(\gamma_1^2 - 2\gamma)\Big)\\
&= (1-r_1)\gamma_1\Big(\mathbf{v}(\theta_2) + \mathbf{v}(\theta_1)\Big) + \gamma_1\beta_1 c_2.
\end{aligned}
$$

As mentioned above the loss of the follower is

$$(1-r_2)r_2(d_1(k(\theta_2)) - \theta_2)^2.$$

Given an interval i

$$
\begin{aligned}
\mathbb{E}((d_1(k(\theta)) - \theta)^2 | i) &= \mathbb{E}\big(\big((1-\gamma_1)\theta_1 + \gamma_1\mu_i - \theta_2\big)^2 \big| \, i\big)\\
&= (1-\gamma_1)^2 V(\theta_1) + \mathbb{E}\big(\big(\theta_2 - \gamma_1\mu_i\big)^2 \big| \, i\big)\\
&= (1-\gamma_1)^2 V(\theta_1) + \mathbb{E}(\theta_2^2 | i) + \gamma_1^2(\bar{m}_i)^2 - 2\gamma_1(\bar{m}_i)^2.
\end{aligned}
$$

Now taking the expectation across intervals

$$
\begin{aligned}
&(1-r_2)r_2\sum_{i=1}^{N} P(i)\Bigg(((1-\gamma_1)^2 V(\theta_1) + \mathbb{E}(\theta_2^2|i) + \gamma_1^2(\bar{m}_i)^2 - 2\gamma_1(\bar{m}_i)^2\Bigg)\\
&= (1-r_2)r_2\Big((1-\gamma_1)^2 V(\theta_1) + V(\theta_2) + (\gamma_1^2 - 2\gamma_1)\mathbb{E}(\bar{m}_i^2)\Big).
\end{aligned}
$$

Note that the first term is independent of i and thus can be pulled out of the sum. To get a similar formula as the leader case, recall that $c_i = V(\theta_i) - \mathbb{E}(\bar{m}_i^2)$, so

$$= (1-r_2)r_2(1-\gamma_1)^2\Big(\mathbf{v}(\theta_2) + \mathbf{v}(\theta_1)\Big) - (1-r_2)r_2(\gamma_1^2 - 2\gamma_1)c_2.$$

■

Proof of Lemma 4. The losses for concurrent decision-making are taken from Rantakari (2008) and adapted to the notation of this paper. ■

Proof of Proposition 1. Given a fixed communication quality, I show a stronger result, namely that the coordination loss is lower for any realization of the informational partition.

Statement ii of Proposition 1:

The coordination loss when unit 1 leads is $(d_2 - d_1)^2 = (1-r_2)^2(d_1 - \theta_2)^2$. Noting now that $d_1 = (1-\gamma_1)\theta_1 + \gamma_1\theta_2$ yields the following for coordination loss:

$$(1-r_2)^2(1-\gamma_1)^2\mathbf{E}(\theta_2^2 + \theta_1^2 | m_1, m_2).$$

Hence coordination loss is less when unit 1 leads if and only if

$$\begin{aligned}
&(1-r_2)^2(1-\gamma_1)^2 < (1-r_1)^2(1-\gamma_2)^2\\
\iff\ &(1-r_2)\frac{(1-r_1)}{(1-r_1)+r_1(1-r_2)^2} < (1-r_1)\frac{(1-r_2)}{(1-r_2)+r_2(1-r_1)^2}\\
\iff\ &(1-r_2)+r_2(1-r_1)^2 < (1-r_1)+r_1(1-r_2)^2\\
\iff\ &r_2(-2r_1+r_1^2) < r_1(-2r_2+r_2^2)\\
\iff\ &-2+r_1 < -2+r_2\\
\iff\ &r_1 < r_2.
\end{aligned}$$

Statement i of Proposition 1:

Meanwhile coordination loss when the units move at the same time is bounded above by $(\frac{(1-r_1)(1-r_2)}{1-r_1r_2})^2\mathbf{E}(\theta_2^2+\theta_1^2|m_1,m_2)$. Comparing this to the sequential structure yields[27]

$$\begin{aligned}
&(\frac{(1-r_1)(1-r_2)}{1-r_1r_2})^2 < (1-r_2)^2(1-\gamma_1)^2\\
\iff\ &\frac{1}{1-r_1r_2} < \frac{1}{1-r_1+r_1(1-r_2)^2}\\
\iff\ &-r_1+r_1(1-r_2)^2 < -r_1r_2\\
\iff\ &r_2 < 1.
\end{aligned}$$

■

Proof of Proposition 2.

Statement i of Proposition 2: As mentioned before, there is perfect communication on the state of the the leader, so the total information loss with sequential decision-making is $V(\theta_i)\frac{1}{4+3\frac{\beta_i}{1-r_i}}$. Note that this is always less than $\frac{V(\theta_i)}{4}$. It is without loss to assume $V(\theta_2) < V(\theta_1)$ and I show below that sequential decision-making with unit 1 leading always incurs less information loss than concurrent decision-making. Note that in concurrent decision-making, information loss occurs on both local states and

$$\begin{aligned}
&\mathcal{L}_{\text{concurrent}} = \frac{V(\theta_1)}{4+3\phi_{\text{concurrent}_1}} + \frac{V(\theta_2)}{4+3\phi_{\text{concurrent}_2}} \overset{?}{>} \frac{V(\theta_2)}{4}\\
\iff\ &\mathcal{L}_{\text{concurrent}} = \frac{1}{4+3\phi_{\text{concurrent}_1}} + \frac{1}{4+3\phi_{\text{concurrent}_2}} \overset{?}{>} \frac{1}{4}\\
\iff\ &\frac{\phi_{\text{concurrent}_1}\phi_{\text{concurrent}_2}-\frac{16}{9}}{(\phi_{\text{concurrent}_1}+\frac{4}{3})(\phi_{\text{concurrent}_2}+\frac{4}{3})} \overset{?}{<} 0\\
\iff\ &\phi_{\text{concurrent}_1}\phi_{\text{concurrent}_2} \overset{?}{<} \frac{16}{9}\\
\iff\ &r_1r_2 < \frac{16}{9}.
\end{aligned}$$

[27] Since I show it is better for all r_1, r_2, it is without loss to compare only to the structure in which unit 1 leads.

Thus the sequential information loss is always bounded below by $\frac{V(\theta_2)}{4}$; meanwhile, the concurrent information loss is bounded above by $\frac{V(\theta_2)}{4}$. The latter statement that there is worse information transmission about the followers state follows by noting that $\phi_{\text{concurrent}} > \phi_{\text{sequential}}$ and that information loss is monotonic in ϕ.

Statement ii of Proposition 2:

Recall that when decisions are made sequentially the only information loss is from the follower. This loss is a monotone decreasing function of ϕ_i when unit i leads. The algebra is as follows:

$$\begin{aligned}
\frac{\beta_1}{(1-r_1)} \geq \frac{\beta_2}{(1-r_2)} &\iff \\
\beta_1(1-r_2) \geq \beta_2(1-r_1) &\iff \\
r_1(1-r_2)^3 \geq r_2(1-r_1)^3 &\iff \\
\frac{r_1}{(1-r_1)^3} \geq \frac{r_2}{(1-r_2)^3} &\iff \\
r_1 \geq r_2.
\end{aligned}$$

Statement iii of Proposition 2:

If $r_1 = r_2$, then $\phi_1 = \phi_2$ and thus the percentage of information lost due to strategic communication is the same under both timings. However, having the unit with less variance communicating via cheap talk as opposed to the unit with more variance will yield less aggregate information loss. ■

6.2 Additional Cheap Talk Communication Analysis

Below I plot the graph of the optimal sequential structure, namely which unit should move first conditional on being sequential. As seen in Figure 1, if the optimal structure is sequential, then the leader should be the unit with a lower r_i except for a small region with intermediate and similar r_i. In this region the optimal structure has the unit that cares more about coordination moving first. The non-monotonicity exists because in this region the gain from higher-quality information outweighs the loss of worse coordination conditional on information. Recall that Proposition 1 says that to increase coordination, conditional on information, the unit which places a higher priority on coordination should move second. In contrast, Proposition 2 says that having this unit move first increases information transmission. However, the difference in coordination between the two timings compared in Proposition 1 is lowest when the two care symmetrically about coordination due to continuity. Additionally the information gain from switching governance structures is largest in intermediate values of r_i and r_j, because for sufficiently high(low) values on coordination, information transmission will be high(low) regardless of who moves first. Hence, there is one region with intermediate and similar r_i and r_j where the unit that cares more about coordination should move first because this has a large gain to information transmission but a small loss to coordination. This region is smaller in Figure 1 than it is below since this region is where concurrent decision-making is optimal.

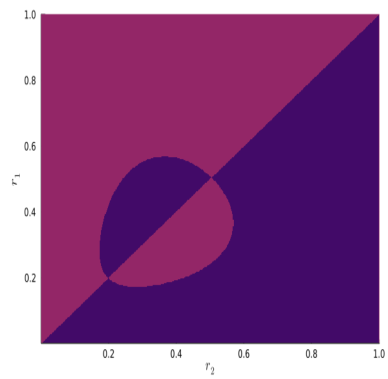


Figure 5: The figure shows which sequential timing minimizes the sum of the units' losses as a function of the coordination motives of the units, r_1 and r_2. The color denotes the optimal timing with Purple corresponding to unit 1 leading and Pink unit 2 leading.

6.3 Public Information Analysis

When the states are known by both parties, this is very broadly a public goods provision problem. The more a unit sacrifices compromising on their own state by moving towards the other unit's state, the more they are contributing to the public good of compromise. The intricacy comes from the fact that the more the leader invests in compromise, the less the follower invests in compromise due to the quadratic losses.

Analyzing the game when there is public information and unit i observes both θ_i and θ_{-i} is equivalent to performing the analysis with cheap talk but ignoring the truth telling constraints and having $m_i = \theta_i$. Since there is no informational component, unit i would always prefer to be the leader. This is because they know they can under invest in coordination and unit j will over invest. Unsurprisingly, the firm with equally weighted divisions never wants a sequential structure since the only benefit of sequential is increased information transmission, and since there is public information, communication does not matter.[28] However, as the weight the headquarters places on unit 1 increases, sequential structures will eventually become optimal because leading is best for unit 1.

6.4 Incomplete Information Analysis

To solve for the utilities of each unit, we can simply take Lemmas 3 and 4 and plug in 0 for $V(\phi)$, since the analysis with no information is still one of cheap talk, but is the least informative equilibrium.

I can now write the losses for unit 1 across the various timings:

- $\frac{(1-r_1)\beta_1}{(1-r_1)+\beta_1}V(\theta_1) + \beta_1 V(\theta_2)$ when unit 1 leads
- $r_1(1-r_1)\left[V(\theta_1) + \frac{\alpha_2^2}{(\alpha_2+\beta_2)^2}V(\theta_2)\right]$ when unit 1 follows
- $r_1(1-r_1)V(\theta_1) + \beta_1 V(\theta_2)$ is the loss when unit 1 decides concurrently.

One can see the loss from miscoordination, i.e., the $V(\theta_2)$ term, is the same for unit 1 when deciding first or at the same time. However, when unit 1 is a leader, they always are able to lower their adaptation loss since

$$\begin{aligned} \frac{(1-r_1)\beta_1}{(1-r_1)+\beta_1} &\overset{?}{<} (1-r_1)r_1 \\ \iff \frac{(1-r_2)^2}{(1-r_1)+r_1(1-r_2)^2} &\overset{?}{<} 1, \end{aligned}$$

which is true since the bottom is a convex combination of the numerator and a term greater than the numerator, 1. Hence, unit 1 would either want to be the leader or the follower as seen below. When they care sufficiently about coordination relative to unit 2, they would prefer to move second to ensure they can coordinate with unit 1.

[28] To see this, one can confirm that the total losses from Lemma 4 when setting $c_i = 0$ are always larger under sequential decision-making.

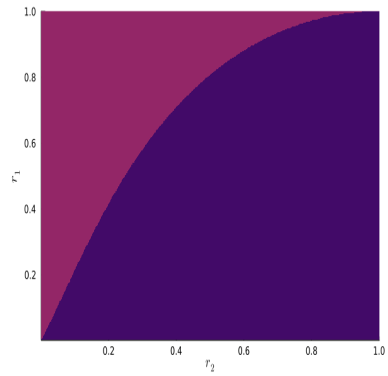


Figure 6: This figure shows the optimal governance structure for unit 1 when the units are not allowed to communicate as a function of the coordination motives of the units, r_1 and r_2. Purple corresponds to unit 1 leads, Pink is unit 2 leads, and Black is Concurrent with equal volatilities.

It is worth noting that any time a unit would want to follow under cheap talk, they would also prefer to follow under incomplete information. This is because followers value the additional information from going second more than the ability to adapt to their state more as the leader. When moving from cheap talk to incomplete information, the first force becomes even stronger as now first movers are unable to know their follower's state and are thus unable to coordinate.

The optimal governance structure for the firm is plotted in Figure 7. The firm places the unit that cares more about coordination as a second mover. Additionally, concurrent decision-making is never optimal, since sequential always allows some information transmission but concurrent gives none.

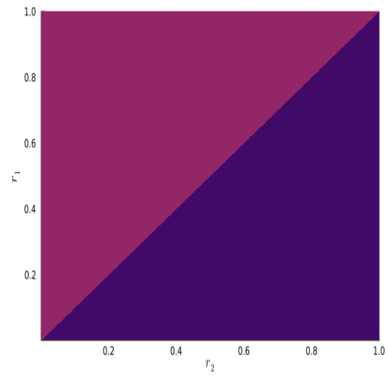


Figure 7: This figure shows the optimal governance structure with equally weighted firms when the units are not allowed to communicate as a function of the coordination motives of the units, r_1 and r_2. Purple corresponds to unit 1 leads, Pink is unit 2 leads, and Black is Concurrent with equal volatilities.

References

Admati, A. R. and Perry, M. (1991). Joint projects without commitment. *The Review of Economic Studies*, 58(2):259–276.

Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations. *Journal of political economy*, 105(1):1–29.

Aguilar, F. J. and Bhambri, A. (1983). Johnson & johnson (a). *Harvard Bussines School Case*, (5).

Akerlof, R. and Holden, R. (2019). Capital assembly. *The Journal of Law, Economics, and Organization*, 35(3):489–512.

Alonso, R., Dessein, W., and Matouschek, N. (2008). When does coordination require centralization? *American Economic Review*, 98(1):145–79.

Alonso, R., Matouschek, N., and Dessein, W. (2010). Strategic communication: prices versus quantities. *Journal of the European Economic Association*, 8(2-3):365–376.

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

Callander, S. et al. (2008). A theory of policy expertise. *Quarterly Journal of Political Science*, 3(2):123–140.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451.

Dessein, W. and Santos, T. (2006). Adaptive organizations. *Journal of Political Economy*, 114(5):956–995.

Dyer, J. H. (1996). How chrysler created an american keiretsu. *Harvard Business Review*, 74(4):42–52.

Ellison, G. (2005). A model of add-on pricing. *The Quarterly Journal of Economics*, 120(2):585–637.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy*, 94(4):691–719.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *Journal of political economy*, 98(6):1119–1158.

Hayes, R. and Wheelwright, S. (1979). Link manufacturing process and product life cycles. *Harvard Business Review*, pages 133–140.

Hortaçsu, A., Natan, O. R., Parsley, H., Schwieg, T., and Williams, K. R. (2021). Organizational structure and pricing: Evidence from a large us airline. Technical report, National Bureau of Economic Research.

Lewis, K. and Mistree, F. (1997). Collaborative, sequential, and isolated decisions in design. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 80456, page V003T30A018. American Society of Mechanical Engineers.

Marino, A. M., Matsusaka, J. G., and Zábojník, J. (2010). Disobedience and authority. *The Journal of Law, Economics, & Organization*, 26(3):427–459.

Moorthy, K. S. and Png, I. P. (1992). Market segmentation, cannibalization, and the timing of product introductions. *Management science*, 38(3):345–359.

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

Rantakari, H. (2013). Organizational design and environmental volatility. *The Journal of Law, Economics, & Organization*, 29(3):569–607.

Safizadeh, M. H., Ritzman, L. P., Sharma, D., and Wood, C. (1996). An empirical analysis of the product-process matrix. *Management Science*, 42(11):1576–1591.

Takeuchi, H. and Nonaka, I. (1986). The new new product development game. *Harvard business review*, 64(1):137–146.

Valle, S. and Vázquez-Bustelo, D. (2009). Concurrent engineering performance: Incremental versus radical innovation. *International Journal of Production Economics*, 119(1):136–148.

Van Zandt, T. (1999). Real-time decentralized information processing as a model of organizations with boundedly rational agents. *The Review of Economic Studies*, 66(3):633–658.

Varian, H. R. (1994). Sequential contributions to public goods. *Journal of Public Economics*, 53(2):165–186.