Actions Speak Louder Than Words: Should We Let Them?

Roi Orzach *

August 31, 2021

Abstract This paper analyzes concurrent vs. sequential organizational 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 communication by allowing first movers to speak with their actions rather than cheap-talk messages. However, first movers also have an incentive to over-adapt to their state, knowing second movers will conform to their decision. Concurrent decisions are optimal if and only if the two units’ have similar volatilities and weights on coordination. Hence, organizations should analyze who decides first, not just who decides. (JEL D21, D23, D82)

1 Introduction

Many decisions between two parties are made sequentially. For example, Apple’s hardware unit designed a computer before their software unit created an operating system. This timing allowed the software unit to know exactly what hardware to write their code for. However, of course, many decisions are made simultaneously. Automobile companies realized that when design occurred before production, designers abused their first-mover

*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.
privileges by not listening to manufacturers and making designs that were challenging to produce.

In this paper I address two questions: when should decisions be made sequentially, and in sequential cases who should go first? I analyze these questions 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 local conditions (adaptation) and to the other unit’s decision (coordination). Before making their decisions, the units engage in strategic communication. Units may asymmetrically weight coordination and adaptation, similar to how a design unit may place a lower weight on coordination than a manufacturing unit. Unsurprisingly, asymmetric firms are most likely to benefit from asymmetric (sequential) timings.

Simultaneous development was studied in Alonso et al. (2008) and Rantakari (2008) who use this model to determine the allocation of decision rights noting that moving decision rights can increase communication quality and coordination. However, in keeping with the adage that actions speak louder than words, staggering the decisions allows the unit that moves second to perfectly observe the decision of the first mover. This is in contrast to simultaneous development where each unit is unsure of the exact decision the other unit will make. Under simultaneous development, units have an incentive to exaggerate their own state knowing their partner will not fully adapt to it. Meanwhile under sequential development, since the decision is always fully revealed there is no need to send a biased message, but they do have an incentive to take a biased action (e.g., excessively adapt to their local state).

Analyzing when the additional informational gain due to revealing the leading unit’s decision outweighs the loss from biased actions creates an additional governance choice for firms. Throughout the paper I will relate my model to the following articles addressing the same sequentiality question as mine.

There exists a large literature analyzing simultaneous decision-making, i.e., “concurrent engineering.” Concurrent engineering is characterized by having functional units typically later in the product development cycle—such as sales, marketing, or supply chain—making decisions simultaneously with design/engineering. Takeuchi and Nonaka (1986) look at why certain industries have concurrent engineering practices and claim that the difference is

---

1These are models with non-contractable decisions and the only organizational power the headquarters has is changing the allocation of decision rights.
determined by the industry the firm is in. The main prediction of their paper is that industries with a research and development focus are less likely to be concurrent. They justify this prediction by reasoning that these industrial research units are sufficiently focused on adaptation to their local conditions, rather than coordination with sales and other units, that the research team needs to move first without considering the implications for other teams. Valle and Vázquez-Bustelo (2009) test this claim using survey evidence. In support of Takeuchi and Nonaka (1986), they show that firms engaged in breakthrough innovation are less likely to adopt concurrent engineering practices.

While Takeuchi and Nonaka (1986) analyze production decisions, a similar question exists later in the production of products. The Product-Process Matrix in Hayes and Wheelwright (1979) describes instances when the various units (e.g. chemical treatment, assembly, paint, etc) within a firm will be called upon sequentially or simultaneously in their production process. They predict that when firms have many different product variants (e.g., because of many consumer niches or fast evolving tastes) and low levels of standardization, all units across the firm will make decisions at the same time. They also predict that a standardized product is likely to be produced in an assembly line, which is a canonical sequential process. This hypothesis was later tested in Safizadeh et al. (1996), which used survey data from firms to show that the simultaneous vs sequential prediction from Hayes and Wheelwright (1979) is seen in the data.

In addition to shedding light on the dynamics of communication and decision-making within a firm, the forces I study also apply between firms. A guiding example for this theory applied to supplier relations is Chrysler’s change in organizational structure, as described in Dyer (1996). Dyer (1996) details how in the early 1990’s Chrysler drastically reduced production costs by changing their timing with suppliers from sequential to concurrent. Previously, Chrysler would design their cars and later hold auctions with many suppliers to determine who would supply a set of parts. Since Chrysler had already finished designing their cars, the exact specifications of the parts were given to the suppliers, yielding perfect communication. But Chrysler engineers would determine these specifications without considering the difficulties of producing these parts, yielding sub-optimal coordination between the firms.

Instead of interacting with each supplier after being selected by an auction via only formal contracts, Chrysler developed an exclusive partnership with their suppliers and started
working more closely with them. Due to this exclusivity, Chrysler began making relationship specific investments. For instance, Chrysler may have invested in a machine that can very efficiently utilize parts from a specific supplier because they know they will continue to use the machine for the next several years. One implication of these investments is Chrysler placing more weight on coordination with their suppliers. Given the change in coordination incentives, Chrysler also moved from sequential development (i.e., Chrysler engineers draw up specs of a part, supplier then figures out how to build it) to simultaneous development (i.e., Chrysler engineers work with supplier engineers as they together decide what part to build). The theory detailed in this paper precisely describes this change in organizational structure.

In this paper I will show that this phenomenon holds true more generally, namely, that if the units place high and similar weights on coordination and have similar local volatilities, then they decide at the same time; in other cases, the unit that cares more about coordination almost always decides second. In the language Takeuchi and Nonaka (1986), if the research unit is conducting breakthrough innovation, then an asymmetry exists between them and sales on how much they value coordination and on how much uncertainty they are encountering, which causes sequential decision-making to be optimal. In the language of Hayes and Wheelwright (1979), when there is a standardized product, then the underlying volatilities in units like paint or detailing are small, yielding sequential decision to be optimal. Finally, in Dyer (1996), when Chrysler and their suppliers started emphasizing coordination similarly, they moved to simultaneous decision-making.

The remainder of this paper is organized as follows. Section 2 reviews the related literature, Section 3 describes the model, Section 4 gives properties of sequential decision-making under three different informational structures, Section 5 provides the results, and Section 6 concludes. The Appendix contains proofs for all statements not proved in the text.

2 Literature Review

This paper connects to the literature on adaptation coordination 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 allocation of decision rights. Given

---

2For a theory of endogenous changes in payoff parameters I refer the reader to Rantakari (2013).
this, I adopt the notation of Rantakari (2008) as much as possible. Rantakari (2008) explores a potential asymmetry between the divisions by allowing one unit to retain their decision right while the other unit has their decision taken by the headquarters. Meanwhile, Alonso et al. (2008) does consider sequential decision-making. However, they ask whether, if units make decisions sequentially under decentralization, does the centralization/decentralization trade off remains at play without comparing sequential decentralization to simultaneous decentralization.\(^3\) In contrast to their paper analyzing if centralization/decentralization is optimal given the different timings, I focus on when should decisions in fact be sequential for a fixed set of decision makers.\(^4\) For this reason, I take it as given that decision rights are inalienable.\(^5\) Additionally, Rantakari (2013) shows 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.

An extensive literature exists on sequential vs simultaneous contributions to public goods, focusing on which yields the largest donations.\(^6\) This research is fundamentally related to the coordination vs 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 to get enough donations for a project’s completion under a sequential move game, players must use a very sophisticated trigger strategy, suggesting sequential donations are harder to achieve. In fact Varian (1994) shows that in a continuous public goods setting, weakly smaller contributions will occur when they happen sequentially. This intuition cannot be blindly extended into the settings of an organization since these papers lack private information. With private information, communication quality increases substantially when moving to a sequential structure because the follower can perfectly observe the decision of the leader, which their models do not capture.

The paper which answers a question most similar to mine is Lewis and Mistree (1997),

\(^3\)Since Alonso et al. (2008) considers symmetric units they also don’t ask which unit should go first when decisions are sequential.

\(^4\)Recent papers, such as Angelucci et al. (2021), have looked into sequential messaging within this framework although not sequential decision-making.

\(^5\)Note that sequential decision-making is optimal for some parameter values even when I consider the larger class of governance structures. This suggests that this paper’s insights are at play in the more general setting in which decision rights can be centralized. [Computations available upon request].

\(^6\)In this literature, the equilibrium which yields the most donations, is also the one that maximizes social surplus.
which looks at a game in which two different units of an airline each get to choose a dimension of their part and they examine times when the aircraft is better designed in a simultaneous or sequential format. Lewis and Mistree (1997) omit the communication stage and the random local condition, and thus also reach the incorrect conclusion that simultaneous development is always better but gets closest to the idea of decision-making orders within an organization.

3 Model

There are two units who wish to match their decision, $d_i$, to both their local conditions (adaptation) and the other unit’s decision (coordination). Each unit is privately informed about their local condition, $\theta_i$ and wish to minimize a weighted sum of their coordination and adaptation losses

$$L_i(d_i, d_{-i}, \theta_i) = (1 - r_i)(\theta_i - d_i)^2 + r_i(d_{-i} - d_i)^2.$$ 

Here, $r_i \in (0, 1)$ measures the weight on coordination as units that value coordination more have a larger coefficient $(d_i - d_j)^2$. I will consider two different governance structures: sequential and concurrent. In both structures, each unit observes their $\theta_i$, and then units engage in cheap talk communication by sending messages, $m_i$. The decision timings of the two governance structures diverge:

- **Sequential**: After observing the messages, unit $j$ chooses $d_j$ and then upon observing $d_j$, unit $i$ makes their decision $d_i$.

- **Concurrent**: After observing the messages, units 1 and 2 choose $d_1$ and $d_2$ at the same time.

I will be solving for a Perfect Bayesian Equilibrium. I focus on the most informative equilibrium of the cheap talk communication game, as is standard in the literature. Following through on the example from Takeuchi and Nonaka (1986), it may be reasonable to think $r_{\text{engineering}} \sim 0, r_{\text{sales}} \sim 1$ in the case of an aerospace corporation. In my setting, this is the Pareto dominant equilibrium. The headquarters chooses the governance structure to maximize the sum of profits across the two units (and this is the only decision

---

7 Following through on the example from Takeuchi and Nonaka (1986), it may be reasonable to think $r_{\text{engineering}} \sim 0, r_{\text{sales}} \sim 1$ in the case of an aerospace corporation.

8 The least informative equilibrium is equivalent to a game with private information, but without cheap talk, and is discussed in Section 4.1.
they make). The timing is summarized below.

<table>
<thead>
<tr>
<th>$T = 1$</th>
<th>$T = 2$</th>
<th>$T = 3$</th>
<th>$T = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governance Structure Chosen</td>
<td>Units observe</td>
<td>Units send their cheap talk message, $m_i$</td>
<td>Decisions $d_i, d_j$ are made</td>
</tr>
<tr>
<td>$\theta_i$</td>
<td>$\theta_i$</td>
<td>$\theta_i$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Timing of the Game

4 Analysis

4.1 Benchmarks

It is worth analyzing the model under two benchmark information structures: private information without cheap talk and full information, respectively. The calculations for both cases can be found in the Appendix. Additionally, the analysis for these two cases does not rely on any distributional assumptions on the $\theta_i$’s. The fundamental benefit of sequential actions, as opposed to simultaneous, is that perfect communication of the decision of the first mover always occurs. The absence of cheap talk exacerbates this effect since units can communicate only by revelation through actions. Given this exacerbation, I prove the following remark.

**Remark 1** Without cheap talk the optimal governance structure is to have the unit that cares more about coordination move second.

In contrast, the complete information benchmark I prove is that both units should move concurrently. To gain intuition for this, one can think of the two units engaging in a public goods game. Each unit can choose to move their action closer to that of the other party at some private cost (not adapting to their local condition). As is standard in public goods papers, sequential “contributions” lead to worse outcomes for the players. Worse outcomes occur because contributions are strategic substitutes across the agents, and thus the first mover has an incentive to invest even less than would be optimal, knowing the second mover

---

9 In Section 5 I discuss how the optimal governance structure would change if the headquarters had heterogeneous weights on the two units.
will over-invest.\textsuperscript{10} Hence, in the complete information benchmark, units would always prefer to act simultaneously as stated below.

\textbf{Remark 2} \textit{With complete information, the optimal governance structure is to have the units decide concurrently.}

As I will show in the next section, private information with cheap talk departs from these extreme cases and allows situations in which sequential timings yields higher surplus and vice versa.

\subsection{Cheap Talk Communication Analysis}

It is without loss of generality to assume that, under sequential decisions, unit 1 is the leader with unit 2 being the follower, since one can relabel the units. Additionally, to be able to get an explicit form of the communication equilibria, I assume $\theta_1, \theta_2$ are independent such that $\theta_i \sim U [-\bar{\theta}_i, \bar{\theta}_i]$.\textsuperscript{11} I will define the volatility of unit $i$ will be defined as the variance of their local conditions.

I can now solve the game by backwards induction and calculate the decisions by the units after receiving any communication and updating their beliefs. Under sequential timing, given the message and decision from the leader, the follower will ignore the message. To see why this is the case, recall the message is only relevant because it is predictive of $\theta_1$ and hence $d_1$, but the follower has already observed $d_1$. Hence, under sequential decision-making, it is without loss to assume that $m_1 = \theta_1$.\textsuperscript{12} For notational simplicity, define $r_1(1 - r_2)^2 := \beta_1$. This is a measure of the wedge between the adaptation preferences of the two units.

\textbf{Lemma 1} \textit{Equilibrium decisions under sequential decision-making are}

\begin{align*}
d_1 &= \frac{1 - r_1}{(1 - r_1) + \beta_1} \theta_1 + \frac{\beta_1}{(1 - r_1) + \beta_1} \mathbb{E}(\theta_2|m_2), \\
d_2 &= (1 - r_2)\theta_2 + r_2 d_1,
\end{align*}

\textsuperscript{10}This closely follows the intuition from Varian (1994), but occurs in more general settings whenever actions are strategic substitutes.

\textsuperscript{11}In the alternative information structures discussed in Section 4.1, this assumption is not needed and only the moments of the distribution affect payoffs.

\textsuperscript{12}It is equivalent to assume $m_1 = \emptyset$, but the assumption $m_1 = \theta_1$ allows a definition of communication loss consistent with the literature.
while under simultaneous decisions they are
\[ d_i = (1 - r_i) \theta_i + r_i E(d_j | m_i). \]

One can note that unit 2’s decision under sequential settings does not involve expectations since they have already observed the decision of unit one, and hence do not care about either \( \theta_1 \) or \( m_1 \). However, in simultaneous timings or for unit 1 under sequential timings, decisions are a function of posterior beliefs generated after receiving the message from their partner.

Looking at the coefficient on \( E(\theta_2 | m_1) \) in decision 1 under sequential timing, which will be denoted by \( \gamma_1 := \frac{\beta_1}{(1-r_1) + \beta_1} \), one can see that the denominator is less than 1. This means the leader is adapting to their state more than just \( (1 - r_1) \) which is how much they would adapt under simultaneous decision making. This inequality implies that the leader is adapting to their state more than if they were a follower. Additionally one can note that \( \gamma_1 \) is increasing in \( r_2 \). This relation implies that the more unit 2 is focused on coordination, the more unit 1 will focus on adapting to her state.

However, there will not always be perfect communication from the follower under sequential timing or from either unit under simultaneous timing. If there were, the unit would then prefer to exaggerate their type to get perfect coordination at no cost to them. Hence, as is standard in the literature, I will 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.

Given the decision 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 \([-1, 1]\) such that \( U_2(k_i, k_{i+1}) = U_2(k_i, k_i) \). Communication loss is defined as \( c_i = V(\theta_i - E(\theta_i | m_i)) \), where \( V \) denotes the variance. As the informational partitions become finer, \( E(\theta_i | m_i) \) approaches \( \theta_i \); hence, better communication can be viewed as a finer partition.

**Lemma 2** Equilibrium communication loss is \( V(\theta_i) \frac{1}{3 \phi + 4} \) where \( \phi_{\text{sequential}} = \frac{\beta_1}{(1-r_1)} = \frac{r_1(1-r_2)^2}{(1-r_1)} \)

and \( \phi_{\text{simultaneous}} = \frac{r_j(1-r_j)}{(1-r_j)} \).

\[ \frac{1}{3 \phi + 4} \] and \( \frac{1}{(1-r_1)} \) for formal details, I refer the reader to Alonso et al. (2008) and Rantakari (2008). However the partition must satisfy the following difference equation \( k_n + 1 - k_n = k_n - k_{n-1} + \frac{4k_n}{\phi} \).

\[ 13 \] 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 communication loss to be zero, since the (observed) decision of the leader perfectly reveals the state.
The $V(\theta_i)$ is the variance of the local state being communicated, and the next term is the percentage of the total variance lost due to imperfect communication. 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 communication would entail $\phi = \infty$. In the sequential structure having perfect communication occurs only when $r_1 = 1$, namely the leader cares about only coordination.

**Lemma 3** Equilibrium loss for the leader is

$$ (1 - r_1)\gamma_1 \left( V(\theta_2) + V(\theta_1) \right) + \gamma_1 \beta_1 c_2. $$

Meanwhile, equilibrium loss for the follower is

$$ (1 - r_2)r_2(1 - \gamma_1)^2 \left( V(\theta_2) + V(\theta_1) \right) - (1 - r_2)r_2(\gamma_1^2 - 2\gamma_1)c_2. $$

In Rantakari (2008) units may want to commit to not listen to the message of another firm; however, in this case one can see the leader always wants to listen to the follower since $\gamma_1 \beta_1 > 0$. The follower would also prefer the leader to listen since $\gamma_1 < 1$, and hence the term in front of $c_2$ is positive. Combining the above lemma with the statements in Rantakari (2008) yield the following lemma.

**Lemma 4** The losses can be summarized as follows:

$$ L_{\text{leader}} = (V(\theta_1) + V(\theta_2))(1 - r_1)\gamma_1 + V(\theta_2)\gamma_1 \beta_1 V(\phi^{\text{eq}}) $$

$$ L_{\text{follower}} = (V(\theta_1) + V(\theta_2))(1 - r_2)r_2(1 - \gamma_1)^2 - V(\theta_2)(1 - r_2)r_2(\gamma_1^2 - 2\gamma_1) V(\phi^{\text{eq}}) $$

$$ L_{\text{simultaneous}} = (V(\theta_1) + V(\theta_2)) \frac{(1 - r_1)\beta_1}{(1 - r_1 r_2)^2} + V(\theta_1) r_1 (1 - r_1) r_1 \frac{(1 - r_1 r_2)^2 - (1 - r_1)^2}{(1 - r_1 r_2)^2} V(\phi^{\text{sim}}_1) $$

$$ + V(\theta_2) r_1 (1 - r_2)^2 \frac{(1 - r_1 r_2)^2 - (1 - r_1)}{(1 - r_1 r_2)^2} V(\phi^{\text{sim}}_2) $$

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

The next two propositions illustrate the stark tradeoff between simultaneous and sequential decision making. Proposition 2 shows that to increase communication quality decisions should be made sequentially and the unit that has a higher need to adapt to local conditions or has more volatile local conditions should move first. Meanwhile Proposition 1 shows that if communication quality is exogenously, then given coordination increases when
decisions are simultaneous and the unit that has a lower need to adapt to local conditions should move first.\textsuperscript{15} Namely, I assume that under all decision orderings the informational partitions determining the messaging are exogenously given, equal, and not subject to truth telling constraints.

**Proposition 1** For a fixed communication quality the following are true:

i Concurrent decision making always yields less coordination loss than either sequential timing.

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 is coordination is a public good and allowing one unit to move first reduces the total amount contributed since first movers use their first mover advantage to strategically under-invest. The intuition for the second statement in the proposition is that since the leader has an incentive to under-invest in coordination, the ideal leader is one who would not have invested much in coordination anyway.\textsuperscript{16} However, the next proposition shows that to induce better communication sequential decision making is optimal. Additionally, even when restricting the comparison to the two sequential timings Proposition 1 says to increase coordination the unit that cares more about coordination should move second whereas Proposition 2 says this unit should move first to increase communication.

**Proposition 2**

i At least one of the two sequential timings induces less aggregate communication loss than a simultaneous timing despite worse communication about the followers state.

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

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

\textsuperscript{15}It is worth noting that a necessary condition for coordination is the ability to know the other unit’s state, which comes from communication.

\textsuperscript{16}This proposition does not rely on the distributional assumption of \( \theta \). Additionally, this proposition holds not just in expectation over \( \theta \) but for any realization of \( m_1(\theta_1), m_2(\theta_2) \) in equilibrium.
The intuition for the first statement is that sequential decision making always yields perfect information revelation about the leaders state and hence there will be at least one way to order the two units to beat the concurrent timing. Meanwhile, while the aggregate communication loss is lower the communication loss for the follower is higher because the leader overadapts to their state relative to a simultaneous timing causing the follower to want to exaggerate their message yielding worse communication 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 try and influence the decisions of the other unit. This causes communication quality to increase when units that don’t care about coordination follow and thus are the ones needing to send cheap talk messages.

Finally, recall that by letting a unit lead there is no information loss about their state. Hence, if the two unit’s have the same incentive to send biased messages, letting the one with more information needing to be transmitted communicate perfectly via revelation of decisions will reduce aggregate communication loss. The next section reports on whether the effects in Proposition 1 or 2 dominates as the parameters of the model vary.

5 Optimal Governance

Given the solutions above, I can calculate the optimal governance structure for any given parameter values. Below are the graphs that compares both versions of sequential with simultaneous. When the units care approximately equally about coordination, the public goods effect described in Proposition 1 begins to dominate the increased communication described in Proposition 2, and the optimal governance is simultaneous. However, upon moving further away from the diagonal, one sees that sequential becomes the optimal timing as now the public goods effect is small in comparison to the coordination effect.

Additionally, when \( r_1 \) and \( r_2 \) are sufficiently small, simultaneous contributions are 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. However, in the sequential game, no matter how uninformative the cheap talk is, having the follower see the payoff relevant decision allows some communication. In fact, for sufficiently small \( r_1 \) and \( r_2 \), either sequential structure improves upon the simultaneous one.

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

![Optimal Governance Structure](image)

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

Figure 2: Black is unit 1 leads, Red is unit 2 leads, and Yellow is concurrent. Note that only the relative volatilities matter in determining the optimal governance structure.

Below is the plot for the optimal governance of unit 1 from the perspective of unit 1. As the weight the firm places on unit 1 increases, Figure 2 will continuously form into the one below. A unit that cares sufficiently about coordination would in fact prefer to move second to ensure 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 invest in adaptation knowing the second mover will coordinate on their behalf. Additionally, as seen in the plot on the right, when unit 1 has more volatility, unit 1 would prefer to move first more frequently. This is because unit 1 now has a great incentive to adapt to their local conditions and being a leader allows them to do so. For these reason, with sufficiently asymmetric values of $r_1$ and $r_2$ or sufficiently asymmetric volatilities, moving to a sequential structure is in fact a Pareto improvement upon simultaneous.\(^{17}\) Moving concurrently balances these effects, but as shown below is almost never preferred by a single unit. However, recall the \(^{17}\)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,
headquarters is assumed to choose the timing that maximizes the sum of profits.

As I mentioned in Section 3, I do not consider other governance structures that involve moving a decision right from one unit to another or the headquarters. However one can show that allowing one unit to make both decisions (Directional Authority)\footnote{The calculations for this structure are quite simple as now $d_i = d_j = \theta_j$, $L_i = 0$, and $L_j = 2(1 - r_j)\mathbf{V}(\theta)$} is \textit{not} optimal with equally weighted units when allowing for sequential decision-making. This dominance occurs despite Rantakari (2008) showing that Directional Authority is the optimal governance structure in the presence of equally weighted units within a broader set of governance structures (all with concurrent decision making, and including centralization). The intuition for Directional Authority is that if unit $i$ cares sufficiently about coordination, then they want $d_i = \theta_j$ to ensure they can coordinate with unit $j$. To accomplish this, one can give unit $j$ both decision rights. However, the firm could instead require that unit $j$ moves first to ensure that unit $i$ knows $\theta_j$ but can still adapt to their own state. Sequential decision making improves the communication, at a much lower adaptation loss than under Directional Authority. Both Directional Authority and Centralization aim to improve coordination, but do so at the expense of communication, meanwhile sequential decision making does so by improving communication. Finally, decentralization, in the simultaneous setting, is opti-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures.png}
\caption{Figure 3: Black is unit 1 leads, Red is unit 2 leads, and Yellow is concurrent.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures.png}
\caption{(a) Equal volatility across units. (b) Unit 1 has double unit 2's volatility.}
\end{figure}


\footnote{The calculations for this structure are quite simple as now $d_i = d_j = \theta_j$, $L_i = 0$, and $L_j = 2(1 - r_j)\mathbf{V}(\theta)$}
mal for units that care sufficiently about adaptation. However, in these parameter values, sequential decision-making is optimal since it allows for much improved communication.

5.1 Implications within Examples

Takeuchi and Nonaka (1986) discuss what types of firms would benefit from abandoning “the old, sequential approach to developing products.” They suggest that firms that are involved in breakthrough innovations, such as chemical research firms, would not benefit from developing their product concurrently. In terms of the model, the research team is so focused on matching their state, which can be viewed as discovering a working chemical, \( r_{\text{research}} \approx 0 \). Additionally the sales units are more concerned with finding the consumer niche that wants the specific chemical that was developed rather than matching their exact local state, \( r_{\text{sales}} \approx 1 \). As Figure 2 shows, this asymmetry is precisely when sequential development is optimal. Additionally, one could assume that the research units experience greater volatility than sales thus Figure 2 would also suggest that the increased volatility adds a second reason why the research unit should move first. However, Takeuchi and Nonaka (1986) also give the example of Canon. Here engineering and sales may not have such asymmetric volatilities or polar opposite values on coordination, and consistently with Figure 2, they suggest concurrent development.

Meanwhile, Hayes and Wheelwright (1979) suggest that firms with non-standardized products should use simultaneous product development. An example of non-standardized products given in the article is a commercial printing company. Since there are many different types of documents needing to be produced and each job is unique, each unit experiences volatility and thus there is enough symmetry in the volatilities that simultaneous is optimal. Meanwhile, an example in the article of a product that is made sequentially is a car manufacturer. A brakes unit is going to be adapting to potentially extreme conditions, yielding a high volatility; units such as paint will not encounter these extreme conditions and have low volatility. Hence, the observed sequential structure with the engine being produced first is consistent with the predictions of Figure 2.

Dyer (1996) describes how Chrysler changed their relationships with suppliers and due to this change moved from sequential to concurrent development. Beforehand, Chrysler never committed to a supplier, leaving Chrysler to have a low value of coordination with a given supplier, low \( r_{\text{Chrysler}} \). Meanwhile, each supplier knew that if in the next year Chrysler found
a cheaper supplier, Chrysler would abandon them. The possibility that Chrysler might terminate its relationship with a supplier made suppliers less interested in coordinating with Chrysler, i.e. \( r_{\text{Supplier}} \) to be greater than \( r_{\text{Chrysler}} \), but both small. As one can see in Figure 2, these low parameter values yields sequential development with Chrysler designing the parts first and then the suppliers deciding how to make the parts.\(^{19}\) When Chrysler decided to commit to their suppliers, Chrysler was now incentivized to make relationship specific investments. These investments make miscoordination more costly, as now they can not instead use these machines with other firms. Additionally, because Chrysler and their manufacturer’s previous levels of coordination were sufficiently asymmetric sequential development was optimal, but now Chrysler’s coordination incentives increased to the point where concurrent became optimal as can be seen in Figure 2.

One explanation Dyer (1996) gives for why the new governance between Chrysler and their suppliers improved performance was communication. Previously, Chrysler would not listen to their suppliers but after moving to concurrent development Chrysler created joint meetings for manufacturers to share their opinions on proposed designs. These joint meetings could be viewed as a relationship specific investment as described above. However, another interpretation is that by moving to concurrent timing Chrysler has a greater incentive to listen to their manufacturers. Proposition 2 suggests that since manufacturing now moves concurrently with design, the dyad is able to engage in better communication about manufacturing’s needs. My results suggest that an important determinant in allowing someone to be a first mover is their ability to first listen to their followers.

6 Conclusion

The organizational economics literature has focused on who should make decisions rather than when the decisions should be made by a fixed set of decision-makers. I present a model of optimal timing of decisions within and between organizations which can determine when to have sequential or simultaneous decision-making. In my setting, the fundamental trade off is that sequential decision-making allows for increased communication due to the revelation of a decision. Because in my model actions always speak louder than words, the decision

\(^{19}\)While the payoff function does not include a “quality” term, one can think of the quality as being a function of the adaptation and coordination losses. Thus, when Chrysler improved these terms the quality increased.
revelation induces better communication. However, sequential decision-making encourages the first mover to under-invest in coordination, knowing the second mover will overcompensate. My model can compute when each of these effects dominates and shows sequential is better when the preferences for coordination between two units are both sufficiently high and similar.

While the focus of this paper has been on the choice of the optimal decision timing given exogenous payoff parameters, it remains to explore 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. Additionally, the headquarters may also have an incentive to make a unit sufficiently focused on coordination so that they have no motivation to preempt the decisions of the primary unit if the primary unit is sufficiently important to the firm. Studying how the timing and payoff parameters move together is an interesting avenue for further research.

7 Appendix

7.1 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 $\theta_2$, unit 2 chooses

$$d_2 = (1 - r_2)\theta_2 + r_2d_1.$$  

(1)

As a function of $d_1$ and $\theta_2$ this gives unit 2 a loss of

$$U_2(\theta_2, d_1) = (1 - r_2)r_2(d_1 - \theta_2)^2.$$  

(2)

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

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

Substituting in equation 1 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
\[ 1 - r_1(d_1 - \theta_1)^2 + \beta_1(m_i - d_1)^2. \]

Taking a First Order Condition yields
\[ d_1 = \frac{1 - r_1\theta_1 + \beta_1 m_i}{1 - r_1 + \beta_1} := (1 - \gamma_1)\theta_1 + \gamma_1 m_i. \]

**Proof of Lemma 2.** Given equation 2, this generates the following indifference condition:

\[ U_2(k_i, k_{i+1}) = U_2(k_i, k_i) \]
\[ \iff (d_1(k_{i+1}) - k_i)^2 = (d_1(k_i) - k_i)^2 \]

Using uniform assumption \( k_i - \gamma_1 m_i = \gamma_1 m_{i+1} - k_i \)

Now \( \iff k_i - \gamma_1 m_i = \gamma_1 m_{i+1} - k_i \)

\( \iff k_{i+1} - k_i = k_i - k_{i-1} + \frac{4}{\gamma_i} k_i \)

\( \iff k_{i+1} - k_i = k_i - k_{i-1} + \frac{4}{\gamma_i} k_i \)

Can take from Alonso et al. (2008) 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 \( V(\theta_i) - V(m_i) = \frac{1}{3} \frac{1}{3^\theta + 4} \). 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

\[
\mathbb{E}\left((1 - r_1)(d_1 - \theta_1)^2 + r_1((1 - r_2)\theta_2 + r_2d_1 - d_1)^2|i\right)
= (1 - r_1)\mathbb{E}((d_1 - \theta_1)^2|i) + r_1(1 - r_2)^2\left(\mathbb{E}((\theta_2 - (1 - \gamma_1)d_1 - \gamma_1 m_i)^2|i)\right)
\]

Recall \( d_1 = (1 - \gamma_1)\theta_1 + \gamma_1 m_i \)
\[ = (1 - r_1)\mathbb{E}((-\gamma_1 \theta_1 + \gamma_1 m_i)^2|i) + r_1(1 - r_2)^2\left(\mathbb{E}((\theta_2 - (1 - \gamma_1)\theta_1 - \gamma_1 m_i)^2|i)\right)
\]
\[ = (1 - r_1)\gamma_1^2\mathbb{E}((\theta_1 - \bar{m}_i)^2|i) + r_1(1 - r_2)^2\left((1 - \gamma_1)^2 V(\theta_1) + \mathbb{E}(\theta_2^2|i) + \gamma_1^2 (\bar{m}_i)^2 - 2\gamma_1(\bar{m}_i)^2\right)
\]
\[ = (1 - r_1)\gamma_1^2(\mathbb{E}(\theta_1^2|i) + \bar{m}_i^2) + \beta_1\left((1 - \gamma_1)^2 V(\theta_1) + \mathbb{E}(\theta_2^2|i) + (\gamma_1^2 - 2\gamma)(\bar{m}_i)^2\right).
\]
Taking the expectation across intervals yields
\[
\sum_{i=1}^{N} P(i) \left( (1 - r_1)\gamma_1^2 (E(\theta_1^2 | i) + \bar{m}_i^2) + \beta_1 (1 - \gamma_1)^2 V(\theta_1) + E(\theta_1^2 | i) + (\gamma_1^2 - 2\gamma)(\bar{m}_i)^2 \right)
\]
\[
= (1 - r_1)\gamma_1^2 (v(\theta_1) + E(\bar{m}_i^2)) + \beta_1 (1 - \gamma_1)^2 V(\theta_1) + v(\theta_2) + (\gamma_1^2 - 2\gamma)(\bar{m}_i)^2)
\]

I can re-arrange this in terms of loss from misc-ordination and loss from bad communication as follows where \(c_i := V(\theta_i) - E(\bar{m}_i^2)\) is the loss from communication:
\[
= (1 - r_1)\gamma_1^2 (v(\theta_1) + V(\theta_2) - c_2) + \beta_1 (1 - \gamma_1)^2 V(\theta_1) + v(\theta_2) + (\gamma_1^2 - 2\gamma)(v(\theta_2) - c_2)
\]
\[
= \left(v(\theta_2) + v(\theta_1)\right) (1 - r_1)\gamma_1^2 + \beta_1 (1 - \gamma_1)^2 - c_2 (1 - r_1)\gamma_1^2 + \beta_1 (\gamma_1^2 - 2\gamma)
\]
\[
= (1 - r_1)\gamma_1 \left(v(\theta_2) + v(\theta_1)\right) + \gamma_1 \beta_1 c_2.
\]

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 \(i\)
\[
E((d_1 (k(\theta)) - \theta)^2 | i) = E((1 - \gamma_1)\theta_1 + \gamma_1 \mu_i - \theta_2)^2 | i)
\]
\[
= (1 - \gamma_1)^2 V(\theta_1) + E((\theta_2 - \gamma_1 \mu_i)^2 | i)
\]
\[
= (1 - \gamma_1)^2 V(\theta_1) + E(\theta_2^2 | i) + \gamma_1^2 (\bar{m}_i)^2 - 2\gamma_1 (\bar{m}_i)^2.
\]

Now taking the expectation across intervals
\[
(1 - r_2) r_2 \sum_{i=1}^{N} P(i) \left( (1 - \gamma_1)^2 V(\theta_1) + E(\theta_2^2 | i) + \gamma_1^2 (\bar{m}_i)^2 - 2\gamma_1 (\bar{m}_i)^2 \right)
\]
\[
= (1 - r_2) r_2 \left( (1 - \gamma_1)^2 V(\theta_1) + V(\theta_2) + (\gamma_1^2 - 2\gamma_1) E(\bar{m}_i^2) \right).
\]

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 \(c_i = V(\theta_i) - E(\bar{m}_i^2)\), so
\[
= (1 - r_2) r_2 (1 - \gamma_1)^2 \left(v(\theta_2) + v(\theta_1)\right) - (1 - r_2) r_2 (\gamma_1^2 - 2\gamma_1)c_2.
\]

**Proof of Lemma 5.** The losses for sequential decision-making are taken from Lemma 3. The losses for simultaneous decision-making are taken from Rantakari (2008). ■
**Proof of Proposition 1.** Given a fixed communication quality, I will 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:

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

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

$$\frac{(1 - r_2)^2(1 - \gamma_1)^2}{(1 - r_1) + r_1(1 - r_2)^2} < \frac{(1 - r_2)}{(1 - r_2) + r_2(1 - r_1)^2}$$

$$\iff (1 - r_2) + r_2(1 - r_1)^2 < (1 - r_1 + r_1(1 - r_2)^2$$

$$\iff r_2(-2r_1 + r_1^2) < r_1(-2r_2 + r_2^2)$$

$$\iff -2 + r_1 < -2 + r_2$$

$$\iff r_1 < r_2.$$

**Statement i of Proposition 1:**

Meanwhile coordination loss when the units move at the same time is bounded above by 

$$(\frac{(1 - r_1)(1 - r_2)}{1 - r_1r_2})^2E(\theta_2^2 + \theta_1^2|m_1, m_2).$$

Comparing this to the sequential structure yields

$$\iff -r_1 + r_1(1 - r_2)^2 < -r_1r_2$$

$$\iff r_2 < 1.$$  

**Proof of Proposition 2.**

**Statement i of Proposition 2:** As mentioned before, there is perfect communication on the state of the leader, so the total communication loss within the sequential timing is $V(\theta_1)\frac{1}{4 + 3\beta_1}$. Note that this is always less than $\frac{V(\theta_1)}{4}$. It is without loss to assume $V(\theta_2) <$ 

20Since I will 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.
V(θ₁) and I will show below that sequential decision making with unit 1 leading always incurs less communication loss than concurrent decision making. Note that in concurrent decision-making, communication loss occurs on both local states and

\[ L_{\text{simultaneous}} = \frac{V(\theta_1)}{4 + 3\phi_{\text{simultaneous}_1}} + \frac{V(\theta_2)}{4 + 3\phi_{\text{simultaneous}_2}} \]

\[ \iff \frac{\phi_{\text{simultaneous}_1}\phi_{\text{simultaneous}_2} - \frac{16}{9}}{(\phi_{\text{simultaneous}_1} + \frac{4}{3})(\phi_{\text{simultaneous}_2} + \frac{4}{3})} < 0 \iff \phi_{\text{simultaneous}_1}\phi_{\text{simultaneous}_2} < \frac{16}{9} \iff r_1 r_2 < \frac{16}{9}. \]

Thus the sequential communication loss is always bounded above by \( \frac{V(\theta_2)}{4} \); meanwhile, the simultaneous communication loss is bounded below by \( \frac{V(\theta_2)}{4} \). The latter statement that there is worse communication of the followers state follows by noting that \( \phi_{\text{simultaneous}} > \phi_{\text{sequential}} \) and that communication loss is monotonic in \( \phi \).

**Statement ii of Proposition 1:**
Recall that when decisions are made sequentially the only communication loss is from the follower. This loss is a monotone decreasing function of \( \phi_i \) when unit \( i \) leads. The algebra is as follows:

\[ \frac{\beta_1}{(1 - r_1)} \geq \frac{\beta_2}{(1 - r_2)} \iff \beta_1(1 - r_2) \geq \beta_2(1 - r_1) \iff r_1(1 - r_2)^3 \geq r_2(1 - r_1)^3 \iff \frac{r_1}{(1 - r_1)^3} \geq \frac{r_2}{(1 - r_2)^3} \iff r_1 \geq r_2. \]

**Statement iii of Proposition 1:**
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 total communication loss. ■
7.2 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 communication 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 communication. 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 communication gain from switching governance structures is largest in intermediate values of $r_i$ and $r_j$, because for sufficiently high(low) values on coordination, communication 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 communication 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.

![Optimal Sequential Governance Structure](image)

Figure 4: Black is unit 1 leads and Red is unit 2 leads.
7.3 Full Information Analysis

When the states are known by both parties, very broadly this is a public goods provision problem. The more a unit sacrifices compromising on their own state and move towards the other unit’s state 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 complete information and unit $i$ observes both $\theta_i$ and $\theta_{-i}$ is equivalent to 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 communication. Since there is full information, communication does not matter, and this was the only benefit to sequential structures. However, as the weight the headquarters places on unit 1 increases, sequential structures will eventually become optimal because leading is best for unit 1.

7.4 Incomplete Information Analysis

To solve for the utilities of each unit, I can simply take Proposition ?? 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:

- $\frac{(1-r_1)\beta_1}{1+r_1} V(\theta_1) + \beta_1 V(\theta_2)$ when unit 1 leads
- $r_1 (1-r_1) \left[ V(\theta_1) + \frac{\alpha_2^2}{(\alpha_2 + \beta_2)} V(\theta_2) \right]$ 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

---

21To see this, one can confirm that the losses from Lemma 1 when setting $c_i = 0$ are always larger under a sequential timing.
able to lower their adaptation loss since

\[
\frac{(1 - r_2)^2}{(1 - r_2) + r_1(1 - r_2)^2} < 1,
\]

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.

Figure 5: Black is unit 1 leads, Red is unit 2 leads, and Yellow is Concurrent.

It is worth noting 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 followers state and are thus unable to coordinate.

The optimal governance structure for the firm is plotted below. The firm places the unit that cares more about coordination as a second mover. Additionally, concurrent timing is never optimal, since sequential always gives communication but concurrent gives none.
Figure 6: Black is unit 1 leads, Red is unit 2 leads, and Yellow is Concurrent.

References


