Dynamics of Group Decision-Making

by

Roi Orzach

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2025

Authored by: Roi Orzach
Department of Economics
May 15, 2025

Certified by: Glenn Ellison
Gregory K. Palm (1970) Professor of Economics, Thesis Supervisor

Certified by: Stephen Morris
Peter A. Diamond Professor of Economics, Thesis Supervisor

Accepted by: INSERT
INSERT
INSERT

THESIS COMMITTEE

Thesis Supervisor

Glenn Ellison
Gregory K. Palm (1970) Professor of Economics
Department of Economics

Thesis Readers

Glenn Ellison
Gregory K. Palm (1970) Professor of Economics
Department of Economics

Stephen Morris
Peter A. Diamond Professor of Economics
Department of Economics

INSERT
INSERT
INSERT

Dynamics of Group Decision-Making

by

Roi Orzach

Submitted to the Department of Economics
on May 15, 2025 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

This thesis comprises three chapters, all focused on Microeconomic Theory, specifically the dynamics of decision-making. The first explores how the desire for conformity results in long-run misperceptions due to uninformative decision-making. The second chapter studies multi-project collaborative experimentation. The final chapter analyzes whether decentralized organizations should utilize sequential or concurrent decision-making.

The first chapter notes that in many settings, individuals imitate their peers' public decisions for one or both of two reasons: to adapt to a common fundamental state, and to conform to their peers' preferences. In this model, the fundamental state and peers' preferences are unknown, and the players learn these random variables by observing others' decisions. With each additional decision, the public beliefs about these unknowns become more precise. This increased precision endogenously increases the desire to conform and can result in decisions that are uninformative about a player's preferences or perceptions of the fundamental state. When this occurs, social learning about peers' preferences and fundamentals ceases prematurely, resulting in inefficient decisions. In line with findings from social psychology, I show that interventions aimed at correcting misperceptions of peers' preferences may lead to more efficient decision-making in settings where interventions aimed at correcting misperceptions of the fundamental state may have no effect.

The second chapter (joint with Charles Angelucci) analyzes collaborative experimentation across multiple independent domains. Each domain contains infinitely many potential projects with asymmetric benefits. In each period and domain, two players can idle, jointly explore a new project, or jointly exploit a known one, with voluntary transfers. For intermediate discount factors, treating domains as independent during experimentation is suboptimal. The optimal experimentation policy exhibits common features of collaborative experimentation: lengthy exploration, temporary project exploitation, recall of past projects, and inefficient initial or terminal idling within certain domains. We connect these findings to research on buyer-supplier dynamics and persistent productivity differences.

The final chapter examines how the timing of decision-making shapes the allocation of decision rights within an organization. Here, I analyze concurrent versus sequential decision-making in a model where two units 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 compared to concurrent decision-making. However, first movers also have an incentive to over-adapt to their state, knowing second movers will conform to their decision. A surplus-maximizing headquarters prefers sequential decision-making to concurrent if and only if (i) the two units' local conditions have sufficiently different volatilities and (ii) their need to coordinate is sufficiently asymmetric or low. Finally, sequential decision-making is shown to be optimal even when allowing for additional governance structures involving the reallocation of decision rights across the units and the headquarters and is shown to render some commonly-analyzed forms of decentralization sub-optimal.

Thesis supervisor: Glenn Ellison
Title: Gregory K. Palm (1970) Professor of Economics

Thesis supervisor: Stephen Morris
Title: Peter A. Diamond Professor of Economics

Acknowledgments

This thesis represents the culmination of a six-year long journey that taught me both how to be a better economist, but also a better person.

First and foremost, I am deeply grateful to my main advisor, Glenn Ellison, whose guidance and generosity have been truly inspiring. His unwavering support and dedication to his students are unparalleled. While there are countless ways in which Glenn has been an exceptional mentor, one particular instance stands out. Less than three months after I first met him, the COVID-19 pandemic began, and he immediately offered me a place to stay should the graduate dorms close. This act of kindness exemplifies the generosity he has continued to show throughout my time as his student.

I am also deeply grateful to the Sloan Applied Microeconomics group, particularly Charles Angelucci, Alessandro Bonatti, and Robert Gibbons. Their attention to detail knew no bounds—whether it was footnote 12 in an online appendix or the proper way to button a suit. Additionally, I extend my sincere thanks to Stephen Morris, an outstanding advisor whose willingness to discuss both contemporary research and revisit his nearly 25-year-old paper—the very work that first inspired my interest in reputational concerns—has been invaluable. While only Robert Gibbons *studies* effective management practices, all of my advisors *practice* incredible management practices and I could not have completed my dissertation without them. Finally, I've also benefited from insightful conversations with Jean Tirole and the late Heikki Rantakari.

One of the joys of graduate school has been learning from my exceptional colleagues. Thanks especially to Sarah Gertler, Pedro Martinez-Bruera, Jackson Mejia, Paul-Henri Moisson, Daniel O'Connor, Kramer Quist, Upamanyu Sharma, and Henry Zhang.

Finally, I wish to thank my family: both old and new. I wouldn't be here had it not been for the 21 years of investment into my education by my parents before entering graduate school. My sister, Shelli, and her dog, Knox, helped me through my first years of graduate school. Finally, I wouldn't have been able to make it through the home stretch without my soon-to-be wife, Maya, and our amazing dog, Mocha.

Contents

List of Figures

Chapter 1

Conformity Concerns: A Dynamic Perspective

Roi Orzach[1]

1 Introduction

Substance abuse is an important public health concern and greatly contributes to mortality rates across the world ((cf. Sheikh et al., 2018)). For decades public health officials educated adolescents on the cost of substance abuse, but these "initial attempts at prevention were ineffective because they focused primarily on lecturing students about the dangers and long-term health consequences of substance use" Griffin and Botvin (2010). These attempts failed because they assumed an "individual-oriented" (Prentice and Miller, 1993) cognitive model about their participants: individuals' decisions about substance use were imagined to be based purely on anticipated health outcomes for the individual. In contrast, informing students about their peers' true preferences towards substance use was much more effective: Schroeder and Prentice (1998) found that undergraduates who learned their peers' preferences towards alcohol reduced their consumption by 40 percent compared to those who were informed about just the health costs. Schroeder and Prentice (1998) call this a "peer-oriented" cognitive model.

Public health research rationalized this finding by noting that many individuals overestimate how much their peers enjoy alcohol (cf. Prentice and Miller, 1993). This misperception, combined with a desire to conform, motivated people to partake in substance abuse. Such phenomena are pervasive. Political endorsements (Loury, 1994; Geiger and Swim, 2016), female labor force decisions (Bursztyn et al., 2020a), and corporate board decisions (Westphal and Bednar, 2005; Chang et al., 2019) are interpreted as influenced in similar ways. Across

[1]Department of Economics, MIT, 50 Memorial Drive, Cambridge MA 02142 (e-mail: orzach@mit.edu).
I thank Alessandro Bonatti, Glenn Ellison, Robert Gibbons, and Stephen Morris for countless conversations and additionally thank Chiara Aina, Charles Angelucci, Ian Ball, Abhijit Banerjee, Inga Deimen, Stefano DellaVigna, Tore Ellingsen, Drew Fudenberg, Ying Gao, Saumitra Jha, Jackson Meija, Navin Kartik, Simone Meraglia, Paul-Henri Moisson, Ellen Muir, Vishan Nigam, Shelli Orzach, Jacopo Perego, Nicola Persico, Kramer Quist, Daniel Rappoport, Rafaella Sadun, Tomasz Sadzik, Lones Smith, Cass Sunstein, Jean Tirole, Jonathan Weinstein, Alexander Wolitzky, as well as participants at the MIT Organizational Economics and Theory groups, 2024 North American Summer Meeting of the Econometric Society, and 2024 Stony Brook Game Theory Conference for helpful comments.

these examples, inefficiencies may arise from two sources: misperceptions of the private benefits of such decisions and misperceptions of peers' attitudes towards these decisions.

In this paper, I provide a general framework that explains how these two misperceptions interact, why they hence may both persist, and why some interventions may be more successful than others. In the model, a community of agents attempts to learn two initially unknown variables: a fundamental state (which could be the health costs of substance abuse) and their peers' average preference type. Learning occurs both privately and socially. Privately, each individual receives a signal about the fundamental state and his own preference type, which is informative about his peers' preference types because they are drawn from the same population distribution. Further, the players may be able to learn the unknown variables by observing the decisions of their peers. In line with the social learning literature (cf. Banerjee, 1992; Bikhchandani et al., 1992), I ask: Given an infinite sequence of decisions, can the public correctly infer these two unknown variables? In answering this question, I depart from the "individual-oriented" model and explicitly model the desire to conform. This approach generates new predictions regarding the success or failure of social learning. Additionally, in cases where social learning fails, it provides novel insights into optimal interventions.

Formally, each individual seeks to maximize a utility function that combines their private utility with a conformist utility component. Their private utility is determined by how well their decision adapts to their private preference type and the fundamental state. Their conformist utility, on the other hand, is determined by the exogenous conformity concerns (a parameter common to all individuals) multiplied by the gap between how the community perceives their preference type and the true average preference type of the community. While the exogenous conformity concerns is constant over time, the endogenous reputational penalty from choosing decisions different from one's peers' decisions changes as the public beliefs about the fundamental state and average preference type become more precise. I refer to this endogenous reputational penalty as the "effective conformity concerns."

The key mechanism potentially preventing social learning is that once the effective conformity concerns are sufficiently large, an individual is unable to adapt to his preference type without incurring a large reputational penalty. Further, because adaptation to a player's private signal about the fundamental state might be attributed to a player's preference type, such adaptations are similarly discouraged. Therefore, when the effective conformity concerns are large enough each decision is independent of a player's private information. This independence implies that no new information is publicly learned in the current period, resulting in an identical situation for the next period, and, by induction, for all subsequent periods. As a result, if the effective conformity concerns ever become sufficiently large, all subsequent players will pool on inaccurate perceptions of the fundamental state and of the average preference type.

Finally, interventions to improve efficiency are needed primarily when social learning fails. Because social learning fails when the effective conformity concerns are large, then when interventions are needed, they should target conformity concerns (via information about the average preference type) as opposed to adaptation loss (via information about the fundamental state).

In Section 3, I first analyze a benchmark static model where players have common knowledge about both the fundamental state and their peers' average preference type but differ in their own preference types, as described above. In this benchmark, exogenous conformity concerns are still a parameter shared by all individuals, however in this static benchmark there is no learning, implying that the effective conformity concerns are equal to the exogenous conformity concerns. In this scenario, conformity concerns place a penalty on the degree to which players adapt to their preference types (cf. Bernheim, 1994). I show that all players choose the same decision independent of their preference types if and only if the conformity concerns exceed a given threshold (Proposition 1).[2] This benchmark shows how, when conformity concerns are significant, an individual's decisions may cease to reflect their true preferences.

In Section 4, I analyze the complete model where conformity concerns interact with uncertainty about the fundamental state and the average preference type. The necessary and sufficient condition for "social learning to occur" (i.e., for the players to learn asymptotically the fundamental state and the average preference type) is that there exist infinitely many periods of a decision rule with revelation. This condition requires that the decision rule involves revelation when the beliefs about the fundamental state and the average preference type are arbitrarily precise. As a result, a necessary condition for social learning to occur is that the conformity concerns must be less than the threshold for revelation in the benchmark (Lemma 2).

Unsurprisingly, when exogenous conformity concerns are sufficiently small, this places an upper bound on the effective conformity concerns. Consequently, if exogenous conformity concerns are sufficiently small, learning occurs regardless of prior beliefs. Finally, I show that when the exogenous conformity concerns are intermediate, the learning outcome depends on the initial beliefs (Proposition 2).

When exogenous conformity concerns are such that learning fails in the limit, the players face three inefficiencies. First, the players are unable to adapt to their preference types. Further, the players' decisions are based on imprecise beliefs about *both* the fundamental state and the average preference type of their peers.

Given these two imprecise beliefs, Section 5 analyzes the different effects of *peer-oriented* interventions, which inform players about their peers' preference types (e.g., how "cool" substance abuse is thought to be by others), versus *individual-oriented* interventions, which inform players about the fundamental state (e.g., the health costs of substance abuse).[3] If individuals had no desire for conformity, they would disregard peers-oriented interventions entirely. In such a scenario, an individual-oriented intervention would always be the preferred approach. However, in such situations interventions are, arguably, not needed; if there were no desire for conformity-then, asymptotically, the players would eventually learn the fundamental state. For this reason, I compare the effect of peer-oriented and individual-oriented interventions based on their ability to break a pooling equilibrium when the exogenous conformity concerns are high.

When exogenous conformity concerns are high, I show that if the beliefs about the fundamental state and the average preference type are uncorrelated, then individual-oriented

[2] In Bernheim (1994), equilibria with partial revelation may exist. I discuss the difference between my benchmark and the environment in Bernheim (1994) in Section 6.2. Reassuredly, in Bernheim (1994) when conformity concerns are sufficiently high, the unique equilibrium is also fully pooling.

[3] Throughout, I assume that interventions are credible. See Benabou and Tirole (2024) for an analysis where the party conducting the intervention has differing preferences from the community resulting in a commitment problem.

interventions can never break a pooling equilibrium. Intuitively, players may deviate from past decisions for one of two reasons: first, if a player's signal implies the fundamental state differs from the public beliefs; second, if a player's preference type motivates a different decision from his peers' decisions. Importantly, because deviations are only partially ascribed to a player's preference type, imprecise beliefs about the fundamental state encourage adaptation, implying that individual-oriented interventions increase the effective conformity concerns and will not break a pooling decision rule.

In contrast, peer-oriented interventions may break a pooling outcome. There are two competing effects of a peer-oriented intervention. First, when players are unsure what the average preference type is, their own preferences serve as an informative signal. Given that a player wants the public perception of his preference type to be close to the average preference type, choosing a decision that is responsive to his own preference type is optimal; hence, a non-pooling equilibrium will exist. The competing force is that if a player is unsure what the average preference type of the population is, then the population is also uncertain of that player's preference type. The latter uncertainty implies that the population's inference about that player's preference type places a greater weight on his decision rather than the prior. I show that if the uncertainty about the fundamental state is large, the competing effect is stronger, and thus peer-oriented interventions may break a herd. Generalizing the definition in Bikhchandani et al. (2021), the pooling equilibrium is "fragile" to peer-oriented interventions but not to individual-oriented interventions (Proposition 4). Moreover, even when neither intervention can break a pooling equilibrium, the peer-oriented intervention may have a more substantial impact on the decision players pool on (Proposition 5). This is because a peer-oriented intervention enables players to pool on a more efficient decision.

My findings are consistent with several key studies (e.g., Schroeder and Prentice, 1998; Bursztyn et al., 2020a; Gulesci et al., 2023) advocating for peer-oriented interventions in situations with significant conformity concerns. Furthermore, my research provides a mechanism supporting Sunstein's (2019) argument that sharing information about others' preferences can lead to the revelation of true beliefs and preferences. As Sunstein notes, this process "unleash[es]" people, allowing them to "reveal what they believe and prefer," ultimately facilitating the discovery of "preexisting beliefs, preferences, and values."

Further, in my model, before a peer-oriented intervention, the players will have incorrect beliefs about the true average preference type of the group. This misperception is referred to as "pluralistic ignorance" in the social psychology literature, which is summarized below in Section 1.1. This literature notes that a primary source of pluralistic ignorance is a failure to recognize and adapt to changes to the groups' preferences. Motivated by this finding, I consider an extension where the underlying preference types of the population change over time and show that conformity concerns exacerbate the lag between the public beliefs and true preference types (Proposition 6).

Finally, the paper explores three further extensions that provide robustness checks for the main analysis. These extensions show that the qualitative features of the analysis do not depend on the following assumptions: short-run players (Proposition 7), linear decision rules (Proposition 8), or Gaussian random variables (Proposition 9).

As I elaborate below, this paper provides a new rationale for why social learning fails: conformity concerns. When conformity concerns are sufficiently small, the results from the earlier literature continue to hold and discrete actions or boundedly informative signals may prevent social learning. However, for higher (but, importantly, non-infinite) values of the conformity concerns social learning will fail due to the conformity concerns. In such settings, my model delivers predictions consistent with the empirical literature about effective interventions, whereas the social learning literature without conformity concerns would conjecture that individual-oriented interventions are always optimal.

1.1 Related Literature

This paper is related to three strands of literature: models of social learning, models of decision-making with reputational concerns, and the empirical literature about pluralistic ignorance and interventions. After presenting my results, I will connect my findings to the empirical literature in Section 4.3 and Section 5.

This paper is closely related to the literature on social learning. In this literature, social learning is the process in which players learn about their environment through their peers' decisions. Social learning is said to occur when asymptotically the players make optimal decisions. For instance, Banerjee (1992) and Bikhchandani et al. (1992) study models where players sequentially receive information and make a decision. If a sequence of past decisions is informative about the signals those players received, then players may rationally choose to stop utilizing their own signals, causing social learning to stop. In contrast, for the quadratic-loss environment I consider Lee (1993) shows that continuous decisions are a sufficient condition for social learning to occur.[4] Further, an implication of Kartik et al. (2024) is that with quadratic-loss payoffs, "directionally unbounded beliefs," which the Normal distribution satisfies, is sufficient for social learning.[5] I allow for both responsive decisions and directionally unbounded beliefs, and yet find that social learning can fail in the presence of conformity concerns. In contrast, the social learning literature has documented other obstacles to social learning such as costs of acquiring information (cf. Burguet and Vives, 2000; Chandrasekhar et al., 2018), misspecified priors (cf. Bohren, 2016; Frick et al., 2020), non-bayesian updating (cf. Golub and Jackson, 2010), changing fundamentals (cf. Dasaratha et al., 2023), or differential observability assumptions (cf. Banerjee and Fudenberg, 2004; Arieli and Mueller-Frank, 2019).

Further, this paper relates to the literature on reputational concerns. This literature typically analyzes settings where an agent takes an observable decision attempting to both (i) adapt the decision to a signal and (ii) make the observer think the agent is a "good type". In Scharfstein and Stein (1990), an agent wants to be perceived as competent, and in Morris (2001), an agent wants to be perceived as un-biased.[6] In both these papers, and most of the literature on reputational concerns, there is a single observer viewing the player's decision and the preferences of this observer are common knowledge.[7] In contrast, I focus

[4] More generally, Ali (2018) shows that if preferences are "responsive," defined as every distinct belief having a distinct optimal action, then social learning occurs.

[5] Relatedly, Smith and Sørensen (2000) show that if preferences are monotone and the state of the world is binary a non-zero probability of an arbitrarily precise signal is sufficient for efficient social learning to occur.

[6] Braghieri (2021); Angeli et al. (2022) provides empirical support for the prediction in Morris (2001) that political correctness may render speech uninformative.

[7] There are many papers which analyze different settings, but have the same feature: there is a known type the agent wants to be perceived as. See Canes-Wrone et al. (2001); Ely and Välimäki (2003); Ottaviani

on environments where the decision-maker has multiple observers and the ideal perceived type for the decision-maker is different for each potential observer, implying that the sender has single-peaked preferences over the receiver's beliefs. A subset of the reputation literature analyzes such preferences: Bernheim (1994), Loury (1994), Manski and Mayshar (2003), Austen-Smith and Fryer Jr (2005), Kuran and Sandholm (2008), Michaeli and Spiro (2015), and Tirole (2023). Further, I explicitly utilize the definition of conformity developed and modeled in Bernheim (1994). However, in all these papers, the interaction is static and the lone decision-maker maximizes over the distribution of observers. Instead, in my dynamic analysis there will exist aggregate uncertainty over the distribution of observers that will be partially resolved as the game unfolds. When this uncertainty is not fully resolved, the players conform to incorrect perceptions of their peers, a finding not present in the literature on reputational concerns.

There is a literature that considers the effect of reputation and coordination on social learning. For instance, Smith et al. (2021) shows contrarianism behavior should be rewarded for efficient informational herding. With conformity concerns, not only is contrarianism not rewarded, it is actively punished, resulting in inefficient learning. Additionally, Angeletos et al. (2007) considers an environment where players may receive benefits from coordinating on similar decisions. These are distinct from conformity concerns: a player's preferences over how his preference type is perceived.[8] This distinction results in different predictions for asymptotic learning and allows for the possibility of peer-oriented interventions.

The impact of conformity concerns on social learning is considered in Li and Van den Steen (2021) and Fernández-Duque (2022). In both models, players can either publicly support or oppose an issue. An individual's support for a particular decision is influenced by two factors: their personal preference type and their perception of social approval. These models are different from my work because in those papers (i) there is uncertainty about only the players' preference types, not the fundamental state of the world, and (ii) decisions and preference types are binary. Distinction (i) allows my work to discuss when an individual-oriented or a peer-oriented intervention is preferred. Distinction (ii) allows my work to predict a failure in social learning where previous work would not because the social learning literature without conformity concerns already predicts failures in social learning with discrete decisions. However, with conformity concerns, herding occurs with continuous decision-making. To see why, note that even a mild adaptation to one's private information may be met with an extreme change in reputation, preventing players from utilizing the continuous decision set.

Finally, this paper also contributes to a century of literature on "pluralistic ignorance," which refers to the systematic misperception of peers' preferences. For a review within economics see Bursztyn and Yang (2022) and within social psychology see Miller (2023).

and Sørensen (2006); Braghieri (2021); Rappoport (2022). Ely and Välimäki (2003) is especially relevant as Ely and Välimäki (2003) show that when players are sufficiently patient, the unique equilibrium is a pooling equilibrium. I show a similar result in Subsection 6.1, however the mechanisms are different.

[8] Bernheim (1994) provides a discussion of the difference between "popularity" concerns where each player's preferred decision is similar to his peer's decision and "conformity" concerns where each player's preferred decision is shaped by how his preference type is perceived. Bernheim (1994) shows popularity concerns yield vastly different equilibria than conformity concerns: for instance, popularity concerns will result in strictly monotone decision rules whereas conformity concerns result in pooling. Importantly, for popularity concerns there should be no difference between private and public decisions, whereas in Bursztyn et al. (2019, 2020b); Braghieri (2021), and many others we see differences.

This literature suggests that peer-oriented interventions (i.e., ones that correct pluralistic ignorance) are preferred in a variety of settings: substance abuse (cf. Schroeder and Prentice, 1998), female labor force participation (cf. Bursztyn et al., 2020a), and religious norms (cf. Gulesci et al., 2023). My paper provides a general framework that develops a prediction that is consistent with these applied literatures: pluralistic ignorance arises and peer-oriented interventions are needed when conformity concerns are high.

The rest of the paper is organized as follows. Section 2 describes the model, equilibrium assumptions, and defines relevant terms. Section 3 analyzes a benchmark of the model with common knowledge. Section 4 contains the main analysis. Section 5 extends the model to include interventions and additional determinants of pluralistic ignorance. Section 6 contains the three robustness extensions. Finally, Section 7 concludes and the Appendix contains proofs for all statements not shown in the text.

2 Model

Consider a community whereby, each period, a player makes a decision attempting to adapt to his private information and private preference type while possessing conformity concerns. The first subsection describes this set-up and the second subsection discusses the equilibrium refinements used throughout the analysis.

2.1 Model Description

Players: There is an infinite sequence of short-run players, $t \in 1, 2, \ldots$. Each player, t, observes the public history, h_t, (which will be specified after defining the utility), and his private information, and then chooses a decision $a_t \in \mathbf{R}$ in period t. The players possess uncertainty about both a common fundamental state, $\theta \sim N(0, \tau_\theta)$, and the average preference type of the group, $\mu \sim N(0, \tau_\mu)$.[9] Each player's private information includes a private signal $s_t = \theta + \epsilon_t$ and his private preference type $v_t = \mu + \nu_t$. I assume ϵ_t and ν_t are independent within and across periods, and that both are Gaussian random variables with a mean normalized to zero and a variance normalized to one.

Utility: Each player's utility has two components. First, the player wants to adapt his decision, a_t, to a combination of the fundamental state and his preference type. The weight of the fundamental state, $\gamma_t \geq 0$, is publicly observable and discussed further below. Second, while each player observes his own preference type, the player prefers the public's perception of his preference type to be close to the average within the community, which represents the conformity concerns.[10] Define by $\phi(b|h_t, a_t)$ the probability distribution over a player's preference type b that take decision a_t given the history of previous decisions.[11] Further, for

[9]The Gaussian assumption allows for clean comparative statics and to solve the model with uncertainty about both the preferences of others and the fundamental state. The benchmark in Section 3 where θ, μ are common knowledge holds for any distribution of v_t, as discussed in Section 6.2. Section 6.3 discusses generalizations of the learning framework to other distributions.

[10]I show in Proposition 10 in Appendix B that one would get qualitatively similar results if, throughout the analysis, each player wanted his perceived preferences to be c units higher than μ.

[11]There are two distinct reasons why previous decisions impact the inference function. First, previous decisions impact the equilibrium perception about θ, and thus a player may choose a higher decision due to a

now, $\phi(\cdot)$ is unconstrained off-path. The total utility for player t is thus,

$$u_t(a_t; v_t, s_t | h_t) := -\mathbf{E}_{\theta,\mu}\Big((a_t - \gamma_t\theta - v_t)^2 - \kappa \int (b-\mu)^2 \phi(b|h_t, a_t) db \big| v_t, s_t, h_t\Big). \qquad (1.1)$$

The first term in the expectation states that the player wants to choose a decision close to a linear combination of the fundamental state and his preference type. The second term is the conformity term, scaled by $\kappa \geq 0$. The term within the parenthesis is a reduced-form representation of conformity: player t wants the community's perception of his preference type, b, to be close to the true average preference type of the community, μ. Given that the loss function is quadratic, one can show that from the stand-point of which decisions are taken, this is equivalent to player t preferring that such an inference be close to that of a randomly drawn preference type in the community.[12]

Information: I assume that the public history takes the form $h_t = \{\gamma_1, a_1, \ldots, \gamma_{t-1}, a_{t-1}, \gamma_t\}$. This assumption states that there is full observability of the decisions and when they were made. Given this assumption, one can interpret the γ_t's as commonly observed time fixed-effects determining whether the fundamental state or one's preference type is comparatively more important. These time fixed effects are modeled as (i) $\gamma_t \leq 1\ \forall\ t$, and (ii) there being an $m \in \mathbf{N}$ such that $\gamma_{t+m} = \gamma_t$ and $\gamma_{t+1} \neq \gamma_t$. These assumptions are only necessary in the analysis with uncertainty about θ and μ. Without these assumptions there is only one "moment condition" for the players to separately infer θ and μ, potentially resulting in incomplete learning. However, in Section 9.4.4, I show that such learning outcomes are not locally stable, and this assumption removes their existence.[13]

I analyze Perfect Bayesian Equilibria (cf. Fudenberg and Tirole, 1991) satisfying the following requirements each period, and I will discuss each requirement below.

1. *Linearity:* Decisions are a linear combination of the public's beliefs about θ, the public's beliefs about μ, a player's private signal, and a player's preference type.

2. *Social Optimality:* The players always play the linear equilibrium that maximizes total surplus.

2.2 Equilibrium Selection:

Linearity: The restriction to linear equilibria is common when studying the normal learning model as it allows for greater tractability. Generally, this assumption is with loss of generality. Despite this, I show in Section 6.2 that when the distribution of preference types has full support and there is no uncertainty over θ (the fundamental state) or μ (the average preference type), the only equilibria with revelation which satisfy D1 from Cho and Kreps (1987) are

higher perception of θ rather than higher preferences. Second, previous decisions impact the equilibrium perception of the average preference type, which impact the equilibrium perception of a given player's preference type.

[12]Both interpretations correspond to Bernheim (1994) when μ is common knowledge.

[13]These learning outcomes resemble the confounded learning outcomes in Smith and Sørensen (2000). Unlike Smith and Sørensen (2000), these outcomes are not locally stable. Intuitively, what this assumption states is that players may place different weights on the fundamental state, e.g., at the end of the semester with respect to drinking or during election cycles in political speech examples.

the linear equilibria. Upon adding uncertainty over θ, through signals s_t, one must add distributional assumptions to ensure linearity is without loss. To see why, note that a_t will be a function of $v_t + \gamma_t \mathbf{E}(\theta|s_t, v_t)$, and thus the public inference of v_t will be based on the sum $v_t + \gamma_t \mathbf{E}(\theta|s_t, v_t)$. If this inference is non-linear, then generally the decision rule will be non-linear. In the Gaussian case, this inference is linear, and in Appendix B, I prove that the only equilibria with revelation that satisfy D1 remain linear. Similarly, upon adding uncertainty over μ, if player t's inference about μ is non-linear in his preference type, v_t, then generally the equilibrium will be non-linear. Again, the Gaussian assumption ensures that the inference is linear, and in Appendix B, I show that this assumption implies that the only equilibria satisfying D1 with revelation are linear.

Social Optimality: I assume that for each period, t, the equilibrium decision rule in period t maximizes the expected surplus out of all linear decision rules in period t.[14] This refinement is identical to one where the sequence of equilbria across periods maximizes the discounted expected surplus of the players from the class of linear equilibria in all periods. This refinement implies the players utilize a decision rule with revelation whenever one exists. Intuitively, players always prefer the linear equilibrium with revelation to pooling on any decision, a^*. In the revealing equilibrium, each type can always choose a^*, and that they do not implies they have a weak preference to not. Further, the conformity loss is independent of the equilibrium chosen by Bayes Plausibility.

These criterion prescribe a unique decision rule in all periods, t. Thus, I refer to the unique Perfect Bayesian Equilibrium satisfying these conditions as *the signaling equilibrium*.

3 Common Knowledge Benchmark

This section analyzes the impact of conformity on decision-making and mutes uncertainty about the fundamental state and the preferences of others. To do so, I assume θ and μ are common knowledge, and, without loss of generality, are both equal to zero.[15] Further, without uncertainty, there is no time dependence, thus it is without loss of generality to consider the decision rule of a player with preference type v. A signaling equilibrium requires that this decision rule is linear in v. Given linearity, there are two cases: a constant and a strictly increasing decision rule. If the decision rule is constant, then this decision rule is defined as “fully-pooling." Further, for any fully-pooling decision rule, there exist off-path beliefs that deter any deviations from the pooling decision.

Any decision rule, $a(v)$ that is non-constant is defined as “revealing" and is determined by $\hat{v}(a) = \alpha a + \beta$, which was defined as the posterior expectation of a player's preference type given his decision. The necessary and sufficient condition for these beliefs to constitute an equilibrium is that given $\hat{v}(a)$, the decision rule that maximizes a players utility, $a(v)$, must result in a consistent conjecture of $\hat{v}(a)$. The first-order condition for the decision rule given

[14]Without the social optimality refinement, then for any sequence of values $x_1, \ldots,$ there exists an equilibrium where all players pool on x_t in period t. Such equilibria are removed by this requirement.

[15]As a result of θ being common knowledge, the players disregard their signals s_t, which removes any need for distributional assumptions about the signals. This simplified analysis is similar to that of Bernheim (1994). Section 6.2 examines this environment without assuming linearity and offers a comprehensive comparison with the findings of Bernheim (1994).

a conjecture $\hat{v}(a) = \alpha a + \beta$ is:

$$a - v + \kappa\alpha(\alpha a + \beta) = 0 \iff v = (1 + \kappa\alpha^2)a + \kappa\alpha\beta. \tag{1.2}$$

Further, these beliefs constitute an equilibrium if and only if:

$$1 + \kappa\alpha^2 = \alpha \text{ and } \beta = \kappa\alpha\beta. \tag{1.3}$$

It is immediate that $\beta = 0$ is a solution to the latter equality, and, further, it is the unique solution for any α that solves the former. While one can solve the former equality, Figure 1.1 provides intuition why an equilibrium with revelation cannot exist for high values of conformity concerns. Figure 1.1 depicts as a function of the conjectured slope of the beliefs, α, the resulting beliefs from the best response, $1 + \kappa\alpha^2$, for two different values of κ. Further, note that Figure 1.1 also provides intuition for why the fully-pooling decision rule is an equilibrium. In the pooling equilibrium, the slope of the decision rule is zero, implying that the slope of the beliefs (i.e., the inverse of the decision rule) is infinity. Finally, the best response to such beliefs is to choose the pooling decision, which will generate such beliefs in equilibrium.

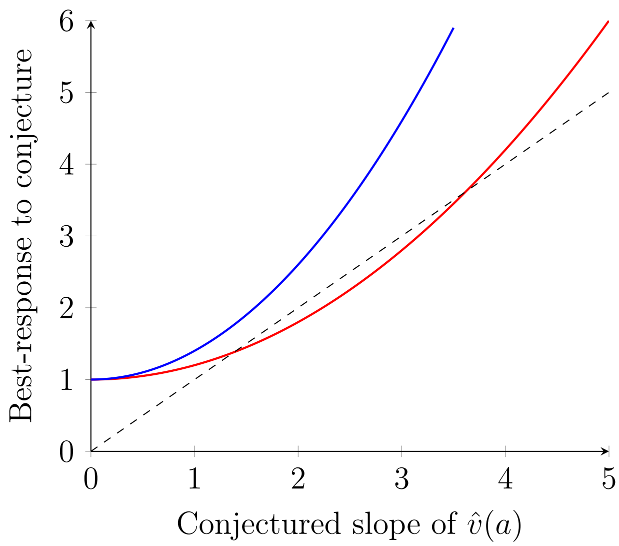


Figure 1.1: Existence of Linear Equilibria with Revelation

The x-axis represents a conjectured slope of the posterior beliefs of the community given a decision, $\hat{v}(a)$. The y-axis depicts the beliefs that result from the best response of the players to such a conjecture as stated in Equation (1.3). In Blue is the best response when the conformity concerns are high ($\kappa = .4$) and in Red the best response when the conformity concerns are low ($\kappa = .2$).

Figure 1.1 provides complementary reasons why no equilibrium with revelation exists for high values of κ. Intuitively, (i) as the conformity concerns increase, the players must conform more resulting in a lesser slope of $a(v)$. As $\hat{v}(a)$ is the inverse of $a(v)$, then the slope of $\hat{v}(a)$ is high. Further, (ii) the players cannot conform too much while still maintaining an equilibrium with revelation. If this occurred, then the slope of $a(v)$ must be small, implying a large slope of $\hat{v}(a)$. If the conjectured slope is too large, the players have a large incentive to conform as their decision greatly impacts their reputation. Combining these intuitions, if the conformity concerns are high then (i) dictates that the slope of $\hat{v}(a)$ must be large, but (ii) dictates that such slopes cannot constitute an equilibrium.

Finally, when the conformity concerns are low enough that an equilibrium with revelation exists, then three different linear Perfect Bayesian Equilibria exist. As the conformity loss is fixed across the equilibria and equal to κ multiplied by the variance of v, we can focus on the adaptation loss. Further, a lesser slope of $\hat{v}(a)$ corresponds to a greater slope of $a(v)$, implying that the equilibrium with the lowest slope of $\hat{v}(a)$ results in the best adaptation loss. Therefore, the social optimality refinement pins down a unique signaling equilibrium.

Proposition 1 (**Commonly Known Environment**)
There exists a threshold value, $\kappa^{c.k.}$, such that the unique signaling equilibrium as defined in Section 2.2 is characterized by the following decision rule:

$$a(v) = \begin{cases} \frac{1+\sqrt{1-4\kappa}}{2} v & \text{if } \kappa \leq \kappa^{c.k.} \\ 0 & \text{if } \kappa > \kappa^{c.k.} \end{cases}, \tag{1.4}$$

where κ denotes the weight on conformity. Given this decision rule,

$$u(v) = \begin{cases} -\frac{1-\sqrt{1-4\kappa}}{2} v^2 & \text{if } \kappa \leq \kappa^{c.k.} \\ -v^2 - \kappa & \text{if } \kappa > \kappa^{c.k.} \end{cases}. \tag{1.5}$$

This proposition summarizes the intuitions from the figure. First, the degree to which a player adapts to his preference type is decreasing with respect to the conformity concerns, κ. To see this, note that when $\kappa = 0$, $a(v) = v$ and as κ increases up to the threshold $\kappa^{c.k.}$, $a(v) = v/2$. The reason for this decrease is that each player has an added incentive to conform when the conformity concerns increase. Finally, when the incentive to conform becomes sufficiently high, i.e., $\kappa > \kappa^{c.k.}$, there is no decision rule with revelation and the signaling equilibrium is fully pooling. Importantly, such fully-pooling decision rules provide no information about a player's private information, and this observation will be key in the main analysis.

4 Analysis

This section begins with an analysis of the complete model where the players learn both the fundamental state and the average preference type of their peers while possessing conformity concerns. I show that the players learn the fundamental state if and only if they learn the average preference type of their peers. Further, such learning fails whenever the conformity concerns are sufficiently high and occurs whenever the conformity concerns are sufficiently low. The second subsection presents additional comparative statics about the asymptotic utility and the asymptotic precision of the beliefs about the fundamental state and show that both of these terms are decreasing in κ. In doing so, I assume away uncertainty about the population's average preference type and analyzes how the players learn the fundamental state.[16] Incorporating both dimensions of uncertainty allows for predictions that neither uni-dimensional learning model will produce. However, doing so complicates the analysis by requiring a joint update in the posterior beliefs each period. In the final subsection, I relate the findings of my baseline theoretical model to the empirical and qualitative literatures discussed in Section 1.1.

Before delving into the analysis, I define notation that appears throughout. Given the Gaussian set-up, the joint distribution of the public beliefs about θ and μ at any time t remain jointly Gaussian.[17] Define $\theta(t) = \mathbf{E}(\theta|h_t)$ and $\mu(t) = \mathbf{E}(\mu|h_t)$.[18] These random variables have the following joint distribution:

$$\begin{pmatrix} \theta(t) \\ \mu(t) \end{pmatrix} \sim N\left(\begin{pmatrix} \bar{\theta}(t) \\ \bar{\mu}(t) \end{pmatrix}, \begin{pmatrix} \tau_{\theta,t} & \rho_t\sqrt{\tau_{\theta,t}\tau_{\mu,t}} \\ \rho_t\sqrt{\tau_{\theta,t}\tau_{\mu,t}} & \tau_{\mu,t} \end{pmatrix}^{-1} \right). \tag{1.6}$$

As is common, it will be convenient to work with the precision matrix, defined as the inverse of the variance matrix in the above equation. Below, I introduce definitions.

Definition 1 (**Social Learning**)
Social learning about fundamentals occurs (respectively, fails) if and only if $\theta(t) \to_p \theta$ (respectively, $\theta(t) \not\to_p \theta$). Social learning about preferences occurs (respectively, fails) if and only if $\mu(t) \to_p \mu$ (respectively, $\mu(t) \not\to_p \mu$).

In the signaling equilibrium, the inference function, $\phi(b, h_t, a_t)$, will be Gaussian, the mean of which will be denoted as $\hat{v}(a_t) = \int b \cdot \phi(b, h_t, a_t)$. Within the quadratic set-up the variance of $\phi(b, h_t, a_t)$ and the variance of μ are independent of the decision taken. These observations allow for the following simplification of the utility function up to some constant c_t as stated below.

$$\begin{aligned} u_t(a_t; v_t, s_t|\theta(t), \mu(t), \gamma_t) :=& -\mathbf{E}_{\theta,\mu}\Big((a_t - \gamma_t\theta - v_t)^2\Big|\theta(t), \mu(t), v_t, s_t\Big) \\ &- \kappa\mathbf{E}_{\theta,\mu}\Big(\big(\hat{v}_t(a_t) - \mathbf{E}(\mu)\big)^2\Big|\theta(t), \mu(t), v_t, s_t\Big) - c_t. \end{aligned} \tag{1.7}$$

The first term is the adaptation loss. The second term is the squared difference from the expectation of player t's preference type and the average preference type of the population, and the final term is the variance stemming from the uncertainty over μ and θ. Given this utility, I define the following.

Definition 2 (**Asymptotic Utility and Adaptation Loss**)
The asymptotic utility is $\lim_{t\to\infty} \mathbf{E}\big(u_t(a_t; v_t, s_t|\theta(t), \mu(t), \gamma_t)\big)$. The asymptotic adaptation loss is $\lim_{t\to\infty} \mathbf{E}\big(-(a_t - \gamma_t\theta - v_t)^2\big)$.

[16] Further, in Subsection 5.5, I conduct an analysis where the fundamental state is common knowledge, but the average preference type, μ, is uncertain. In that analysis, conformity concerns prevent the players from learning μ, and provides similar comparative statics.

[17] Notably, while the prior beliefs about θ and μ in period 1 are independent, in any subsequent period, the beliefs about θ and μ are dependent as both condition on the same sets of decisions.

[18] Throughout, $\theta(t), \mu(t)$ are sufficient statistics for the probability distribution determining the players' beliefs at a given point in time. As a shorthand, I therefore refer to these random variables as "the beliefs."

Throughout I provide results discussing both notions of asymptotic efficiency. However, the comparative statics for one notion of asymptotic efficiency coincide with the comparative statics for the other because the difference in these terms is the expected conformity loss which is pinned down by Bayes plausibility.

4.1 Determinants of Social Learning

This subsection analyzes the complete model where the players attempt to socially learn θ and μ as described in Section 2.1. To analyze this environment, I consider an arbitrary period, t. Recall that the beliefs about θ and μ follow a bivariate normal distribution. These beliefs are sufficient statistics for the sequence of past decisions, $a_1, \ldots, a_{t-1}$, and the relative weights on the fundamental state in each period $\gamma_1, \ldots, \gamma_{t-1}$ in a linear equilibrium.

To analyze the signaling equilibrium, fix any equilibrium that is not fully pooling and a conjecture for $\hat{v}_t(a_t) = \alpha_t a_t + \beta_t$, where the consistent conjectures for α_t and β_t will, of course, depend on the beliefs at period t. In this equilibrium, player t's first-order condition for his decision rule is:

$$a_t(1 + \kappa\alpha_t^2) = \gamma_t \mathbf{E}\big(\theta \mid \theta(t), \mu(t), s_t, v_t\big) + v_t + \kappa\alpha_t\beta_t + \kappa\alpha_t \mathbf{E}\big(\mu \mid \theta(t), \mu(t), s_t, v_t\big). \tag{1.8}$$

Such an equilibrium exists if the posterior expectation of v_t given such a decision rule is consistent with the equilibrium conjecture of $\hat{v}(a_t) = \alpha_t a_t + \beta_t$. Given the distributional assumptions, one can write this posterior expectation as follows,

$$\mathbf{E}\Big(v_t \mid a_t(1 + \kappa\alpha_t^2)\Big) := a_t(1 + \kappa\alpha_t^2) \cdot \tilde{s}(\alpha_t, \kappa, \theta(t), \mu(t)) + \iota(\alpha_t, \kappa, \beta_t, \theta(t), \mu(t)), \tag{1.9}$$

where $\tilde{s}(\cdot)$ will be thought of as determining the sensitivity of the decision rule to v_t and $\iota(\cdot)$ as determining the intercept. Thus the necessary and sufficient condition for a signaling equilibrium to involve revelation is whether there exist an α_t and β_t which satisfy,

$$(1 + \kappa\alpha_t^2)\tilde{s}(\alpha_t, \kappa, \theta(t), \mu(t)) = \alpha_t \tag{1.10}$$

$$\iota(\alpha_t, \kappa, \beta_t, \theta(t), \mu(t)) = \beta_t. \tag{1.11}$$

These equations resemble those in Equation (1.3) where, in that benchmark, $\tilde{s}(\cdot) = 1$ and $\iota(\cdot) = \kappa\alpha_t\beta_t$. Similar to that benchmark, whenever there exists a solution to Equation (1.10), there will exist a unique solution to Equation (1.11), thus shifting the focus to Equation (1.10). Further, Equation (1.10) is independent of β_t. Finally, the sensitivity, $\tilde{s}(\cdot)$, is independent of the means of the beliefs, $\theta(t), \mu(t)$. This independence arises because the sensitivity captures how variation in the decision corresponds to variation in a player's preference type, which is independent of the mean beliefs in a linear equilibrium. The following lemma formalizes this intuition and provides additional properties of the learning process.

Lemma 1 *Fix κ, $\{\gamma_t\}$, and initial beliefs about θ and μ. Social learning about fundamentals occurs if and only if social learning about preferences occurs. Further, whether or not social learning about fundamentals occurs (symmetrically, preferences) is independent of the realizations of a_t.*

The intuition behind the first statement in the lemma is that if the players socially learn θ, the players must have observed infinitely many periods of informative decisions. Given the knowledge of what θ is, the players can use these infinitely many periods to infer μ and vise versa. Since the conditions for social learning about preferences and fundamentals are identical, for brevity, I will refer to social learning as when the players socially learn both the preferences and fundamentals. Further, the second statement formalizes the intuition that whether an equilibrium involves revelation is determined only by the sensitivity in the conjectured type as a function of the decision. Importantly, that the sensitivity is influenced by the precision of the beliefs in a given period is exactly why the *effective conformity concerns*, corresponding to the endogenous reputational penalty of adaptation, change over time. Adaptation imposes a change in one's perceived preference type determined by the sensitivity, and the conformity loss is equal to κ multiplied by this sensitivity. To build intuition, if the beliefs about θ are sufficiently imprecise, each player will put comparatively more weight on his signal s_t. As the decision rule now puts a larger weight on s_t than v_t, the sensitivity will be lower because variation in a_t will be ascribed to a player's signal as opposed to his preference type. In contrast, if the players are sufficiently certain about θ and μ, then each player effectively disregards his signal and his preference type when computing the posterior expectations of θ and μ, implying that the right-hand side of Equation (1.8) is approximately equal to v_t and that the sensitivity is approximately equal to 1. One can use this intuition to generate the following lemma, providing a sufficient condition for a failure in asymptotic learning.

Lemma 2 (**Sufficient Condition for Failure of Social Learning**)
If $\kappa > \kappa^{c.k.}$, social learning about fundamentals and preferences fails for any initial beliefs about the fundamental state and average preference type.

This lemma states that when the conformity concerns exceed the threshold for revelation in the common knowledge environment, the players are unable to socially learn θ or μ. To gain intuition, note that by Lemma 1 an equilibrium with revelation exists if and only if there exists an α which solves Equation (1.10). That $\kappa > \kappa^{c.k.} = 1/4$ implies that if the sensitivity were equal to one (or within an ϵ window of 1), then there would exist no solution, as the left-hand side of Equation (1.10) would be strictly greater than the right-hand side for any α. Hence, for a signaling equilibrium with revelation to exist, the sensitivity cannot converge to one. Further, for social learning to occur, there must exist infinitely many periods of revelation even as the beliefs converge to the truth. However, if the beliefs converge to the truth, the right-hand side of Equation (1.8) converges to v_t, implying that the sensitivity converges to 1, yielding a contradiction. As a result, when $\kappa > \kappa^{\text{c.k.}}$ the beliefs do not converge and the players face three inefficiencies in the limit. First, as the players use a pooling decision rule in the limit, the players are unable to adapt to their preference types. The subsequent two inefficiencies stem from the players utilizing a pooling decision rule based on inaccurate perceptions of *both* θ and μ.

Given Lemma 2, it suffices to analyze $\kappa < \kappa^{\text{c.k.}}$. First note that for any $\kappa < \kappa^{\text{c.k.}}$, there exist initial beliefs such that social learning succeeds. To see why, note that for sufficiently precise beliefs about θ and μ, the noise stemming from adaptation to changes in the perceptions of θ or μ is sufficiently small. As a result, for any ϵ, there exist sufficiently precise beliefs about θ and μ, such that the sensitivity in the decision to v_t is less than $1 + \epsilon$. By Equation (1.10), an equilibrium with revelation exists if and only if $1 + \kappa\alpha_t^2$ multiplied by the sensitivity is equal to α_t. Since $\kappa < \kappa^{\text{c.k.}}$, one can increase the left-hand side by ϵ (corresponding to an increase in κ) and there will still exist a solution. Hence, for sufficiently precise beliefs, there will exist an equilibrium with revelation. Further, as the beliefs in the subsequent period will be more precise, the next period will involve revelation, and by induction, all subsequent periods. Finally, the assumed time fixed-effects in γ_t imply that the players can separately identify θ and μ.[19] This intuition implies that if $\kappa < \kappa^{\text{c.k.}}$, there exist sufficiently precise beliefs for which social learning will succeed. The following lemma formalizes this intuition.

Lemma 3 (**Existence of Social Learning**)
Fix $\{\gamma_t\}$ and $\kappa < \kappa^{c.k.}$. There exists an open set of initial beliefs such that social learning about preferences and fundamentals occurs.

The question now turns to whether $\kappa < \kappa^{\text{c.k.}}$ is a sufficient condition for social learning for all possible initial beliefs. Intuitively, for social learning to fail, there must exist a period in which the signaling equilibrium is pooling. That the signaling equilibrium is pooling is equivalent to there not being a solution to $1+\kappa\alpha_t^2$ multiplied by the sensitivity in the decision rule to v_t equaling α_t (at a high level, the effective conformity concerns being larger than the conformity concerns). As one example of why this may occur, note that given a sequence of high decisions each player is unsure if the decisions were high due to a high fundamental state or high average preference type. If the player has a low preference type, he updates that the fundamental state must be comparatively high. If this inference is sufficiently strong, the sensitivity in the posterior expectation of v_t given a change in the decision may be greater than one. When this sensitivity is strictly greater than one, the condition for the existence of an equilibrium with revelation is strictly tighter than $\kappa < \kappa^{\text{c.k.}}$. As a result, there may exist an open set of initial beliefs such that social learning about both preferences and fundamentals fails, despite $\kappa < \kappa^{\text{c.k.}}$. The following proposition formalizes this logic and unifies the previous lemmas.

Proposition 2 (**Characterization of Long-Run Learning**)
There exists a threshold $\underline{\kappa} \in (0, \kappa^{c.k.})$ such that:

1. *If $\kappa \leq \underline{\kappa}$, then for any initial beliefs, social learning about preferences and fundamentals occurs.*

2. *If $\kappa \in (\underline{\kappa}, \kappa^{c.k.})$, then there is an open set of initial beliefs such that social learning about preferences and fundamentals occurs. Further, there exists an open set of parameter values for which social learning about preferences and fundamentals fails despite $\kappa \in (\underline{\kappa}, \kappa^{c.k.})$.*

3. *If $\kappa > \kappa^{c.k.}$, then for any initial beliefs, social learning about preferences and fundamentals fails.*

The first statement can be seen by analyzing the condition for the existence of a signaling equilibrium with revelation. If κ is sufficiently small, then for any prior beliefs one can show

[19] As discussed in Footnote 13, without this assumption the beliefs may converge to an unstable learning outcome.

that the conformity concerns are sufficiently small such that there exists an equilibrium with revelation in every period. Further, the third result is a direct consequence of Lemma 2. The second result states that when the conformity concerns take an intermediate value social learning may occur or fail. That social learning may occur is a consequence of Lemma 3. That social learning may fail is due to the effective conformity concerns being larger than the conformity concerns when there is a sufficiently strong negative correlation between the beliefs about θ and μ.

This section shows that the condition for social learning-about either fundamentals or the preference types-depends on an intuitive fundamental: the magnitude of conformity concerns.[20] I discuss this finding in the context of the empirical literature in Section 4.3.

4.2 Comparative Statics

In the previous subsection the players possess uncertainty about both a fundamental state and the average preference type of others. Allowing for both types of uncertainties allows for the comparison between interventions addressing misperceptions of the preferences of others and misperceptions about the fundamental state in Section 5. However, the finding that social learning about fundamentals occurs if and only if conformity concerns are sufficiently small continues to hold absent uncertainty about the preferences in the population. Further, the analysis absent such uncertainty allows for additional comparative statics about the long-run behavior of the players.

In this subsection, I assume, μ, the average preference type of the community is common knowledge and without loss of generality equal to zero. Further, as the learning problem has only one dimension of uncertainty, one can normalize $\gamma_t = 1$ without loss of insights (see Footnote 13). As such, a sufficient statistic for the history is the current public belief about θ, $\theta(t) \sim N(\bar{\theta}(t), \tau_{\theta,t})$.

The evolution of $\theta(t)$ uniquely determines the equilibrium behavior of the players. Further, as argued in Section 4.1, the mean of $\theta(t)$, $\bar{\theta}(t)$, does not influence whether or not an equilibrium will be revealing nor the degree of revelation. Thus, the equilibrium dynamics are determined by the precision of $\theta(t)$, $\tau_{\theta,t}$. In this simplified model, the effect of greater uncertainty about θ is that deviations in the decision are increasingly ascribed to s_t as opposed to v_t. As a result, the greater the uncertainty, the lesser the effective conformity concerns. These intuitions combine to generate the following proposition.

Proposition 3 (**Social Learning when Average Preference Type is Known**)
Fix any prior beliefs about θ and let μ be common knowledge. Social learning about fundamentals occurs if and only if $\kappa \leq \kappa^{c.k.}$. Further, when $\kappa > \kappa^{c.k.}$ the long-run precision of the beliefs $\tau_\theta(\kappa) := \lim \tau_{\theta,t}(k) < \infty$ is decreasing in κ. Finally, the asymptotic adaptation loss and asymptotic utility of the players is decreasing in κ with a discontinuity at $\kappa = \kappa^{c.k.}$.

This proposition shows that conformity concerns impacts not only the binary outcome of social learning, but also the degree of asymptotic learning. The intuition behind this result

[20]At this level of generality, the learning outcomes are not monotone in κ. For instance, a marginally higher value of κ alters whether the decision rule places marginally more weight on s_t or v_t, which results in differential beliefs in the subsequent period. In either analysis with only one dimension of uncertainty, the learning outcomes are strictly monotone with respect to κ.

is that uncertainty about θ, in this simplified environment, always decreases the effective conformity concerns. As a result, if the equilibrium involves revelation in the environment with common knowledge, then the equilibrium is revealing in all periods with uncertainty about θ. This implies that if $\kappa \leq \kappa^{c.k.}$, the players learn θ for any initial beliefs. In contrast, if $\kappa > \kappa^{c.k.}$, then in the common knowledge environment the players cannot adapt to their private information. As social learning necessitates adaptation even as the beliefs become arbitrarily precise, the players necessarily stop adapting to their private information before learning occurs.

Further, when the conformity concerns are higher, the players switch to the pooling equilibrium earlier. Given this earlier switch, an increase in κ when $\kappa > \kappa^{\text{c.k.}}$ results in pooling on less accurate perceptions of the fundamental state, ultimately resulting in both a worse asymptotic adaptation loss and a worse asymptotic utility. In contrast, for $\kappa \leq \kappa^{c.k.}$, the asymptotic utility of the players converges to that in the common knowledge benchmark. Finally, the discontinuity at $\kappa = \kappa^{\text{c.k.}}$ occurs because the players can adapt to their private type in the limit if and only if $\kappa \leq \kappa^{\text{c.k.}}$. The following figure showcases this intuition, by plotting the asymptotic adaptation loss as a function of κ in both the benchmark (blue) and the dynamic model (red).

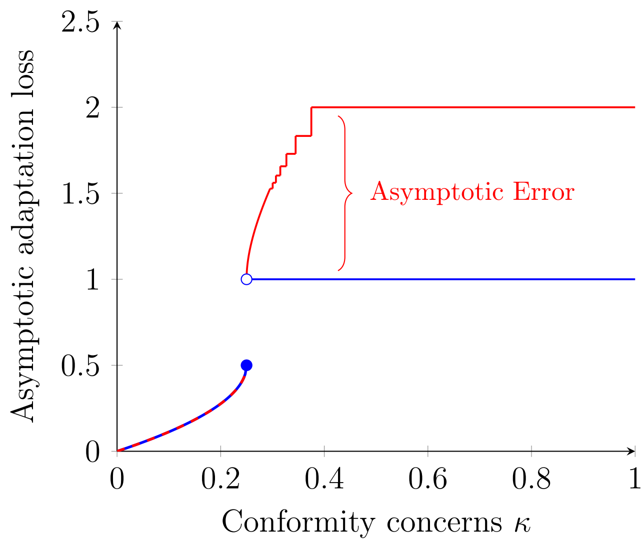


Figure 1.2: Asymptotic Adaptation Loss

In Blue is the asymptotic adaptation loss in the benchmark. In Red is the asymptotic adaptation loss when the players have common knowledge about μ, the average preference type of their peers, but have prior $\theta \sim N(0,1)$ about the fundamental state. These two functions coincide when $\kappa \leq \kappa^{\text{c.k.}}$ but differ for higher values of conformity concerns. The difference between these two functions is due to the players pooling on $\bar{\theta}(t)$, which will not equal θ, when the players fail to socially learn θ.

This simplified model showcases how conformity concerns impede efficient learning. Despite the players having access to continuous decisions and sufficiently informative signals, the players fail to learn the fundamental state for any prior beliefs when the conformity concerns are sufficiently high. Further, in addition to the extensive margin of whether the beliefs perfectly converge to the truth, the conformity concerns effect the intensive margin of the precision of the beliefs: the greater the conformity concerns, the more imprecise beliefs the players ultimate harbor.

4.3 Applied Relevance

My model has predictions for both the asymptotic efficiency of decisions and whether the players will learn the average preference type of their peers. I now argue that my predictions are more consistent with the empirical literature than the existing theoretical literature discussed in Section 1.1.

Determinants of Efficient Decisions: In my analysis, the magnitude of conformity concerns are the main predictor for whether decisions will be asymptotically efficient. In contrast, the social learning literature predicts that continuous decisions or unboundedly informative signals are sufficient for asymptotic efficiency. I now review the empirical literature in support of my predictions.

Continuous Decisions: As discussed in Lee (1993), examples of successful social learning include financial markets and business forecasting where investment decisions are continuous. The literature notes that despite a potential for initial erroneous mistakes and herding, that firms asymptotically learn whether a given asset is valuable. In contrast, my model predicts that the conformity concerns must be low for efficient decisions asymptotically. Returning to the case study of financial markets, one might think the conformity concerns are low relative to the financial stakes. Given these low conformity concerns, my model produces a similar prediction to the historical literature. However, if decisions are continuous, but the conformity concerns are high, such as alcohol consumption (cf. Prentice and Miller, 1993), drug use (cf. West and O'Neal, 2004), and many others, then in line with my model, we see more inefficiencies and worse beliefs amongst the community.

Unboundedly Informative Signals: The social learning literature defines a signal as unboundedly informative signals if with positive probability the player is arbitrarily certain of the optimal decision after observing one's private signal. If these signals occur, then even if the players are herding on a wrong decision, when such a signal occurs, the player who received such a signal will break the herd and choose the correct decision. This observation is in contrast to the famous conformity experiment in Asch (1953). Participants were grouped and shown a series of lines, then asked to identify the one matching a reference line. Unbeknownst to the participants, the experimenters planted an actor into the group to deliberately provide incorrect answers. Without the actors, success exceeded ninety-nine percent, but with the actors, over seventy-five percent of participants conformed. The social learning literature predicts that individuals should not copy the actor because each individual can identify the correct answer. In contrast, my model predicts that if conformity concerns are large, the individuals will copy the actor's incorrect decision.[21] Finally, Franzen and Mader (2023) replicated the original study and found that monetary incentives decrease the probability of conformity by 13 percentage points. Consistent with my model, these incentives increase the

[21] In my model, decisions are continuous. However, one could alter my model to consider discrete decisions, and indeed one would find that players would conform even in the presence of unboundedly informative signals.

importance of adaptation, resulting in less conformity.

Pluralistic Ignorance: My model predicts that "pluralistic ignorance" can arise in equilibrium. Pluralistic ignorance is defined as, "a situation in which group members systematically misestimate their peers' attitudes" (Miller, 2023). In a review article, Bursztyn and Yang (2022) document that such misperceptions lead to inefficient social norms and are rampant, occurring in a wide range of environments: political movements, macroeconomic expectations, vaccination status, and many others.

The extent of pluralistic ignorance corresponds to the magnitude by which individuals systematically misestimate the preference types of their peers, for a given realization of their peers true preference types.[22] In the model, if the public beliefs about μ converge to the truth, then (tautologically) every player correctly predicts the average preference type. In contrast, if the public beliefs do not converge, then there exists uncertainty about μ, implying that each player's estimate of μ combines both the public beliefs about μ and his preference type, v_t, which is predictive about μ. In such cases, the distribution of estimates will be non-degenerate, and with a probability equal to one, will not be perfectly centered around μ. Further, the greater the uncertainty in the public beliefs, the more likely such beliefs are centered around the prior beliefs about μ as opposed to the true value. Thus, one can view the uncertainty about the public beliefs about μ (the inverse of $\tau_{\mu,t}$) as describing the expected degree of pluralistic ignorance.

While there exist numerous behavioral explanations for pluralistic ignorance, the model presented in this paper provides an additional explanation: the desire for conformity necessitates "self-censorship in public discourse" (cf. Loury, 1994), resulting in insufficient information for others to gauge the views of the public.[23]

5 Policy Interventions

In this section, I first review two case studies about interventions. Next, I extend the model to analyze informational interventions and show that when conformity concerns are high, interventions addressing misperceptions of the average preference type outperform interventions addressing misperceptions of the fundamental state. Finally, motivated by the literature in social psychology on pluralistic ignorance, I extend the model to consider changing preference types and connect this finding to peer-oriented interventions.

[22]In the model, $\mu(t)$ is an unbiased estimator for μ. However, the object of interest is the gap between $\mu(t)$ and μ for a given realization of the preferences, μ. One may object that, in practice, $\mu(t)$ is greater than μ in every school in the context of alcohol (rather than being unbiased). However, all these students observe similar sets of celebrities on social media or television, implying that the beliefs across schools should not be viewed as independent samples.

[23]This theory finds empirical backing in Braghieri (2021) which documents that participants are likely to skew their answers to politically sensitive questions in the direction of public support when these answers are public and that such skewing decreases the information from any given statement. In contrast to Braghieri (2021), I allow for uncertainty about the direction of the public support. For instance, college students may be uncertain how "cool" drinking is. Allowing for such uncertainty allows for a tighter connection with the case studies regarding pluralistic ignorance.

5.1 Empirical Examples

In this subsection I detail two empirical studies about interventions from social psychology and economics, respectively.

Case 1: Alcohol Use on Campus: Prentice and Miller (1993) conducted a survey amongst Princeton undergraduates to show that students over-estimate their peers' preferences towards alcohol by 32 percent and that such misperceptions are correlated to the over-consumption of alcohol on campus.[24] Given the results in Prentice and Miller (1993), Schroeder and Prentice (1998) causally tested whether these misperceptions are the primary cause of excessive drinking on campus as opposed to potential misperceptions of the deleterious health consequences of excess drinking. To do so, the authors divided incoming college students into two groups. The first group received information about the health effects of alcohol consumption (θ in my model), and the second group received information about their peers' preferences towards alcohol (μ in my model). At the end of the semester, students were surveyed about their drinking behavior, and the intervention targeting misperceptions of peers' preferences reported 40 percent less drinking than the intervention addressing health consequences.

Case 2: Women Working Outside the Home: A recent randomized experiment from Bursztyn et al. (2020a) finds that, in the context of Saudi Arabia, the vast majority of young married men privately support women working outside the home (WWOH) and substantially underestimate support by other similar men. Further, the low levels of WWOH, suggests that such misperceptions may be impeding efficient behavior. The authors then randomly assign information about the correct perceptions of one's peers to the participants and show that correcting the misperceptions increases men's willingness to have their wives work by 36 percent. In a final step towards showing that misperceptions of peers' preferences are the main inhibitor of efficient decision-making, Bursztyn et al. (2020a) consider whether misperceptions about the fundamental state may be the root cause, noting, "if so many people, in fact, support WWOH then there are probably many firms willing to hire women for jobs outside the home" (p. 3018). They show that information about preferences of their peers caused no updated inference in the number of jobs for women, consistent with the fact that the effect is not caused through updated beliefs about the fundamental state.[25]

[24] Prentice and Miller (1993) even notes, "Princeton reunions boast the second highest level of alcohol consumption for any event in the country after the Indianapolis 500," implying that such perceptions are sufficiently strong to continue a decade after graduation at a reunion and have deleterious outcomes for the community.

[25] Bursztyn et al. (2020a) include a model with a few key differences to mine. First, in Bursztyn et al. (2020a) agents are endowed with incorrect beliefs about their peers' preference types. In contrast, these misperceptions arise in equilibrium in my model. Second, in Bursztyn et al. (2020a) decisions are discrete, whereas my paper shows learning may fail for continuous decisions. Finally, my model allows for misperceptions about θ and μ and gives conditions when information about μ is preferred. In contrast, in Bursztyn et al. (2020a) θ (which can be viewed as the economic benefits of WWOH) is common knowledge and the only possible interventions are about μ.

5.2 Modeling Interventions

I consider four different types of interventions composed of the intersection of whether the information shared with individuals is made common knowledge and whether the information is about the fundamental state or the average preference type.

Before analyzing "common-knowledge" interventions, I analyze "private interventions." One can think of a private intervention as giving player t access to additional information; however, such information is private and is not accounted for by the community when inferring player t's preference type given his decision. Without formally stating the result, one can see that such an intervention has no ability to break a pooling equilibrium nor influence which decision the players pool on. To see why, suppose player t is told the value of the fundamental state, θ. Given that each player wants to match his decision to the fundamental state, player t has an identical incentive to adapt to θ as a hypothetical player who received a signal whose implied posterior mean of θ matches the fundamental state. Further, since the signal distribution has full support, and the hypothetical player cannot adapt to such information, neither can player t. Thus, the equilibrium in period t remains identical. Finally, as such information was private information to player t, and no change in behavior occurs in period t, then no change in behavior follows for any subsequent periods.[26]

Given the stark irrelevance result for private interventions, I now focus on common knowledge interventions. In the standard framework absent interventions, the public history at time t is $\mathbf{h}_t = \{\gamma_1, a_1, \ldots, \gamma_{t-1}, a_{t-1}, \gamma_t\}$, namely the sequence of past decisions and the environments in which such decisions were chosen. I consider an intervention where information is released before period t, but after a_{t-1}. Such an intervention leaves the prior histories unchanged (and further the prior sequence of events remains unchanged as each player is short-lived). This information could be about either θ or μ, which will be referred to as individual-oriented and peer-oriented interventions, respectively.

5.3 The Effects of Interventions

I begin with a definition of when a signaling equilibrium with pooling is fragile. I call a pooling decision rule "fragile" to an individual-oriented intervention with n pieces of information if there exists a hypothetical public disclosure of n i.i.d. signals about θ that are distributed identically to s_t which cause the equilibrium in period t to be non-pooling when it would otherwise be pooling. Similarly, it is fragile to a peer-oriented intervention with n pieces of information if n i.i.d. signals to v_t causes an equilibrium to be non-pooling when it would otherwise be pooling. This definition mirrors the definition of fragility in Bikhchandani et al. (2021) but is augmented to allow for a signal about μ.

Proposition 4 (**Fragility**)
The following are true:

1. *If $\kappa < \kappa^{c.k.}$, for any pooling equilibrium there exists an N such that the pooling equilibrium is fragile to both an individual-oriented intervention and to a peer-oriented intervention*

[26] In support of this theory, Tevyaw et al. (2007) shows that the magnitude of the reduction in alcohol use was 3 times larger on average for group-level interventions than individual-level interventions.

with $n \geq N$ pieces of information. Further, after either intervention, social learning about fundamentals and preferences occurs.

2. *If $\kappa > \kappa^{c.k.}$, for any $n \in \mathbf{N} \cup \infty$, an individual-oriented intervention (respectively, peer-oriented intervention) with n pieces of information will never result in social learning about μ (respectively, θ).*

3. *If $\kappa > \kappa^{c.k.}$ and the correlation between the beliefs about θ and μ is equal to zero, the equilibrium is never fragile to an individual-oriented intervention with n pieces of information but may be fragile to a peer-oriented intervention with n pieces of information.*

The intuition behind the first result is that if $\kappa < \kappa^{c.k.}$, then when there is complete information the signaling equilibrium involves revelation. Further, that the equilibrium is pooling despite a $\kappa < \kappa^{c.k.}$ is necessitated by a strong negative correlation between the beliefs about θ and μ. When this correlation is strong, each player negatively updates about θ given his preference type, v_t. As a result, when the correlation is negative, changes in v_t cause a smaller change in a_t, increasing the reputational penalty from adapting to one's private information (as this is increasingly ascribed to v_t). A sufficiently large amount of information about either θ or μ will weaken the negative correlation in the beliefs about θ and μ, allowing the players to adapt to private information once more.

The second result says that if the conformity concerns are high, then giving information about only one dimension of uncertainty will be unsuccessful in spurring social learning. The proof for this result is precisely Proposition 2, which states that, for any prior beliefs about θ and μ, the players' beliefs about θ and μ cannot converge to the truth. The implication of this result is that even if a social planner designs a perfect individual-oriented intervention, the players will necessarily continue to pool on inaccurate perceptions of μ.

Finally, the intuition for the final result comes from the different effects of these interventions on the effective conformity concerns. If there is no correlation between the beliefs about θ and μ the first-order condition defining a player's decision in Equation (1.8) simplifies to:

$$a_t(1 + \kappa\alpha^2) = \gamma_t \mathbf{E}(\theta|\theta(t), s_t) + v_t + \kappa\alpha \mathbf{E}(\mu|\mu(t), v_t). \tag{1.12}$$

An individual-oriented intervention always decreases the weight players place on s_t, thus making the decision rule more sensitive to v_t. This increased sensitivity implies that increasing the information about θ magnifies the effective conformity concerns and thus cannot break a pooling equilibrium.

In contrast, a peer-oriented intervention has two effects. First every players wants to be perceived as the average preference type. Consequently, when $\tau_{\mu,t}$ is low, players with different preference types will have different perceptions of what the average preference type is. Intuitively, if each player has different perceptions of the population's average, then each player will adapt his decision to his preference type because doing so will ensure the public perception of his preference type will be in line with the population's average. This logic implies that when $\tau_{\mu,t}$ is low, the players have an added incentive to adapt.

The countervailing force is that when $\tau_{\mu,t}$ is low, the uncertainty over a given player's preference type is also high. As is standard in signaling games, when the uncertainty over a given player's preference type is higher, the same player has a greater incentive to signal, and thus a lower incentive to adapt.

Note that the relative value of $\tau_{\theta,t}$ has no affect on the first force but does effect the latter. To see why $\tau_{\theta,t}$ impacts the latter force, note that when $\tau_{\theta,t}$ is high, each decision is mostly determined by a player's preference type, v_t, and not their signal, s_t. As the decision is primarily a function of v_t, a sufficiently precise signal of v_t is generated. This precise signal implies the community's inference about a player's preference type is less sensitive to changes in the prior, such as an increase in $\tau_{\mu,t}$. As a result, when $\tau_{\theta,t}$ is high, increasing $\tau_{\mu,t}$ has a comparatively small effect on the community's inference about a player's preference type and a comparatively large effect on the player's inference about the community's average preference type. Finally, increasing $\tau_{\mu,t}$ makes the player's inference about the community's average preference type less sensitive to the player's own preference type, which gives that player a lower incentive to adapt to their private information. This intuition is seen in Figure 1.3 below: when $\tau_{\mu,t} > \tau_{\theta,t}$ (respectively, $\tau_{\mu,t} < \tau_{\theta,t}$) an increase in $\tau_{\mu,t}$ causes a change to a revealing equilibrium (respectively, pooling equilibrium).

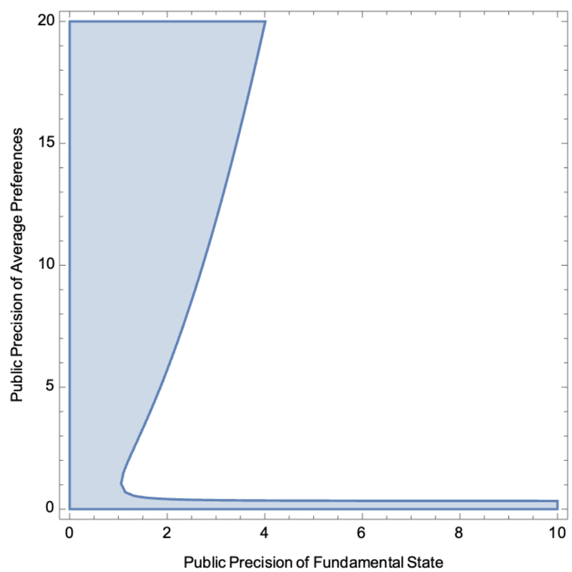


Figure 1.3: When the Signaling Equilibrium Involves Revelation

In the figure, the x-axis corresponds to $\tau_{\theta,t}$, the public precision of the fundamental state, and the y-axis corresponds to $\tau_{\mu,t}$, the public precision of the average preference type. The shaded region corresponds to when the signaling equilibrium involves revelation. The non-shaded region corresponds to when the signaling equilibrium is pooling. In the figure $\kappa = \gamma_t = 1$, which correspond to the weight of conformity and the relative value players place on adapting to the fundamental state.

However, even if neither a peer-oriented intervention nor an individual-oriented intervention break a pooling equilibrium, these interventions will influence which decision the players pool on. To gain intuition into the forces behind the change, I consider the following situation: suppose the equilibrium in period 1 was revealing and thereafter the decision rule is pooling. Recall a_1 denotes the decision chosen in period 1. This decision influences what the pooling decision will be in period 2, a^*, where for simplicity I assume $\gamma_2 = 1$ to derive:

$$a^*(a_1) = \mathbf{E}(\theta|a_1) + \mathbf{E}(\mu|a_1). \tag{1.13}$$

Further, recall that a_1 is a linear combination of both player one's private signal about θ, s_1, and player one's preference type, v_1, yielding:

$$a_1 = \lambda_\theta s_1 + \lambda_v v_1, \tag{1.14}$$

where λ denotes such weights. I now consider the following intervention where the public history is adapted to be either $\mathbf{h}_2(\theta) = \{a_1, \gamma_1, \theta\}$ or $\mathbf{h}_2(\mu) = \{a_1, \gamma_1, \mu\}$ and analyze the change in the pooling decision rule that follows. Suppose that the players utilize the decision-rule in Equation (1.14), resulting in a pooling decision rule denoted by $a^*(a_1)$ as in Equation (1.13) for all subsequent periods.

Proposition 5 (**Interventions**)
Denote by $\Delta(\theta)$ (respectively, $\Delta(\mu)$) as the difference between the new decision the players pool on compared to a^*. Then,

$$\Delta(\mu) = \mu\left(1 - \frac{\lambda_\theta\lambda_v}{\frac{1+2\tau_\theta}{1+\tau_\theta}\lambda_\theta^2 + \frac{\tau_\theta}{1+\tau_\mu}\lambda_v^2}\right) + \alpha_\mu a_1 \tag{1.15}$$

$$\Delta(\theta) = \theta\left(1 - \frac{\lambda_\theta\lambda_v}{\frac{1+2\tau_\mu}{1+\tau_\mu}\lambda_v^2 + \frac{\tau_\mu}{1+\tau_\theta}\lambda_\theta^2}\right) + \alpha_\theta a_1, \tag{1.16}$$

for some constants α_θ, α_v.

To understand the expressions above, let us now consider the effect of an individual-oriented intervention revealing θ, where a symmetric analysis occurs for μ. Upon revealing θ, the updated equilibrium perception of θ is θ. Further, the players re-evaluate the perception of μ as a function of both θ (the second term in the parentheses of Equation (1.16)) and a_1. The object of interest is how much the players decision changes with respect to μ. One can see that given information that θ is positive (respectively, negative) the players update that μ is negative (respectively, positive). Further, this update could be larger in magnitude than the update about the value of θ. Specifically, these cases occur when τ_μ is small (i.e., the players are uncertain about their peers' true preferences). Such a counter-update provides one rationale why the individual-oriented interventions have a small (if not negative) effect on behavior.

5.4 Designing Effective Interventions

While both interventions have their merits in different circumstances, the model predicts differential effectiveness. In the model, when conformity concerns are large, the players enter into a pooling equilibrium based on inaccurate perceptions of their peers' preference types and the fundamental state. Proposition 4 suggests that peer-oriented interventions may be preferred due to their ability to break a pooling equilibrium. Further, Proposition 5 suggests that even a perfect individual-oriented intervention alone may fail to shift the pooling decision in the direction of efficiency.

These predictions are broadly consistent with the results in Schroeder and Prentice (1998) and Bursztyn et al. (2020a) for two reasons. First, interventions addressing misperceptions of the average preference type are preferred. Second, in both settings conformity concerns are arguably high. If instead conformity concerns were low (or in the limit equal to zero), then similar to the theoretical literature, the optimal intervention would be individual-oriented. When the conformity concerns are low, knowledge of the average preference type is less decision-relevant and individual-oriented interventions allow the players to reach an efficient decision faster.

5.5 Changing Preferences and Peer-Oriented Interventions

The social psychology literature notes that "a society's perception of itself tends to lag behind actual changes in people's private beliefs and values," and argues this lag necessitates peer-oriented interventions to improve decision-making (Miller, 2023). To address these phenomena, I generalize the main model to allow μ to be time dependent and follow an autoregressive process parameterized as,

$$\mu^*(t) = \rho\mu^*(t-1) + (1-\rho)\psi_t, \tag{1.17}$$

where ψ_t is independent across time and distributed as $N(0, \tau_\mu)$, with τ_μ being the same τ_μ as in Section 2 where $\mu \sim N(0, \tau_\mu)$. Making this equivalence allows the ex-ante uncertainty about $\mu^*(t)$ to be equal to that of μ for all t. The first implication of changing preference types is that switches to pooling decision rules are temporary.

In the primary analysis, once the players switch to a pooling decision rule, the players pool for all subsequent periods. However, given the changing average preference type, if the players pool in all subsequent periods, the public beliefs eventually converge back to beliefs where revelation occur. Finally, in this extension, the players utilize the decision rule with revelation for arbitrarily many periods and θ remains fixed, implying that it is without loss to assume that θ is common knowledge and fixed at zero when analyzing the asymptotic behavior of the community. The following proposition analyzes the asymptotic behavior of the players as a function of the conformity concerns.

Proposition 6 (**Shifting Preferences**)
Suppose μ follows an autoregressive process as defined in Equation (1.17) and denote by $\tau_{\mu,t}$ the precision about the public beliefs about $\mu^*(t)$ in period t. The signaling equilibrium in period t involves revelation if and only if

$$\tau_{\mu,t} < \frac{\kappa^{c.k.}}{\kappa - \kappa^{c.k.}}. \tag{1.18}$$

As a result, there exists a threshold $\kappa^* \in \left(\kappa^{c.k.}, \kappa^{c.k.} \cdot (\tau_\mu + 1)/\tau_\mu\right)$ such that

1. *If $\kappa \leq \kappa^*$ the signaling equilibrium involves revelation in all periods and $\tau_{\mu,t} \to \tau(\rho)$. Further, $\tau'(\rho) < 0$ and $\lim_{\rho \to 0} \tau(\rho) = \infty$.*

2. *If* $\kappa \in \left(\kappa^*, \kappa^{c.k.}(\tau_\mu + 1)/\tau_\mu\right]$ *the signaling equilibrium involves revelation (i.e., the precision in the beliefs is less than the right-hand side of Equation* (1.18)*) for infinitely many periods and the players pool for infinitely many periods (i.e., the precision in the beliefs is greater than the right-hand side of Equation* (1.18)*), and* $\tau_{\mu,t}$ *does not converge.*

3. *If* $\kappa > (\tau_\mu + 1)/\tau_\mu$*, the signaling equilibrium is pooling in all periods and* $\tau_{\mu,t} = \tau_\mu$ *in all periods.*

Before analyzing the asymptotic results, let us build intuition for when the signaling equilibrium in period t involves revelation. Note that in any equilibrium with revelation a player's preference type is fully revealed in equilibrium, thus $\tau_{\mu,t}$ does not impact the posterior beliefs about a player's preference type. As a result, the only effect of an increase in $\tau_{\mu,t}$ is that each player's beliefs about μ places less weight on v_t. As each player wants his perceived preference type to equal μ, and thus chooses a decision which is responsive towards his beliefs about μ, an increase in $\tau_{\mu,t}$ decreases each player's incentive to respond to v_t. Such a force is seen in Equation (1.18), which notes that as $\tau_{\mu,t}$ increases the condition on κ for an equilibrium with revelation to exist becomes more stringent.

Given Equation (1.18), it is immediate that if $\kappa \leq \kappa^{\text{c.k.}}$, then the signaling equilibrium would involve revelation for all periods. If the signaling equilibrium involved revelation for all periods, $\tau_{\mu,t}$ would monotonically increase to a constant $\tau(\rho) < \infty$, implying the players have imprecise beliefs about $\mu^*(t)$. Note that the players can never perfectly infer $\mu^*(t)$ because $\mu^*(t)$ changes each period. Thus the condition for an equilibrium with revelation in all periods is determined by κ^*, which binds Equation (1.18) when $\tau_{\mu,t} = \tau(\rho)$.

When the conformity concerns are greater than this value but small enough such that the equilibrium involves revelation in period 1, the players oscillate between a decision rule with revelation and pooling in the limit. This occurs because as more periods with revelation occur (respectively, pooling), the beliefs become more precise (respectively, imprecise), eventually necessitating a pooling (respectively, revealing) equilibrium. As a result, the beliefs do not converge, but rather oscillate around the value $\tau_{\mu,t}$ which binds equation (1.18). This value is decreasing in κ, implying that higher values of the conformity concerns implies worse beliefs in the limit, and these worse beliefs are caused by a greater number of periods utilizing a pooling equilibrium.[27] Finally, if the conformity concerns are large enough that the players pool in period 1, then, by induction, the players pool in all periods.

What this proposition implies is that, all else equal, groups who have stronger weights on conformity concerns wait longer to adapt to the underlying conditions. This result is broadly consistent with the social psychology literature on pluralistic ignorance and how norms change. In a review article, Miller (2023) states, "widespread changes in private attitudes change are not sufficient for social norm change. The group's recognition that its collective attitudes have changed is also necessary. Without this recognition, norm change will be impeded." One can interpret this through the model as follows: if the players are pooling in period t on a low decision due to a low belief $\mu^*(t)$, despite a high value for v_{t+1} signaling a change in private attitudes has occurred, this change is *not* sufficient to change the decision rule.

[27] A precise characterization of the asymptotic distribution of $\tau_{\mu,t}$ when $\kappa > \kappa^*$ is challenging because $\tau_{\mu,t}$ evolves according to a non-continuous discrete dynamical system.

Rather, enough periods must pass for the group to be certain that their collective attitudes have changed.

Such a process where the players utilize a responsive decision rule for some number of periods before switching to a pooling decision rule and vice versa, is also in line with the psychology literature on changing norms. For instance, Miller and Prentice (1994) suggest that one major source of pluralistic ignorance is a "conservative lag" whereby opinions change but not decisions. This can be viewed as the periods in which the players utilize a pooling decision rule despite $\mu^*(t)$ changing. As $\mu^*(t)$ has changed over this time frame, the subsequent player's decision differs from what the previous pooling decision was. Such a change exemplifies the "liberal leap" also described in Miller and Prentice (1994).[28] This extension shows how groups that have higher conformity concerns will have greater degrees of pluralistic ignorance and adapt slower to changes in private attitudes. Finally, note that changing preference types are an additional motivation behind the peer-oriented interventions considered in Section 5.2. Without these interventions, the social norms will lag behind the true attitudes in the population, causing inefficiencies.

6 Extensions

In this section I consider three extensions that serve as robustness checks for the assumptions in the main analysis. First, a common assumption in the social learning literature is that each player is short-lived. While long-lived players traditionally are able to observe more signals, and thus make more efficient decisions absent conformity concerns, the conformity concerns are amplified through a ratchet effect. Second, I show that when off-path beliefs satisfy D1 from Cho and Kreps (1987), non-linear decision rules fail to exist in the benchmark environment in Section 3. Finally, I discuss how the results of the model generalize to alternative distributional assumptions with non-linear decision rules.

6.1 Long-Lived Players

This subsection analyzes the incentives of a longer-lived player. To model such a phenomena, I continue to index time by $t \in \{0, 1, \dots\}$, but label the players by $i \in \{0, 1, \dots\}$. I denote by $t(i)$ the first period in which player i appears. Further, each player, i, continues onto the next period with uniform probability p and with probability $1 - p$ is replaced by player $i + 1$. Each player has a discount factor $\delta < p$, which includes both the probability of continuation and inter-temporal discounting. The model in the primary analysis considers $p = 0$, and as a result $\delta = 0$, and each player makes a single decision. When the players are long-lived, the utility for player i is as follows:

$$u^{t(i)} + \delta u^{t(i)+1} + \delta^2 u^{t(i)+2} + \dots. \tag{1.19}$$

[28]The economics literature produces similar findings. For instance, an immediate implication of the model is that when the autoregressive process is more volatile, the players will spend less periods in the pooling decision rule. This result is in line with Giuliano and Nunn (2021), which shows that populations in uncertain climates have less persistent norms.

This game featuring persistent private information entails well-known non-trivial modeling choices. Assuming θ is common knowledge greatly simplifies the analysis. Further, once θ is assumed to be common knowledge, it is without loss of generality to consider $\theta = 0$.

The final simplification I make is a restriction to the following class of equilibria that mimic the characterization in the primary analysis. In period $t(i)$, player i utilizes a linear decision rule. If there exists a linear decision rule with revelation, such a decision rule is utilized in period $t(i)$. As player i's preference type is then fully revealed on-path, in all subsequent periods player i utilizes a pooling decision rule as determined by the posterior mean of player i's perceived preference type. On path, this corresponds to player i's true preference type and is hence socially optimal. When there does not exist a revealing linear decision rule that constitutes an equilibrium in period $t(i)$, the decision rule is fully pooling in period $t(i)$. As no new information is learned, period $t(i) + 1$ is equivalent to $t(i)$, and thus a pooling decision rule will be used in period $t(i) + 1$, and by induction, for all subsequent periods for player i. Note that this set of equilibrium refinements prescribes a unique decision rule for each player because the pooling decision rule is always an equilibrium.

Given the equilibrium selection, the trade-off for player i in period $t(i)$ is that a higher decision today allows for a higher decision to be taken in all subsequent periods because a higher decision increases the perception of player i's preference type. However, the perception being higher also implies a higher conformity loss in all subsequent periods. The following proposition states how this trade-off changes as the discount factor, δ, increases.

Proposition 7 (**Long-Lived Players**)
When players are long-lived with a discount factor, δ, have a conformity weight, κ, and θ is common knowledge, then an equilibrium with revelation exists if and only if:

$$\kappa < \kappa^{c.k.}(1-\delta)\frac{\tau_{\mu,t}+1}{\tau_{\mu,t}}, \tag{1.20}$$

where $\tau_{\mu,t}$ denotes the precision about the public beliefs about the average preference type.

To understand this expression, first note that when $\delta = 0$ the condition for an equilibrium is equivalent to the condition in Proposition 6. Let us now turn to how the discount factor, δ, affects the condition. To understand why an increase in δ increases the effective conformity concerns, note that, in equilibrium, choosing a marginally more conforming decision in period $t(i)$ has only a second-order effect on the adaptation loss in all subsequent periods. Recall, in equilibrium, player i's preference type is correctly inferred, and thus the decision in all future periods is v_i. If player i selected a marginally more conforming decision, this deviation would generate only a second-order loss in adaptation in subsequent periods, as the adaptation loss is quadratic. However, a marginally more conforming decision in period $t(i)$ has a first-order impact on the conformity loss in all subsequent periods. Thus, an increase in the discount factor effectively scales the conformity concerns, as there is no future benefit to adaptation; however, there is a greater future benefit from conformity.

Proposition 7 generates the stark prediction that the degree of misperceptions about the average preference type, μ, is shaped by the players' discount factor. Ignoring integer constraints, the players cease utilizing decision rules with revelation when $\tau_{\mu,t}$ binds Equation (1.20). Solving Equation (1.20) implies that when the players' discount factor is higher, the players learn less about μ. This result adds a competing force to those suggested by the social learning literature. In that literature, long-lived players have an ability to observe more data and thus make more accurate decisions. This extension highlights that when conformity concerns are present, long-lived players might make worse decisions.

6.2 Non-Linear Equilibria

The main analysis analyzed signaling equilibria: linear and socially optimal Perfect Bayesian Equilibria. This subsection shows that while non-linear Perfect Bayesian Equilibria may exist, they do not satisfy the D1 refinement from Cho and Kreps (1987).[29] This refinement states that off-path beliefs are concentrated on the types who have the largest incentive to deviate to such a decision. As in Section 3, when θ and μ are common knowledge it suffices to consider a static version of the game and drop any time-dependence. Further, Equation (1.1) reduces to:

$$u(v,a) = -(a-v)^2 - \kappa \int b^2 \phi(b,a)db. \tag{1.21}$$

In doing so, I first restate the definition of a "central pooling equilibrium" from Bernheim (1994). A central pooling equilibrium is an equilibrium in which $a(v) = c \quad \forall v \in [\underline{v}, \bar{v}]$ where $\underline{v} \leq 0 \leq \bar{v}$ and $a(v)$ is strictly monotone when $v \notin [\underline{v}, \bar{v}]$. The lemma below shows that any equilibrium which satisfies D1 is a central pooling equilibrium.

Lemma 4 (**Class of Equilibria**)
Any equilibrium satisfying D1 is a central pooling equilibrium. In this equilibrium, $a(v) = c^* \quad \forall v \in [\underline{v}, \bar{v}]$ *where* $\underline{v} \leq 0 \leq \bar{v}$. *Further, for* $v \notin [\underline{v}, \bar{v}]$ $a(v)$ *is continuously differentiable with a derivative that satisfies:*

$$a'(v) = \frac{\kappa v}{v - a(v)} > 0. \tag{1.22}$$

For the proof, I refer the reader to Theorem 3 in Bernheim (1994). The intuition behind the result is that the D1 refinement implies that the decision rule is monotone. Given a monotone decision rule, one can show that the D1 refinement further implies that a jump discontinuity cannot arise outside of a central pool. Finally, outside the central pool one can show strict monotonicity of the decision rule, which implies a well-defined inverse of the decision rule. This inverse can be substituted in for $\phi(b,a)$ to generate the differential equation in the lemma.

The primary difference relative to Bernheim (1994) is that I assume that the support of the distribution of v equals the real line whereas Bernheim (1994) assumes the support of v equals a bounded interval. As such, the solution to Equation (1.22) must exist over the

[29] Cho and Kreps (1987) define D1 for signaling games. Their definition applies to my setting if one views the reputation as stemming from the player interacting with one randomly drawn member of the community who gives the player a reward in accordance with this community members' perception of the player's preference type.

real line in my setting, but need only exist over a bounded interval in Bernheim (1994).[30] As there are multiple solutions to the differential equation (cf. Figure 1), one cannot rule out non-linear equilibria with an equilibrium uniqueness result. To do so necessitates solving the differential equation in Equation (1.22) in closed form, which is done in Appendix A to prove the following proposition.

Proposition 8 (**Non-existence of Non-Linear Equilibria**)
Any equilibrium satisfying D1 is linear.

This result gives support for the restriction to linear equilibria in the main analysis. Further, this result implies that if $\kappa > \kappa^{\text{c.k.}}$, no equilibrium with revelation satisfies the D1 refinement.

6.3 General Distributions

Throughout the analysis I focused on the Gaussian distribution which allowed for closed-form solutions and precise comparative statics. In this subsection, I will discuss to what extent these results generalize to different distributions. Recall that the analysis in the previous subsection, which assumed no uncertainty over the fundamental state θ, or the average preference type μ, allowed for any distribution over v_t, the preference type of player t, with a continuous density with support equal to the real line. In that analysis the equilibrium is pooling if and only if the conformity concerns, κ, exceed $\kappa^{\text{c.k.}}$.

The environment with general distributions and both dimensions of uncertainty is intractable.[31] Therefore, I will conduct two separate analyses, each focusing on a different dimension of uncertainty. In this subsection I will provide the intuition for the case where θ is common knowledge and the players are learning μ, and Appendix B contains a parallel analysis when θ is uncertain and μ is known. Recall that the analysis when θ is common knowledge with the Gaussian assumptions is contained in Section 5.5 when $\rho = 0$ or Section 6.1 when $\delta = 0$. In both analyses, when players have more uncertainty about μ, the players can coordinate on an equilibrium with revelation for higher values of κ. This can be used to show that a higher κ implies the players switch to a pooling equilibrium with less precise beliefs and, thus, ultimately harbor these less precise beliefs. This subsection analyzes to what extent this point relies on the properties of the Gaussian distribution.

When the distributions of μ and v_t are Gaussian, the public beliefs about μ at time t will satisfy $\mu(t) \sim N(\bar{\mu}(t), \tau_{\mu,t})$, for some mean $\bar{\mu}(t)$ and precision $\tau_{\mu,t}$. Further, the updating rule

[30] In Bernheim (1994), the decision rule from this differential equation is non-linear and equilibria with partial revelation exist. The reason for the stark difference is that Bernheim (1994) considers a bounded support of v. If, for example, 1 is the supremum of the support, then the definition of a central pooling equilibrium implies that the player with $v = 1$ has his preference type revealed in equilibrium, and thus might as well choose his preferred decision. Hence, the differential equation in Bernheim (1994) has an initial condition of $a(1) = 1$. In the setting with an infinite support, no player chooses $a(v) = v$ outside the central pool, which generates different conditions for equilibrium existence. This distinction is further discussed in Appendix B.

[31] In Section 9.4, I detail the analytical challenges that arise when both factors are present. Intuitively, the differential equations governing the decision rules transform into partial differential equations, significantly complicating the analysis.

for the conditional expectation of μ given the realization of v_t has the following closed-form expression as shown in the utility of player t below:

$$-(a_t - v_t)^2 - \kappa\Big(\hat{v}_t(a_t) - \underbrace{\frac{v_t + \bar{\mu}(t)\tau_{\mu,t}}{1+\tau_{\mu,t}}}_{\mathbf{E}(\mu|v_t)}\Big)^2. \tag{1.23}$$

Here, the constant equal to 1 is the precision of player t's preference type relative to μ. Uncertainty over μ (i.e., a lower value of $\tau_{\mu,t}$) gives player t an additional incentive to respond to v_t, which implies a higher threshold value of κ for the existence of an equilibrium with revelation when $\tau_{\mu,t}$ is higher. This point does not rely on the Gaussian assumption.

Let us consider any full support and atomless distribution of v_t. Further, I denote by $g_t(v_t|a_1,\ldots,a_{t-1}) = \mathbf{E}(\mu|v_t, a_1, \ldots, a_{t-1})$, which, with an abuse of notation, will be denoted as $g_t(\cdot)$. Given this notation, the generalization of the differential equation in Equation (1.8), which describes the decision rule in any equilibrium with revelation satisfying D1, is:

$$a_t'(v_t) = \frac{\kappa\big(v_t - g_t(v_t)\big)}{v_t - a_t(v_t)}. \tag{1.24}$$

Equipped with this differential equation, one can show the following result.

Proposition 9 (**General Distributions with θ known**)
Denote by $i_t := \inf_x g_t'(x) \le \sup_x g_t'(x) := s_t$, where $g_t(x) = \mathbf{E}(\mu|x, a_1, \ldots, a_{t-1})$. There exists an equilibrium with revelation that satisfies D1 if the conformity concerns, κ, satisfy

$$\kappa \le \frac{\kappa^{c.k.}}{1 - i_t}. \tag{1.25}$$

Further, no equilibrium with revelation satisfies D1 if

$$\kappa > \frac{\kappa^{c.k.}}{1 - s_t}. \tag{1.26}$$

This proposition gives a separate necessary and sufficient condition for the existence of an equilibrium with revelation. In the Gaussian analysis, $g_t(v_t)$ is linear which implies $s_t = i_t$. When $s_t = i_t$, these necessary and sufficient conditions coincide, implying these bounds are tight. The basic insight is that an equilibrium with revelation exists if and only if the differential equation in Equation (1.24) has a solution. For a solution to exist, $a_t(v_t)$ must never exceed v_t to ensure that the denominator is non-zero. Further, a higher value of conformity concerns, κ, uniformly increases the solution to the differential equation in Equation (1.24) given any initial condition. For this reason, when κ is sufficiently low (respectively, high) an equilibrium with revelation exists (respectively, does not exist).

This proposition also characterizes the effect of uncertainty about μ on the degree of revelation. One can view both s_t and i_t as measures of the degree of uncertainty. Similar to the Gaussian analysis, this greater population uncertainty gives an added incentive for the players to adapt to their private information and ultimately increases the cutoff κ^*.

As a result of this proposition, one can recover similar results to the main analysis. Namely, whenever the beliefs become sufficiently precise relative to the conformity concerns, the players switch to a pooling equilibrium based on imprecise perceptions of μ. In the Gaussian analysis, $s_t = i_t$ and these values follow a deterministic process. This determinism implied a monotone relation between κ and the asymptotic uncertainty about μ. In this general analysis, such a claim need not be true.[32] However, the result that pluralistic ignorance exists in all equilibria if and only if κ exceeds $\kappa^{\text{c.k.}}$ continues to hold. To see why, note that if $\kappa < \kappa^{\text{c.k.}}$ the sufficient condition for the existence of an equilibrium always holds, and thus with an infinite number of periods with revelation, the players learn the true average. In contrast, if $\kappa > \kappa^{\text{c.k.}}$, then there must exist sufficiently precise beliefs such that the necessary condition for equilibrium existence fails to hold.

7 Conclusion

This paper studies how conformity concerns impact social learning and what interventions are effective when social learning fails. To do so, I enrich a standard model of social learning by adding: (i) a player's desire to adapt to not only a fundamental state but also his private preference type, (ii) an assumption that players have conformity concerns over how the community perceives their private preference type, and (iii) an assumption that there is aggregate uncertainty about the distribution of private preference types in the population. I show that as the players' beliefs about the fundamental state become more precise, the equilibrium penalty experienced by a player who adapts to his private information or his private preference type increases, creating endogenous self-censorship. Further, I show that if the initial conformity concerns are sufficiently high, the endogenous self-censorship not only dampens but eliminates the player's adaptation, resulting in a switch from a revealing to a pooling equilibrium in finite time. Such a switch to pooling implies that forever after the players hold imprecise beliefs about both the fundamental state and the preference types of their peers; the latter is a common finding in social psychology, defined as pluralistic ignorance. Not only are the players pooling (and thus unable to adapt to their private preference types), they pool on an inefficient decision based on these imprecise beliefs. Finally, information about the fundamental state has a lower ability than information about peers' preferences to break a pooling equilibrium. My theoretical result that providing information about the preferences types of one's peers is more effective than information about the fundamental state provides a framework to formalize intuitions extensively discussed empirically in social psychology and economics.

This paper introduced a theoretical methodology that can be used to analyze pluralistic ignorance and how decisions change upon dispelling pluralistic ignorance. I hope this framework can be used to analyze related topics in the social sciences. For instance, related to pluralistic ignorance, there is a large literature on "false polarization" whereby individuals of two distinct subgroups will incorrectly perceive the preferences of the two groups as further apart than reality. Further, related to interventions addressing pluralistic ignorance, there exist numerous empirical and qualitative studies on "risky and cautious shifts," whereby upon learning whether the members in their group have risky (respectively, cautious) opinions, the

[32] Further, at this level of generality equilibria with partial revelation may exist and maximize the expected utility of the players.

opinions of the group will shift to be more polarized than the opinions of the group members themselves (cf. Sunstein, 2009).

8 Appendix A

Proof of Proposition 1. Solving Equation (1.3) proves the existence of $\kappa^{\text{c.k.}}$ and the results when $\kappa \geq \kappa^{\text{c.k.}}$. Finally, the equilibrium with revelation is Pareto superior to the pooling equilibrium, which completes the proof. □

Proof of Lemma 1. Fix a given period t and a conjectured linear belief $\alpha_t a_t + \beta_t$. The first-order condition given a conjectured belief of $\alpha_t a_t + \beta_t$ is equal to Equation (1.8). The normality assumption implies that the conditional expectation of both θ and μ will be linear in both v_t and s_t with an intercept. Using this observation, then for some exogenous constants, $c_{1,t}, \ldots, c_{6,t}$,

$$a_t(1 + \kappa\alpha_t^2) = \kappa\alpha_t\beta_t + c_{1,t} + c_{2,t}\alpha_t\kappa + s_t(c_{3,t} + \kappa\alpha_t c_{4,t}) + v_t(1 + c_{5,t} + \kappa\alpha_t c_{6,t}). \tag{1.27}$$

Simplifying implies,

$$\frac{a_t(1 + \kappa\alpha_t^2) - \kappa\alpha_t\beta_t - c_{1,t} - c_{2,t}\alpha_t\kappa}{c_{3,t} + \kappa\alpha_t c_{4,t}} = s_t + v_t\frac{1 + c_{5,t} + \kappa\alpha_t c_{6,t}}{c_{3,t} + \kappa\alpha_t c_{4,t}}. \tag{1.28}$$

Given this sufficient statistic, the posterior belief about v_t is,

$$\mathbf{E}\Big(v_t | h_t, a_t\Big) = c_{7,t} + c_{8,t}(\alpha_t, \kappa)\frac{a_t(1 + \kappa\alpha_t^2) - \kappa\alpha_t\beta_t - c_{1,t} - c_{2,t}\alpha_t\kappa}{c_{3,t} + \kappa\alpha_t c_{4,t}}, \tag{1.29}$$

where $c_{8,t}(\alpha_t, \kappa) \neq 0$ is determined by both the prior beliefs and the conjectured equilibrium slope α_t. In equilibrium, the conjecture must be consistent implying the right-hand side of Equation (1.29) must equal $\alpha_t a_t + \beta_t$. This equality is stated below:

$$c_{8,t}(\alpha_t, \kappa)(1 + \kappa\alpha_t^2) = \alpha_t(c_{3,t} + \kappa\alpha c_{4,t}) \tag{1.30}$$

$$c_{7,t}(c_{3,t} + \kappa\alpha_t c_{4,t}) + c_{8,t}(\alpha_t, \kappa)(-\kappa\alpha_t\beta_t - c_{1,t} - c_{2,t}\kappa\alpha_t) = \beta_t(c_{3,t} + \kappa\alpha_t c_{4,t}). \tag{1.31}$$

Note that for any solution to Equation (1.30), there exists a solution to the equation for β_t. This is because one can simplify Equation (1.31) to,

$$\beta_t\Big(c_{3,t} + \kappa\alpha_t c_{4,t} + \kappa\alpha_t c_{8,t}(\alpha_t, \kappa)\Big) = c_{7,t}(c_{3,t} + \kappa\alpha_t c_{4,t}) + c_{8,t}(\alpha_t, \kappa)(-c_{1,t} - c_{2,t}\kappa\alpha_t). \tag{1.32}$$

One can simplify the coefficient on β_t above as follows:

$$\begin{aligned} c_{3,t} + \kappa\alpha_t c_{4,t} + \kappa\alpha_t c_{8,t}(\alpha_t, \kappa) &= c_{8,t}(\alpha_t, \kappa)\frac{(1 + \kappa\alpha_t^2)}{\alpha_t} + \kappa\alpha_t c_{8,t}(\alpha_t, \kappa) \\ &= c_{8,t}(\alpha_t, \kappa)(\frac{1}{\alpha_t} + 2\alpha_t\kappa), \end{aligned} \tag{1.33}$$

where the first equality comes from Equation (1.30). Finally, as $c_{8,t}(\alpha_t, \kappa) \neq 0$ and in any equilibrium with revelation $\alpha_t \neq 0$, a unique solution for β_t always exists.

Thus, the necessary and sufficient condition for an equilibrium with revelation is Equation (1.30). Finally, all of the terms in this equation are independent of the means of the prior beliefs ($c_{1,t}$ and $c_{2,t}$) and only condition on the precision matrix.

The proof of this first result implies that the realizations of a_t have no impact on whether or not beliefs converge, because in the Gaussian learning model, the realizations of a_t effect only the mean. Denote by τ_t the precision matrix of the beliefs at time t with the following parametrization.

$$\tau_t = \begin{pmatrix} \tau_{1,t} & \tau_{2,t} \\ \tau_{2,t} & \tau_{3,t} \end{pmatrix}. \tag{1.34}$$

The beliefs update as follows in any equilibrium with revelation,

$$\begin{aligned} \tau_{1,t+1} &= \tau_{1,t} + \phi_t \\ \tau_{3,t} &= \tau_{3,t} + 1 - \phi_t \\ \tau_{2,t} &= \tau_{2,t} + \sqrt{(1-\phi_t)\phi_t} \\ \phi_t &= \frac{(c_{3,t} + \kappa\alpha c_{4,t})^2}{(c_{3,t} + \kappa\alpha_t c_{4,t})^2 + (1 + c_{5,t} + \kappa\alpha_t c_{6,t})^2}. \end{aligned} \tag{1.35}$$

Further, $\theta(t) \rightarrow_p \theta$ if and only if:

$$\lim \frac{\tau_{3,t}}{\tau_{1,t}\tau_{3,t} - \tau_{2,t}^2} \rightarrow 0, \tag{1.36}$$

namely the variance about θ converges to zero. If this variance converges to zero, however, then $c_{3,t}, c_{4,t}$, and $c_{5,t}$ (how s_t impacts the conditional expectation of θ, how s_t impacts the conditional expectation of μ, and how v_t impacts the conditional expectation of θ, respectively) all converge to zero. Further, $c_{6,t}$ (how v_t impacts the conditional expectation of μ) remains weakly positive. As a result, if $\theta(t) \rightarrow_p \theta$ and $\mu(t) \not\rightarrow_p \mu$, then $\phi_t \rightarrow 0$. However, if $\phi_t \rightarrow 0$ and there exist infinitely many periods with revelation, then $\tau_{3,t} > \tau_{1,t} > 0$ implying

$$\lim \frac{\tau_{1,t}}{\tau_{1,t}\tau_{3,t} - \tau_{2,t}^2} \rightarrow 0, \tag{1.37}$$

namely the variance about μ converges to zero. This convergence implies $\mu(t) \rightarrow_p \mu$.

Similarly, if $\mu(t) \rightarrow_p \mu$, but $\theta(t) \not\rightarrow_p \theta$, then $c_{5,t}, c_{6,t}$ and $c_{4,t}$, but $c_{3,t}$ remains bounded away from zero. As a result, ϕ_t converges to a constant, $\phi \in (0, 1)$. Hence,

$$0 < \lim \frac{\tau_{3,t}}{\tau_{1,t}\tau_{3,t} - \tau_{2,t}^2} < \lim \frac{(\phi + \epsilon)\tau_{1,t}}{\tau_{1,t}\tau_{3,t} - \tau_{2,t}^2}. \tag{1.38}$$

However, $\theta(t) \rightarrow_p \theta$ implies the outer limit converges to zero, and as a result so too does the inner limit. □

Proof of Lemma 2. I proceed by contradiction. Note that Lemma 1 implies $\theta(t) \rightarrow_p \theta \iff \mu(t) \rightarrow_p \mu$. Hence, we may suppose by contradiction that $\theta(t) \rightarrow_p \theta$ and $\mu(t) \rightarrow_p \mu$.

For the beliefs to converge, infinitely many periods of a decision rule with revelation must occur which is equivalent to Equation (1.30) holding for infinitely many periods. This implies

(i) Equation (1.30) must hold as $t \to \infty$, $\theta(t) \to_p \theta$, and $\mu(t) \to_p \mu$. Further, (ii) because $\kappa > \kappa^{c.k.}$, there exists an $\epsilon > 0$ such that,

$$(1 + \kappa\alpha^2) - \alpha > \epsilon \quad \forall \alpha \geq 0. \tag{1.39}$$

Combining (i) and (ii), it must be that

$$\lim \frac{c_{8,t}(\alpha_t, \kappa)}{c_{3,t} + \kappa\alpha_t c_{4,t}} \not\to 1. \tag{1.40}$$

However, recall that this ratio in Equation (1.40) is defined by the following equation, as can be seen by manipulating Equation (1.28).

$$\begin{aligned} c_{9,t} + \frac{c_{8,t}(\alpha_t, \kappa)}{c_{3,t} + \kappa\alpha_t c_{4t}}{}_t \, (c_{3,t} + \kappa\alpha_t c_{4,t}) + v_t(1 + c_{5,t} + \kappa\alpha_t c_{6,t})\Big) \\ = \mathbf{E}(v_t | s_t(c_{3,t} + \kappa\alpha_t c_{4,t}) + v_t(1 + c_{5,t} + \kappa\alpha_t c_{6,t})), \end{aligned} \tag{1.41}$$

for some constant $c_{9,t}$. Finally, because the beliefs converge, then $c_{3,t}$, $c_{4,t}$, $c_{5,t}$, and $c_{6,t}$ converge to zero implying that the right-hand side converges to $\mathbf{E}(v_t|v_t)$. This implies that the limit in Equation (1.40) does converge to 1, which derives our contradiction. □

Proof of Lemma 3. It is sufficient to show that in every period the equilibrium involves revelation. To do so, we must show that there exists a solution to Equation (1.30). First, one can show that as $\alpha_t \to \infty$ the left-hand side is greater than the right-hand side, implying that a sufficient condition is that there exists conjectured beliefs where,

$$c_{8,t}(\alpha_t, \kappa)(1 + \kappa\alpha_t^2) < \alpha_t(c_{3,t} + \kappa\alpha_t c_{4,t}). \tag{1.42}$$

Further, as $\kappa < \kappa^{\text{c.k.}}$, upon setting $\alpha_t = 2$ implies that a sufficient condition is

$$\frac{c_{8,t}(2, \kappa)}{c_{3,t} + 2\kappa c_{4,t}} \leq 1 + \epsilon. \tag{1.43}$$

However, for any ϵ this inequality holds for sufficiently precise beliefs about θ and μ, because (i) the left-hand side is continuous with respect to the variance matrix of the public beliefs about θ, μ, (ii) the left-hand side is equal to one when the variance matrix is equal to zero. (i) and (ii) imply that there exists sufficiently precise beliefs where such a condition holds for not only those beliefs but any beliefs more precise. □

Proof of Proposition 2. The proof of statement (3) is a direct consequence of Lemma 2 and the proof of statement (2) is a direct result of Lemma 3.

The proof of statement (1) is as follows. I claim that $\tilde{s}(0, 0, \theta(t), \mu(t)) < 1 - \epsilon$ is a sufficient condition, where $\tilde{s}(\cdot)$ is as defined in Equation (1.10). This condition is sufficient as (i) the left-hand side of Equation (1.10) is larger as $\alpha \to \infty$ (ii) that such a solution holds when $\kappa = 0$, and (iii) the condition is continuous with respect to κ.

When $\kappa = \alpha = 0$, Equation (1.8), simplifies to:

$$\gamma_t \mathbf{E}(\theta | \theta(t), \mu(t), s_t, v_t) + v_t. \tag{1.44}$$

Further, given the definition of $\tilde{s}(\alpha, \kappa, \theta(t), \mu(t))$ in Equation (1.9), Equation (1.44) has a sensitivity strictly less than $1-\epsilon$ for any beliefs concluding the proof.

The proof of the final statement, whereby there exist an open set of parameter values where the beliefs do not converge, despite $\kappa < \kappa^{\text{c.k.}}$ can be seen by noting that Equation (1.30) is cubic in α_t. One can further show that because the coefficient on the term α_t^3 is negative the signaling equilibrium involves revelation if and only if there exists a positive root. Further, by Descartes rule of signs, there may be either zero or two positive roots. Hence, the necessary and sufficient condition for a positive solution to exist is for there to exist three roots. Further, for any cubic polynomial $A + Bx + Cx^2 + Dx^3$, there exist three real roots if and only if

$$(-27 \cdot D^2 \cdot A^2 + 18 \cdot B \cdot C \cdot D \cdot A - 4 \cdot D \cdot B^3 - 4 \cdot C^3 \cdot A + B^2 \cdot C^2) \geq 0. \tag{1.45}$$

Hence it suffices to show that the Equation (1.45) is strictly negative despite $\kappa < \kappa^{\text{c.k.}}$. As such an expression is continuous, demonstrating that this expression may be strictly negative will prove the results. Further, upon simplifying this expression one can find a precision matrix of the form in Equation (1.34), where Equation (1.45) is negative when $\gamma = 1$ and $\kappa = .9\kappa^{\text{c.k.}}$.[33] □

Proof of Proposition 3. This analysis uses τ_ϵ and $\tau_{v_{\text{ind}}}$ to denote the precision of a player's signal and precision of a player's preference type, which were normalized to one in the main analysis by assuming that the variance of ϵ_t and ν_t were equal to one. A linear decision rule can be written as,

$$\begin{aligned} a_t &= \beta_t \bar{\theta}(t) + \alpha_t \big(v_t + \mathbf{E}(\theta | \theta(t), s_t) - \bar{\theta}(t)\big). \\ \frac{a_t - \beta_t \bar{\theta}(t)}{\alpha_t} &= \big(\mathbf{E}(\theta) - \bar{\theta}(t)\big) + v_t \\ &= \Big(\frac{\tau_{\theta,t}\bar{\theta}(t) + \tau_\epsilon s_t}{\tau_{\theta,t} + \tau_\epsilon} - \bar{\theta}(t)\Big) + v_t \\ &= \frac{\tau_\epsilon(\theta - \bar{\theta}(t) + \epsilon_t)}{\tau_{\theta,t} + \tau_\epsilon} + v_t, \end{aligned} \tag{1.46}$$

for some slope α_t and intercept β_t. Note that the left-hand side can be computed given a_t and yields v_t plus a Gaussian random variable $\frac{\tau_\epsilon(\theta - \bar{\theta}(t) + \epsilon_t)}{\tau_{\theta,t} + \tau_\epsilon}$. Next, note that $\theta - \bar{\theta}(t)$ and ϵ_t are mean zero, hence the random variable $(a_t - \beta_t \bar{\theta}(t))/\alpha_t - v_t$ is mean zero. Further, its precision is,

$$\Big(\frac{\tau_{\theta,t} + \tau_\epsilon}{\tau_\epsilon}\Big)^2 \Big(\frac{\tau_{\theta,t}\tau_\epsilon}{\tau_{\theta,t} + \tau_\epsilon}\Big) = \frac{\tau_{\theta,t}(\tau_{\theta,t} + \tau_\epsilon)}{\tau_\epsilon}, \tag{1.47}$$

as the first term notes $\theta - \bar{\theta}(t) + \epsilon_t$ is scaled by $\frac{\tau_\epsilon}{\tau_\epsilon + \tau_{\theta,t}}$. Hence, the precision is scaled by the reciprocal of such a constant squared and the second term is the precision of $\theta - \bar{\theta}(t) + \epsilon_t$. Therefore,

$$\hat{v}_t = \frac{a_t - \beta_t \bar{\theta}(t)}{\alpha_t} \frac{\tau_{\theta,t}(\tau_{\theta,t} + \tau_\epsilon)}{\tau_{\theta,t}(\tau_{\theta,t} + \tau_\epsilon) + \tau_{v_{\text{ind}}}\tau_\epsilon} := \frac{a_t - \beta_t \bar{\theta}(t)}{\alpha_t} \tilde{\tau}_t. \tag{1.48}$$

[33] Mathematica code available upon request.

Given $\hat{v}_t$, one can compute the first-order condition for player t:

$$\begin{aligned}
a_t - \mathbf{E}(\theta) - v_t + \kappa \frac{(a_t - \beta_t \bar{\theta}(t))\tilde{\tau}_t^2}{\alpha_t^2} &= 0 \\
\iff a_t(1 + \frac{\kappa \tilde{\tau}_t^2}{\alpha_t^2}) &= \mathbf{E}(\theta) + v_t + \kappa \frac{\beta_t \bar{\theta}(t) \tilde{\tau}_t^2}{\alpha_t^2} \\
&= \bar{\theta}(t) + \frac{\mathbf{E}(s_t - \bar{\theta}(t))}{\tau_{\theta,t} + \tau_\epsilon} + v_t + \kappa \frac{\beta_t \bar{\theta}(t) \tilde{\tau}_t^2}{\alpha_t^2} \\
&= \bar{\theta}(t)(1 + \kappa \frac{\beta_t \tilde{\tau}_t^2}{\alpha_t^2}) + v_t + \mathbf{E}(\theta) - \bar{\theta}(t). \qquad (1.49)
\end{aligned}$$

Thus,

$$a_t = \bar{\theta}(t) \frac{1 + \kappa \frac{\beta_t \tilde{\tau}_t^2}{\alpha_t^2}}{1 + \frac{\kappa \tilde{\tau}_t^2}{\alpha_t^2}} + \frac{v_t + \mathbf{E}(\theta) - \bar{\theta}(t)}{1 + \frac{\kappa \tilde{\tau}_t^2}{\alpha_t^2}}. \qquad (1.50)$$

Finally, the conjecture must be correct in equilibrium implying,

$$\beta_t = \frac{1 + \kappa \frac{\beta_t \tilde{\tau}_t^2}{\alpha_t^2}}{1 + \frac{\kappa \tilde{\tau}_t^2}{\alpha_t^2}} \iff \beta_t = 1 \qquad (1.51)$$

$$\alpha_t = \frac{1}{1 + \frac{\kappa \tilde{\tau}_t^2}{\alpha_t^2}} \iff \alpha_t^2 - \alpha_t + \kappa \tilde{\tau}_t^2 = 0. \qquad (1.52)$$

Given this decision rule, an equilibrium with revelation exists if and only if $\kappa \tilde{\tau}_t \leq 1/4$. Further, $\tilde{\tau}_t$ is monotone in $\tau_{\theta,t}$ which implies that social learning about fundamentals occurs if and only if $\kappa \leq \kappa^{\text{c.k.}}$. When $\kappa > \kappa^{\text{c.k.}}$ the players switch to a pooling equilibrium when $\kappa \tilde{\tau}_t < 1/4$. As a result, fixing any prior beliefs, the limit of $\tilde{\tau}_t$ is monotone decreasing in κ, implying the limit of $\tau_{\theta,t}(\kappa)$ is monotone decreasing in κ.

I now prove the comparative statics regarding the asymptotic adaptation loss and utility. When $\kappa \leq \kappa^{\text{c.k.}}$, the asymptotic adaptation loss and utility converges to that of the common knowledge benchmark where the comparative statics with respect to κ were already established. When $\kappa > \kappa^{\text{c.k.}}$ the adaptation loss, in the limit, is equal to,

$$\mathbf{E}(\theta + v_t - \bar{\theta}(t))^2 = \mathbf{E}(v_t^2) + \frac{1}{\tau_{\theta,t}}. \qquad (1.53)$$

Further, given the comparative static with respect to κ of $\tau_\theta(\kappa)$, it follows that the adaptation loss is decreasing in κ. Finally, the asymptotic utility is equal to the asymptotic adaptation loss plus the conformity loss, and the conformity loss is mechanically decreasing in κ which proves the result. □

Proof of Proposition 4. The proof of the first statement is a direct consequence of the proof of statement 2 of Proposition 2.

The proof of the second statement can be seen by the third statement of Proposition 2 when n is finite. If n is infinite, then the proof follows from either of the uni-demnsional uncertainty analyses in Proposition 3 when $\rho = 0$ or Proposition 7 when $\delta = 0$.

The proof of the third statement contains two parts. The proof that a pooling equilibrium may be fragile to a peer-oriented intervention is provided in Figure 1 which shows that an increase in $\tau_{\mu,t}$ may break a pool. The proof that an increase in $\tau_{\theta,t}$ never breaks a pooling equilibrium can be seen by noting that when there is no correlation in the beliefs of θ and μ, Equation (1.8) reduces to

$$a_t(1 + \kappa\alpha^2) = \gamma_t \mathbf{E}\big(\theta \mid \theta(t), s_t\big) + v_t + \kappa\alpha\beta + \kappa\alpha\mathbf{E}\big(\mu \mid \mu(t), v_t\big), \tag{1.54}$$

due to independence. One can now see that uncertainty about θ decreases $\tilde{s}(\alpha, \kappa, \theta(t), \mu(t))$ in Equation (1.10). Finally, as argued in Lemma 3 as $\alpha \to \infty$ the left-hand side is greater than the right-hand side, hence the necessary and sufficient condition is whether there exists an α such that the left-hand side is less than the right-hand side. As $\tilde{s}(\alpha, \kappa, \theta(t), \mu(t))$ uniformly decreases when uncertainty about θ is decreased, one can see that a decrease in uncertainty about θ will never break a pooling equilibrium. □

Proof of Proposition 5. The players pool on $a^* = \mathbf{E}(\theta) + \mathbf{E}(\mu)$ where both expectations are derived from the decision of player 1. Player 1 chooses a_1 as follows:

$$a_1 = \lambda_\theta \mathbf{E}(\theta|s_1) + \lambda_v \mathbf{E}(\mu|v_1) \tag{1.55}$$

$$= \lambda_\theta\big(\theta + N(0, 1 + \tau_\theta)\big) + \lambda_v\big(\mu + N(0, 1 + \tau_\mu)\big). \tag{1.56}$$

As a result, a^* is a linear function of a_1. Now consider an individual-oriented intervention where θ is revealed. One must compute $\mathbf{E}(\mu|a_1, \theta)$ by noting:

$$\frac{a_1 - \lambda_\theta \theta}{\lambda_v} = \mu + N(0, 1 + \tau_\mu) + \frac{\lambda_\theta}{\lambda_v} N(0, 1 + \tau_\theta). \tag{1.57}$$

As a result,

$$\mathbf{E}(\mu \mid a_1) = \frac{\frac{a_1 - \lambda_\theta\theta}{\lambda_v}\left(\frac{1}{1+\tau_\mu} + (\frac{\lambda_\theta}{\lambda_v})^2 \frac{1}{1+\tau_\theta}\right)^{-1}}{\tau_\mu + \left(\frac{1}{1+\tau_\mu} + (\frac{\lambda_\theta}{\lambda_v})^2 \frac{1}{1+\tau_\theta}\right)^{-1}} = \frac{a_1 - \lambda_\theta\theta}{\lambda_v} \frac{1}{1 + \tau_\mu\left(\frac{1}{1+\tau_\mu} + (\frac{\lambda_\theta}{\lambda_v})^2 \frac{1}{1+\tau_\theta}\right)}. \tag{1.58}$$

Simplifying the posterior expectation of μ produces the result in the proposition. Further, a symmetric calculation occurs when considering an individual-oriented intervention. □

Proof of Proposition 6. Note that when θ is common knowledge the condition for the existence of a revealing equilibrium reduces to a function $\kappa^*(\tau_{\mu,t})$ whereby a decision rule with revelation exists if and only if $\kappa < \kappa^*(\tau_{\mu,t})$ (for a proof, set $\delta = 0$ in the characterization in Proposition 7).

The threshold κ^* is determined by the value of κ that binds Equation (1.18) when $\tau_{\mu,t} = \max_{t'} \tau_{\mu,t'} := \tau^*$ in an equilibrium where every period involves revelation. As $\tau_{\mu,t} < \tau_\mu/\rho^2$ and $\rho > 0$, then $\kappa^* > \kappa^{\text{c.k.}}$.

Statement (1) and (3) follow immediately from the criterion for when the decision rule in period t involves revelation.

Finally, statement (2) follows from noting that when $\kappa > \kappa^*$, the equilibrium must have infinitely many periods of pooling. If the equilibrium had finitely many periods of pooling, then consider the final period of pooling. After this period, there will exist a period where the precision is τ^*. Further, as $\kappa > \kappa^*$, the decision rule must involve revelation in this period deriving a contradiction. Similarly, if there are only finitely many periods of revelation, then after all the periods of pooling, $\tau_{\mu,t} \to \tau_\mu$. However, at this point, the decision rule must involve revelation. Finally, that there are infinitely periods of pooling and infinitely many periods of revelation implies $\tau_{\mu,t}$ is not Cauchy and thus does not converge. □

Proof of Proposition 7. As in the proof of Proposition 3, I will denote by $\tau_{v_{\text{ind}}}$ the precision of ν_t which was normalized to one, where ν_t was defined to satisfy $v_t := \mu + \nu_t$. Further, sufficient statistics for the decision rule in period t are the current mean and precision of $\mu(t)$ which are denoted as $\bar{\mu}(t)$ and $\tau_{\mu,t}$. I will now consider whether the decision rule for player i can involve revelation in period $t(i)$, which will be denoted as t for the rest of the proof. A linear equilibrium with revelation, when it exists, is of the form,

$$a_t = \alpha_t v_i + \beta_t. \tag{1.59}$$

The public perception of v_i given a_t is equal to $(a_t - \beta_t)/\alpha_t$. Therefore, one can write the utility as a function of a_t as follows,

$$\begin{aligned} u_i(a_t) = &- (a_t - v_i)^2 - \kappa\Big(\frac{a_t - \beta_t}{\alpha_t} - \frac{\bar{\mu}(t)\tau_{\mu,t} + v_i\tau_{v_{\text{ind}}}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}\Big)^2 \\ &- \frac{\delta}{1-\delta}\big(\frac{a_t - \beta_t}{\alpha_t} - v_i\big)^2 - \frac{\delta}{1-\delta}\kappa\Big(\frac{a_t - \beta_t}{\alpha_t} - \frac{\bar{\mu}(t)\tau_{\mu,t} + v_i\tau_{v_{\text{ind}}}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}\Big)^2, \end{aligned} \tag{1.60}$$

where the first line denotes the payoffs in period t given player i's perception of μ and the second line denotes the continuation payoff in all subsequent periods. One can now take a first-order condition to generate:

$$0 = a_t - v_i + \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\Big(\frac{a_t - \beta_t}{\alpha_t} - \frac{\bar{\mu}(t)\tau_{\mu,t} + v_i\tau_{v_{\text{ind}}}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}\Big) + \alpha_t\frac{\delta}{1-\delta}\big(\frac{a_t - \beta_t}{\alpha_t} - v_i\big). \tag{1.61}$$

Finally, in equilibrium, the beliefs are correct which implies $v_t = (a_t - \beta_t)/\alpha_t$. Simplifying the remaining terms implies:

$$\begin{aligned} 0 &= a_t - v_i + \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\Big(v_i - \frac{\bar{\mu}(t)\tau_{\mu,t} + v_i\tau_{v_{\text{ind}}}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}\Big) \\ &= a_t - v_i + \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\frac{v_i\tau_{\mu,t} - \bar{\mu}(t)\tau_{\mu,t}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}} \\ \iff a_t &= v_i\big(1 - \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\frac{\tau_{\mu,t}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}\big) + \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\frac{\tau_{\mu,t}\bar{\mu}(t)}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}}. \end{aligned} \tag{1.62}$$

In equilibrium, α_t must equal the coefficient on v_i in the decision rule:

$$\alpha_t = 1 - \frac{\kappa}{\alpha_t}\frac{1}{1-\delta}\frac{\tau_{\mu,t}}{\tau_{\mu,t} + \tau_{v_{\text{ind}}}} \iff \alpha_t = \frac{1 \pm \sqrt{1 - 4\kappa\frac{1}{1-\delta}\frac{\tau_{\mu,t}}{\tau_{\mu,t}+\tau_{v_{\text{ind}}}}}}{2}. \tag{1.63}$$

Further, note that given α_t, the solution to β_t is uniquely determined by Equation (1.62). Hence, a solution exists if and only if the condition in the text holds because this condition corresponds to the term in the square root being non-negative. □

Proof of Lemma 4. This is precisely Theorem 3 in Bernheim (1994). □

Proof of Proposition 8. The proof of this Proposition is a direct consequence of Proposition 11 in Appendix B. □

Proof of Proposition 9. First note that by an identical argument to Bernheim (1994) all equilibria must be central pooling. Given a central pooling equilibrium, Equation (1.24) characterizes the equilibrium outside of the central pool.

Let us begin showing that if κ exceeds the threshold in the proposition no solution to the differential equation exists for any initial condition, and, as a direct consequence, no equilibrium with revelation exists. To do so, one can differentiate Equation (1.24) and determine,

$$a_t''(v_t) = \frac{-\kappa\big(v_t - g_t(v_t)\big)\big(1 - a_t'(v_t)\big)}{\big(v_t - a_t(v_t)\big)^2} + \frac{\kappa\big(1 - g_t(v_t)\big)}{v_t - a_t(v_t)}$$

$$\iff \big(v_t - g_t(v_t)\big)a_t''(v_t) = a_t'(v_t)\left(1 - g_t'(v_t) - \frac{a_t'(v_t)\big(1 - a_t'(v_t)\big)}{\kappa}\right). \tag{1.64}$$

Further, for any value $x \geq 0$, $x(1-x)/\kappa \leq (1/4)/\kappa$ and given the condition on $\sup_x g_t'(x)$, one can show that for any κ greater than $1/4(1 - s_t)$ there exists an $\epsilon_3 > 0$ such that,

$$\big(v_t - g_t(v_t)\big)a_t''(v_t) \geq a_t'(v_t)\epsilon_3 \iff a_t''(v_t) \geq \frac{a_t'(v_t)\epsilon_3}{v_t}, \tag{1.65}$$

where the final inequality comes from noting that $g_t(v_t) \geq 0$ for positive values of v_t. Let $a_0(v_t)$ be the decision rule that binds differential inequality in Equation (1.65) and satisfies $a_t(\bar{v}_t) = a_0(\bar{v}_t)$, where $\bar{v}_t$ denotes the supremum of the central pool.

One can use the Picard-Lindelöf Theorem to show $a_0(v_t)$ has a unique solution up to this initial condition. This solution satisfies

$$a_0(v_t) = c_1 v_t^{1+\epsilon_3} + c_2, \tag{1.66}$$

where $c_1 > 0$. However, one can show that

$$a_t(v_t) = a_t(\bar{v}_t) + \int\int a_t''(v_t) \geq a_0(\bar{v}_t) + \int\int a_0''(v_t) = c_1 v_t^{1+\epsilon_3} + c_2. \tag{1.67}$$

This inequality implies that $a_t(v_t) > v_t$ for a positive value of v_t which is a contradiction because the players always have an incentive to choose a mildly more conforming decision and receive both a better adaptation loss and a better conformity loss.

Now I will show that if κ is less than the condition provided in Proposition 9 an equilibrium with full revelation exists. To do so, first note that there always exists a v_t such that $v_t = g_t(v_t)$, by Bayes Plausability. To see why, note that if there existed two values v_t' and v_t'' such that

$v_t' > g_t(v_t')$ and $v_t'' < g_t(v_t'')$ then by continuity there exists a value v_t where $v_t = g_t(v_t)$. So for no solution to exist, up to symmetry, it must be the case that $v_t' > g_t(v_t')$ for all values of v_t. By monotonicity of the integral, $\mathbf{E}(v_t) > \mathbf{E}\big(g_t(v_t')\big) = \mathbf{E}(\mu)$ which is a contradiction of Bayes Plausability. As a similar argument could be made if instead $v_t' < g_t(v_t')$ for all values, there must exist a value v_t^* such that $v_t^* = g_t(v_t^*)$. I will now show that if κ satisfies the condition in the proposition there exists an equilibrium with full revelation where $a_t(v_t^*) = v_t^*$. To do so, I will use Carathéodory's existence theorem.

To be able to apply this theorem to Equation (1.24) the decision rule, $a_t(v_t)$ must never cross v_t except for v_t^* so that the implicit function in Equation (1.24) is continuous on its domain. I will analyze the differential equation to the right of v_t^* and a symmetric analysis occurs to the left. A sufficient condition for $a_t(v_t) < v_t$ is that $a_t'(v_t) \leq 1/2$ for all $v_t \geq v_t^*$. Taking Equation (1.24), this condition can be stated as

$$\begin{aligned} & 2\kappa\big(v_t - g_t(v_t)\big) \leq \frac{1}{2}\big(v_t - a_t(v_t)\big) \\ \Longleftarrow\ & 2\kappa\big(v_t - g_t(v_t)\big) \leq \frac{1}{2}\big(v_t - v_t^*\big) \\ \Longleftrightarrow\ & \kappa \leq \frac{1}{4}\frac{1}{1 - \frac{g_t(v_t) - v_t^*}{v_t - v_t^*}}. \end{aligned} \tag{1.68}$$

Finally, the expression in the denominator of Equation (1.68) can be simplified as follows to generate the condition in the proposition:

$$\frac{g_t(v_t) - v_t^*}{v_t - v_t^*} \geq \frac{g_t(v_t^*) + i_t(v_t - v_t^*) - v_t^*}{v_t - v_t^*} = i_t. \tag{1.69}$$

Applying this inequality to Equation (1.68) finishes the proof. □

9 Appendix B: Online

This section details extensions then analyzes the equilibria that satisfy D1 and Pareto-Optimality. I begin this analysis when θ, μ are common knowledge. Such an analysis holds for a general class of distributions. Next, I detail how one can extend the results to allow for learning about these two random variables.

9.1 Extensions

Let us begin with an analysis with conformist preferences where individuals want their perceived preference type to be c units higher than the average in the population. Further, I consider the environment of Section 3 without uncertainty allowing us to drop any time-dependence. A player with preference type v's utility is,

$$-(a - v)^2 - \kappa\big(\hat{v}(a) + c\big)^2. \tag{1.70}$$

Proposition 10 (Shifted Preferences)
Suppose the preferences of the players are as in Equation (1.70), then in the linear equilibrium where $\hat{v}(a) = \alpha \cdot a + \beta$, α is independent of c.

This result shows that while choosing a value $c \neq 0$ will lead to different equilibrium decisions, the level of adaptiveness remains the same. This result can easily be generalized to show that the results regarding social learning are not sensitive to this exact parametrization of conformity concerns.

Proof of Proposition 10. One can conjecture equilibrium beliefs of $\hat{v}(a) = \alpha a + \beta$. Given these conjectures the first-order condition of a player is,

$$a - v + \kappa\alpha(\alpha a + \beta + c) = 0 \iff v = (1 + \kappa\alpha^2)a + \kappa\alpha(\beta + c). \tag{1.71}$$

In equilibrium the conjectures are correct which implies,

$$\alpha = 1 + \kappa\alpha^2 \text{ and } \beta = \kappa\alpha(\beta + c). \tag{1.72}$$

One can note the solution to α is independent of c, and if a solution for α exists, there will always exist a solution for β. □

9.2 Sufficiency of Linear Equilibria

Throughout this subsection the only distributional assumptions that are needed about v are that (i) the distribution admits a continuous density and (ii) the support of the distribution is the real line.

Let us first begin with equilibria that satisfy D1. As detailed in the text, Bernheim (1994) implies that Equation (1.8) must hold outside the central pool. Given this result, one can show the following Proposition.

Proposition 11 (**Linear Equilibria**)
Any equilibrium that satisfies D1 has an empty central pool. On either side of the empty central pool, the decision rule is characterized by $a = \alpha v$ where α takes one of two possible values: $\frac{1-\sqrt{1-4\kappa}}{2}, \frac{1+\sqrt{1-4\kappa}}{2}$ which exist if and only if $\kappa \leq 1/4$, where κ denotes the weight on conformity. However, there always exists a fully-pooling decision rule in which all types take the same decision.

Proof of Proposition 11. As mentioned in the text, there always exists the fully pooling decision rule where the central pool is the entire domain. Hence, let us consider central pooling equilibria with central pools that are not the entire domain. Without loss of generality let us assume $\bar{v} < \infty$, and an identical characterization follows if $\underline{v} > -\infty$. The first-order condition implies:

$$a - v + \kappa\hat{v}(a)\hat{v}'(a) = 0. \tag{1.73}$$

Substituting $a = x$ and $ay(x) = \hat{v}(x)$ to Equation (1.73) yields,

$$0 = x - xy(x) + \kappa xy(x)\big(y(x) + x\dot{y}(x)\big) \iff \dot{y}(x) = \frac{-1 + y(x) - \kappa y(x)^2}{\kappa xy(x)}. \tag{1.74}$$

This simplification is well defined because the denominator is non-zero for all x outside the central pool. Next, if the equilibrium decision rule is linear, then $\dot{y}(x) = 0 \; \forall \; x$. Further, if $\dot{y}(x) = 0$ for some x, one can calculate the $\ddot{y}(\cdot)$ as,

$$\begin{aligned}\ddot{y}(x) = \frac{-y(x)^2(kx\dot{y}(x)+1) + ky(x)^3 + x\dot{y}(x) + y(x)}{kx^2y(x)^2} &= \frac{-y(x)^2 + ky(x)^3 + y(x)}{kx^2y(x)^2} \\ &= -\frac{\dot{y}(x)}{x} = 0. \end{aligned} \tag{1.75}$$

Equation (1.75) implies if $\dot{y}(x) = 0$ for any x, then $\dot{y}(x) = 0 \; \forall \; x$. Thus any non-linear solution satisfies the following integral expression derived by re-arranging Equation (1.74):

$$\int \frac{\dot{y}(x)\kappa y(x)}{-1 + y(x) - \kappa y(x)^2} = \int \frac{-1}{x} \iff$$

$$\frac{-\log(1 - y(x) + \kappa y(x)^2)}{2} + \frac{\log(1 + \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}) - \log(1 - \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}})}{2\sqrt{1-4\kappa}} + c = -\log(x), \tag{1.76}$$

where c is the constant of integration. If $\kappa = 1/4$, then Equation (1.76) is ill-defined and thus all solutions are necessarily linear. Further, assuming $\kappa < 1/4$, any non-linear solution satisfies the following equalities:

$$\log\left(\sqrt{1 - y(x) + \kappa y(x)^2} \cdot \Big(\frac{1 - \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}{1 + \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}\Big)^{\frac{1}{2\sqrt{1-4\kappa}}}\right) = \log(e^c x) \tag{1.77}$$

$$\iff \sqrt{1 - y(x) + \kappa y(x)^2} \cdot \Big(\frac{1 - \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}{1 + \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}\Big)^{\frac{1}{2\sqrt{1-4\kappa}}} = e^c x \tag{1.78}$$

$$\iff (1 - y(x) + \kappa y(x)^2)\Big(\frac{1 - \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}{1 + \frac{1+2\kappa y(x)}{\sqrt{1-4\kappa}}}\Big)^{\frac{1}{\sqrt{1-4\kappa}}} = x^2 e^c. \tag{1.79}$$

Recall that the map $a(v)$ is defined on $(\bar{v}, \infty)$ and thus one can consider the limit of the above equation as $x \to \infty$. The right-hand side of the equation diverges as $x \to \infty$ implying the left-hand side must also diverge. However, if $y(x)$ is bounded, then both the first and second terms of the left-hand side will remain bounded. Therefore, one can assume that $y(x)$ is unbounded. By definition, then $v/a(v)$ is unbounded. However, if $v/a(v)$ is unbounded, then there must exist a player for which ϵ less conformity is preferred.

Thus no non-linear solution exists when $\kappa < 1/4$. If instead $\kappa > 1/4$, then the simplification to Equation (1.76) is

$$\frac{1}{2}\log\left(1 - y(x) + \kappa y(x)^2\right) + \frac{\tan^{-1}\left(\frac{2ky(x)-1}{\sqrt{4k-1}}\right)}{\sqrt{4k-1}} + c = \log(x). \tag{1.80}$$

As argued above, $y(x)$ must be bounded. However, this implies the left-hand side of the above equation cannot diverge, but the right-hand side necessarily diverges. As no non-linear solution exists for any κ, the only solutions to the differential equation outside the central pool are linear. I will now show that the central pool is empty.

Equation (1.3) shows that the linear equilibria must satisfy:

$$\alpha(1-\alpha) = \kappa \iff \alpha = \frac{1 \pm \sqrt{1-4\kappa}}{2}. \tag{1.81}$$

Thus for a central pooling equilibrium to exist one must find $a^*, \underline{v}, \bar{v}$ such that the player with preference type $\bar{v}$ is indifferent between a^* and $\alpha\bar{v}$. I consider the case where $a^* \leq 0$ and an identical argument holds if $a^* \geq 0$.[34] Note that the conformity loss following a^* is at best zero. Fixing the slope of the linear decision rule $a(v) = \alpha v$, then the following inequality must hold:

$$-(\bar{v} - a^*)^2 \geq -(1-\alpha)^2\bar{v}^2 - \kappa\bar{v}^2, \tag{1.82}$$

where the left-hand side is an upper bound on the utility in the central pool and the right-hand side is the utility in the linear equilibrium. One can notice that $\bar{v}^2$ cancels out and that for the solutions to α such an inequality never holds. Given that such an inequality cannot hold, the central pool must be empty in any equilibrium with revelation. □

While Proposition 11 implies that multiple equilibria with revelation exist, only one equilibrium is Pareto optimal. Note that the conformity loss is identical across the equilibria by Bayes Plausibility. Hence, the equilibrium that maximizes utility is the one that minimizes the adaptation loss. In the class of linear equilibria, this corresponds to the equilibrium with the highest slope (as all the slopes are strictly less than one). These refinements select a unique equilibrium as stated in Proposition 1 in the text.

9.3 Generalized Distributions

Subsection 6.3 analyzed the impact of conformity concerns on social learning about μ when θ was common knowledge for a general class of distributions. This subsection will analyze the impact of conformity concerns on social learning about θ when μ is common knowledge for a general class of distributions.

Suppose each player has a preference type $v_t \sim F$ where F admits a continuous and differentiable density $f(\cdot)$ with support equal to the real line and mean normalized to zero. Further, denote by ω_t the player t's posterior beliefs about θ in period t after observing s_t. I will assume the distribution of ω_t admits a continuous and differentiable density. The only decision relevant information for the player is $v_t + \omega_t := \tilde{v}_t$ and it is without loss to consider equilibria that only condition on the sum. Further, by an identical argument to Bernheim (1994), the only equilibria satisfying D1 are central pooling equilibria as a function of $\tilde{v}_t$.

Let us detail the conformity concerns in this general framework. As previously noted, outside the central pool the decision rule will be strictly increasing in $\tilde{v}_t$ which implies the existence of an inverse. Denote by x an arbitrary decision where $a^{-1}(x) = \tilde{v}_t$, then the conformity loss upon choosing the decision x is

$$\kappa\Big(\tilde{v}_t - h_1\big(\tilde{v}_t\big)\Big)^2 + \kappa h_2\big(\tilde{v}_t\big). \tag{1.83}$$

[34] One can use $\underline{v}$ analogously to $\bar{v}$. If $\underline{v}$ was negative infinity, then the equilibrium could not satisfy D1 as no beliefs satisfying D1 following $a^* - \epsilon$ would prevent a player with an arbitrarily negative preference type from deviating.

Note that a player has concerns over the second moment of their perceived preference type. This second moment has two components the mean and the variance. The mean of the perceived preference type is simply $\tilde{v}_t$ less $h_1\big(\tilde{v}_t\big)$, where $h_1(t)$ is defined such to make such an equality hold. Without uncertainty, $h_1(\tilde{v}_t) = 0$, however, in general $h_1(\tilde{v}_t) \neq 0$ because a high value of $\tilde{v}_t$ may stem from either a high value of v_t or a high value of ω_t. Further, $h_2(\tilde{v}_t)$ represents the impact of the variance of players with different preference types choosing the same decision on the conformity loss. Formally, $h_1\big(\tilde{v}_t\big) = \tilde{v}_t - \mathbf{E}(v_t|v_t + \omega_t = \tilde{v}_t)$ and $h_2\big(v_t|v_t + \omega_t = \tilde{v}_t)\big) = \mathbf{Var}(v_t|v_t + \omega_t = \tilde{v}_t)$.

Rather than dealing with the distributions of v_t and ω_t, I will work with $h_1(\cdot)$ and $h_2(\cdot)$. Differentiating the utility function in Equation (1.83) implies

$$0 = 2(a_t - \tilde{v}_t) + 2\kappa\Big(\tilde{v}_t - h_1(\tilde{v}_t)\Big)\frac{1}{a'(\tilde{v}_t)}\big(1 - h_1'(\tilde{v}_t)\big) + \frac{\kappa h_2'(\tilde{v}_t)}{a'(\tilde{v}_t)}$$

$$\iff a'(v_t) = \frac{\kappa\Big(\big(\tilde{v}_t - h_1(\tilde{v}_t)\big)\big(1 - h_1'(\tilde{v}_t)\big) + \frac{1}{2}h_2'(\tilde{v}_t)\Big)}{a_t - v_t}. \tag{1.84}$$

However, such an expression has a similar formulation to Equation (1.24) in Proposition 9. One can replace $g_t(v_t)$ in Equation (1.24) with the expression in Equation (1.84).

9.4 Generalized Results With Social Learning

The first two subsubsections show that linearity is implied by D1 when either μ or θ are common knowledge, respectively. The third subsubsection discusses the difficulty with both dimensions of uncertainty. The final subsubsection discusses the learning outcomes if $\gamma_t = 1$ in all periods.

9.4.1 Learning About the Fundamental State:

In this extension, I assume μ is common knowledge, and without loss of generality equal to zero. By induction, it suffices to consider the incentives of an arbitrary player and thus drop any time dependence. Here, $\theta \sim N(0, \tau_\theta)$ and $v \sim N(0, \tau_\mu)$, resulting in the following utility for player 1:

$$-(a - \theta - v)^2 - \kappa \int \phi(b, a) b^2 db. \tag{1.85}$$

Proposition 12 (**Linear Equilibria without Population Uncertainty**)
The only equilibria with revelation that satisfy D1 are the linear equilibria.

Proof of Proposition 12. As $\theta + v$ is the only decision-relevant variable for the players it is without loss to consider decision rules such that a is a function of the sum alone. Note that in any equilibrium in which a is a function of $\theta + v$, an identical proof to Bernheim (1994) can be used to show that D1 implies central pooling. Finally, note that outside the central pool if $a(x) = a^*$, then the equilibrium inference given a^* is a Gaussian random variable with mean $\frac{\tau_\theta}{\tau_\theta + \tau_\mu}x$ and a fixed variance. Given the quadratic loss framework the variance is irrelevant in determining which decision to take outside the central pool. Hence, the decision outside the central pool is chosen to maximize,

$$-(a-v-\theta)^2-\kappa\frac{\tau_\theta}{\tau_\theta+\tau_\mu}\tilde{a}^{-1}(a)^2, \tag{1.86}$$

given a conjectured decision rule $\tilde{a}(\cdot)$.

However, note that this is an identical framework as the framework assuming θ is common knowledge if one defines $\tilde{\kappa}=\kappa\frac{\tau_\theta}{\tau_\theta+\tau_\mu}$. Given Proposition 11, outside the central pool the decision rule must be linear. All that remains to show is that the central pool is empty.

Suppose by contradiction, a non-empty central pool existed with pooling decision $a^*\leq 0$ and a pool determined by $\underline{v},\bar{v}$. In this equilibrium, the player with preference type $\bar{v}$ must be indifferent between $\alpha\bar{v}$ and a^*. As the linear equilibrium can be extended to be an equilibrium over the entire real line, the player with preference type $\bar{v}$ prefers choosing $\alpha\bar{v}$ where $\phi(b,\alpha\bar{v})\sim N(\frac{\tau_\theta}{\tau_\theta+\tau_\mu}\bar{v},\frac{1}{\tau_\theta+\tau_\mu})$ to choosing $a=0$ where $\phi(b,0)\sim N(0,\frac{1}{\tau_\theta+\tau_\mu})$. Finally, note that choosing the pooling decision incurs a worse adaptation loss than $a=0$ in the linear equilibrium (as $a^*\leq 0$), a worse conformity loss (as $a_1=0$ generates the ideal mean perceived preference type), and a worse variance as the equilibrium variance is higher in the pool. Finally, note if instead $a^*>0$, one can utilize an identical argument with $\underline{v}$ to derive a contradiction. □

9.4.2 Learning About the Preferences of Others:

Let us now consider an alternative environment in which $\theta=0$ is common knowledge, but there exists aggregate uncertainty over the average preference type of the population, μ. By induction, we can again drop any time dependence. Note that v is the only decision-relevant private information, implying a is a function of v alone. Moreover, a player's utility is:

$$-(a-v)^2-\kappa\int\Big(\int\phi(b,a)(b-\mu)^2db\Big)f(\mu\mid v)d\mu. \tag{1.87}$$

As in the proof of Proposition 3, I will denote by $\tau_{v_{\text{ind}}}$ the precision of v conditional on a realization of μ. I can now characterize the equilibria which satisfy D1.

Proposition 13 (**Linear Equilibria with State Uncertainty**)
The only equilibria that satisfy D1 off-path are the linear equilibria.

Proof of Proposition 13. Again, using a proof identical to Bernheim (1994), one can show that D1 implies central pooling. Outside the central pool, a player's preference type is fully revealed. Hence, the loss function outside the central pool is:

$$-(a-v)^2-\kappa\int\big(v^{-1}(a)-\mu\big)^2f(\mu\mid v)d\mu. \tag{1.88}$$

However a player's equilibrium perception of μ is a Gaussian distribution with a mean of $\frac{\tau_{v_{\text{ind}}}}{\tau_{v_{\text{ind}}}+\tau_\mu}v$ and a fixed variance. Given the quadratic-loss framework, the variance is additively separable from the decision rule and can be ignored when considering which decision to take outside the central pool. Let us recall the notation $\hat{v}(a)$ to denote the expected perceived preference type given decision a. Thus, the optimization reduces to choosing an a which maximizes,

$$-(a-v)^2 - \kappa\Big(\hat{v}(a) - \frac{\tau_{v_{\text{ind}}}}{\tau_{v_{\text{ind}}} + \tau_\mu} v\Big)^2. \tag{1.89}$$

One can differentiate to derive,

$$a - v + \kappa\Big(\hat{v}(a) - \frac{\tau_{v_{\text{ind}}}}{\tau_{v_{\text{ind}}} + \tau_\mu} v_1\Big)\hat{v}'(a) = 0. \tag{1.90}$$

Further, in equilibrium $\hat{v}(a) = v$, and thus,

$$a - v + \kappa\Big(v - \frac{\tau_{v_{\text{ind}}}}{\tau_{v_{\text{ind}}} + \tau_\mu} v\Big)v'(a) = 0. \tag{1.91}$$

However, this expression is identical to Equation (1.73) with a rescaling of κ. Thus the only differential equations that satisfy such a constraint are linear. Further, by an identical intuition to Proposition 11, the central pool must be empty in any equilibrium with revelation. These results imply any equilibrium with revelation that satisfies D1 is linear. □

9.4.3 Learning About Both the Preferences of Others and the Fundamental State:

This analysis is qualitatively different to the previous analysis. Now a_t is a function of two-dimensional private information: s_t, v_t. Thus, the equilibrium decision rule in general will be determined by a partial differential equation. Further, for any conjectured non-linear decision rule, $\phi(b, a_t)$ will be non-Gaussian. As one must integrate $\phi(b, a_t)$ to determine the conformity loss, one cannot solve (let alone, write down) the partial differential equation in closed form as was done for the differential equation in Proposition 11.

9.4.4 Unstable Confounded Learning Outcomes

Given the updating rule in Equation (1.35), there is a possibility that despite an equilibrium with revelation in every period that social learning fails. For instance, if $a_t = \lambda s_t + (1-\lambda)v_t$ in every period, the players' will never be able to disentangle θ and μ. The following proposition formalizes this intuition.

Lemma 5 *If there exists infinitely many periods of revelation, the variance matrix of the belief, Σ_t, either converges to the zero-matrix or Σ_t converges to*

$$\begin{pmatrix} c_1 & -\sqrt{c_1 c_2} \\ -\sqrt{c_1 c_2} & c_2 \end{pmatrix} \tag{1.92}$$

for $c_1, c_2 > 0$.

Proof of Lemma 5. Suppose that $\Sigma_t \not\to_p 0$, then Lemma 1 implies that $\theta(t) \not\to_p \theta$ and $\mu(t) \not\to_p \mu$. Further, for all t, a_t is a linear combination of s_t and v_t plus a constant. As a result, for all t a sufficient statistic for a_t is,

$$\lambda_t s_t + (1-\lambda_t)v_t, \tag{1.93}$$

for some constant $\lambda_t \in (0,1)$. If λ_t does not converge, then there exist two different convergent subsequences that converge to λ^1, λ^2, respectively. By the law of large numbers then the players learn $\lambda^1\theta + (1-\lambda^1)\mu$ and $\lambda^2\theta + (1-\lambda^2)\mu$. However, as $\lambda^1 \neq \lambda^2$ the players can separately learn θ and μ. Thus, because the beliefs do not converge, λ_t must converge.

If λ_t converges, then ϕ_t from Equation (1.35) necessarily converges to an interior value. Further, ρ_t is equal to the negative square root of $\tau_{2,t}^2/(\tau_{1,t}\tau_{3,t})$ as stated below:

$$\rho_t = -\sqrt{\frac{(\sum_{i=1}^t \sqrt{\phi_t(1-\phi_t)})^2}{(\sum_{i=1}^t \phi_t)(\sum_{i=1}^t 1-\phi_t)}}. \tag{1.94}$$

As ϕ_t converges, one can re-write ρ_t asymptotically as follows,

$$\rho_t = -\sqrt{\frac{t^2 \lim \phi_t(1-\phi_t) + o(t^2)}{t^2 \lim \phi_t(1-\phi_t) + o(t^2)}} \to -1. \tag{1.95}$$

Finally, the variance of θ and μ are each bounded and decreasing sequences, and thus both variances converge completing the proof. □

I will call the above such beliefs “confounded learning outcomes." In this model, such beliefs are not asymptotically stable because after a small perturbation (such as decreasing the variance of θ) the beliefs will never return to this confounded learning outcome. Generally for such outcomes to be stable, there must exist multiple preference types who have opposing preferences (cf. Smith and Sørensen, 2000). By introducing γ_t, this ensures that if the signaling equilibrium involves revelation in each period then λ_t does not converge.

Chapter 2

Multi-Project Collaborations

Charles Angelucci and Roi Orzach[1]

1 Introduction

In many settings, actors collaborate to experiment simultaneously in multiple domains. In buyer-supplier relationships, companies co-innovate in various product lines or geographies. In the pharmaceutical sector, an R&D alliance may combine resources across research areas. Inside firms, continuous improvement methods involve managers and workers collaborating to identify and implement improvements throughout the production process.

The success of these collaborations relies on keeping interests aligned, so that each party finds ongoing value in maintaining the partnership. In multi-domain collaborations, the ongoing value of continued participation is determined by the aggregate value across all domains of cooperation. This aggregate value—representing what parties stand to lose if cooperation ends—creates interdependencies across domains. For instance, a breakthrough in one domain will increase the parties' perceived value of the collaboration, mitigating opportunism in the other domains. As a result, parties must approach their joint experimentation in each domain of cooperation by balancing the domain-specific outcomes with the broader implications for the overall collaboration. This raises critical questions: How does multi-domain experimentation shape exploration and exploitation choices? How does the number of active domains evolve over time? Does starting with fewer domains foster

[1]Angelucci: MIT Sloan School of Management, 50 Memorial Drive, Cambridge MA 02142 (e-mail: cangeluc@mit.edu); Orzach: Department of Economics, MIT, 50 Memorial Drive, Cambridge MA 02142 (e-mail: orzach@mit.edu).
We thank Daniel Barron, Alexander Bloedel, Alessandro Bonatti, Renee Bowen, David Byrne, Yifan Dai, Joyee Deb, Wouter Dessein, Glenn Ellison, Dana Foarta, Guillaume Fréchette, George Giorgiadis, Robert Gibbons, Marina Halac, Johannes Horner, Matthias Lang, Alessandro Lizzeri, Jin Li, Rocco Macchiavello, Chiara Margaria, Simone Meraglia, Stephen Morris, Aroon Narayanan, Juan Ortner, Jacopo Perego, Hazhir Rahmandad, Daniel Rappaport, Patrick Rey, Kareen Rozen, Klaus M. Schmidt, Steven Tadelis, Michael Ting, Jean Tirole, Marta Troya Martinez, Jonathan Weinstein, Michael Whinston, and seminar audiences at the Oliver E. Williamson Seminar at Berkeley Haas, LMU Munich, the MIT Organizational Economics and Theory groups, the 2023 Montreal IO Conference, the Fall 2023 NBER Organizational Economics Workshop, the Organisational Economics and Leadership Workshop at Exeter University, the Paris School of Economics, the 2023 SIOE Conference, and the TSE Microeconomic Theory Workshop.

cooperation? When does experimentation expand to all domains?

To address these questions, we develop a model of multi-domain collaborative experimentation and use it to interpret key findings from the applied literature that studies settings such as those mentioned above. In our model, the number of domains is exogenous, and domains are technologically independent. Each period, in each domain, two players can choose to idle, exploit a known project, or explore a new one from an infinite set of potential projects. Cooperation on a project requires the participation of both players; working individually is not an option. Project benefits are time-invariant but initially uncertain, and they may be asymmetric across players. The benefits of a project are revealed in the first period of cooperation on that project. Moreover, all projects entail a constant fixed cost for the players, during both exploration and exploitation phases. As a result, players might be reluctant to collaborate in exploring projects if they expect that their individual benefit will not exceed this cost, and they may similarly be reluctant to collaborate in exploiting a project if their realized individual benefit falls below the cost. To align incentives, players can transfer money to each other. However, these transfers are voluntary, so any experimentation policy—a rule determining for each domain whether to idle, exploit a known project, or explore a new one—must be self-enforcing.

We focus on Subgame Perfect Equilibria (relational contracts) that maximize the players' discounted cumulative joint payoffs (their "surplus"). As a first benchmark, Proposition 1 examines the single-player scenario, providing a straightforward solution in which the optimal experimentation policy treats each domain independently: within each domain, exploration continues until a project's value exceeds a time-invariant threshold, after which permanent exploitation of this project is optimal. We refer to this optimal policy for the single-player scenario as the "first-best experimentation policy." Notably, this first-best policy would be optimal for the two players if all projects benefited them equally.

However, because experimentation in our setting requires both players' participation, asymmetric project benefits create incentive challenges. Our analysis centers on this scenario by assuming each project benefits only one player. The beneficiary's identity is revealed when players first cooperate on a project and is independently and identically distributed across projects. These asymmetric benefits create the key friction that may impede first-best experimentation, as implementing this policy requires credible promises of transfers between players. Such promises may lack credibility when players making transfers have insufficient continuation value in the collaboration. As mentioned above, since a player's continuation value equals the sum of continuation values across all domains of cooperation, experimentation choices in one domain affect all others.

In the spirit of Levin (2002, 2003), we show in Proposition 2 that (i) any optimal experimentation policy is governed solely by the value of the most profitable projects identified in each domain to date, and (ii) a single implementability constraint, dependent on only these values and the experimentation policy, fully captures all deviation temptations across players, domains, and transfers. These results imply that any experimentation policy satisfying this constraint can be implemented through a relational contract with appropriately designed transfers. As a result, the optimal experimentation policy is characterized by a multi-dimensional Bellman equation subject to the implementability constraint. However, unlike the single-player benchmark, this constraint precludes an index characterization of the optimal policy. Despite this challenge, we establish key properties of the players' optimal experimentation policy, offering insights into the dynamics of innovation-driven collaborations.

Since the first-best experimentation policy treats each domain independently, we can explicitly determine the conditions under which this policy is implementable through a relational contract and, consequently, chosen by the players. Proposition 3 provides a necessary and sufficient condition: the joint value of the most valuable projects identified in each domain must be sufficiently high to ensure that the collaboration's continuation value supports the implementation of the first-best policy. For low discount factors, this condition binds, implying that, in expectation, players transition to permanently exploiting the most valuable projects found in each domain later than if they could implement the first-best from the start (Corollary 1). In some cases, this transition never occurs, as discussed below. Moreover, this condition enables a complete characterization of optimal experimentation in our second benchmark: the single-domain case. Here, exploration continues until a project's value exceeds a fixed threshold—higher than in the single-agent case—after which permanent exploitation becomes optimal (Corollary 2).

Next, we analyze the second-best experimentation policy, which arises when the first-best policy is not implementable in the current period. We first examine the players' exploration and exploitation decisions, abstracting from the number of domains they engage in. Due to cross-domain interdependencies, the player's exploitation criterion becomes dynamic. Nonetheless, Proposition 4 shows that, unlike the first-best policy where explored projects are either permanently exploited or never used, the second-best policy is such that, with strictly positive probability, players (i) exploit projects temporarily or (ii) exploit previously unexplored projects rather than the most recently explored ones.

We then examine, under the second-best policy, the dynamics of the players' scope of experimentation—defined as the number of domains involving either exploration or exploitation in a given period—where m represents the exogenous maximum number of potential domains. We analyze both initial and terminal (asymptotic) scope of experimentation. We show that starting with limited scope—such as one domain instead of m—reduces players' initial deviation temptation by a factor of m. The potential for later scope expansions, if the continuation value increases, further mitigates initial deviation temptations. However, the continuation value increases only through exploration, and conducting one exploration (versus m) reduces these increases by a factor on the order of m. Proposition 5 shows that, although these opposing forces cannot generally be ranked, for large m, an initially limited scope allows implementation over a wider range of discount factors than immediate exploration in all domains. Moreover, the continuation value of the collaboration need not increase monotonically over time: for instance, a domain's continuation value decreases when players switch from exploration to exploitation. Thus, exploiting projects in some domains may create inefficiencies in others, including being permanently idle. Building on this observation, Proposition 6 shows that initially limited experimentation policies may never reach maximal—and thus efficient—scope asymptotically, and even policies starting with maximal scope may become permanently limited.

In Section 5, we examine how the potential scope of experimentation impacts its feasibility and profitability, drawing connections to the seminal work of Bernheim and Whinston (1990) on multilateral interactions. Further, we discuss extensions of the model included in the Online Appendix, in which the domains of cooperation are asymmetric or exhibit technological interdependencies.

Section 6 connects our main findings to the existing applied literature on buyer-supplier relationships and persistent productivity differences across firms. The buyer-supplier relationships literature stresses experimentation and credibility as critical factors for successful collaborations, and corroborates the prevalence of gradualism and strong path dependence. In addition, we argue that our framework provides novel insights into how managerial practices can generate productivity differences among seemingly similar firms.

The rest of the paper is structured as follows. Section 1.1 reviews the relevant theoretical literature. Section 2 presents the model. Section 3 characterizes the first-best experimentation policy. Section 4 provides the main analysis. Section 5 discusses various model extensions. Section 6 examines the applied literature in light of our theoretical findings. Section 7 concludes the paper.

1.1 Related Theoretical Literature

In this section, we review the theoretical literature related to our work. We postpone the discussion of the applied literature to Section 6.

Firstly, our research connects to the literature on multi-armed bandit problems (Robbins, 1952) and on optimal search (Lippman and McCall, 1976; Weitzman, 1979), contributing to the strand that examines strategic interactions.[2] Bolton and Harris (1999) and Keller et al. (2005) consider settings in which players free-ride on each others' experimentation (see Hörner et al., 2022, for more recent work on this topic). Further, Liu and Wong (2023) consider an environment in which players compete to explore alternatives. In Strulovici (2010), players vote between a safe and a risky arm, with its asymmetric benefits revealed over time through experimentation (see also Anesi and Bowen, 2021). Further, Albrecht et al. (2010) examine a search problem where a committee determines which project to exploit. Chan et al. (2018) and Reshidi et al. (2025) compare group and individual decision-making, examining the effects of static vs. sequential information acquisition and voting rules. In contrast to these papers, our setting allows for voluntary transfers and requires the players to cooperate for both the exploration and exploitation of projects. Most significantly, players experiment simultaneously across multiple domains.

Multi-domain experimentation poses analytical challenges. As noted in Bergemann and Välimäki (2008),"it is well known that [a Gittins] index characterization is not possible when the decision maker must or can select more than a single arm at each t," due to the optimality of recalling past projects.[3] With infinitely many ex ante identical projects, as we assume, project recall is absent and a Gittins index exists in the single-agent case (Bergemann and Välimäki, 2001). In our setting, cooperation between players is needed to experiment and project benefits are asymmetric. Therefore, an experimentation policy must form an

[2]Our setting resembles standard search problems by modeling many alternatives for players to explore. However, unlike typical search problems where rewards come only at the end from the best explored alternative, our model allows players to benefit each time they cooperate on a project, without settling on one. For this reason, we use the broader term "experimentation" rather than "search." Moreover, existing models of strategic experimentation with bandits often limit options to a few alternatives, like a risky and a safe project. We assume an infinite number of i.i.d. projects to eliminate aggregate uncertainty, making the dynamics driven purely by strategic factors.

[3]Bergemann and Välimäki (2008) go on to note that even if such an index existed, "it is normally impossible to obtain analytical solutions for the problem."

equilibrium and we demonstrate that this requirement generates rich dynamics—such as project recall, temporary exploitation, and idling on previously explored domains—even though these very dynamics complicate the problem by precluding a Gittins representation.

Secondly, this work relates to the literature on relational contracts (see e.g., Bull, 1987; Macleod and Malcomson, 1989; Baker et al., 1994, 2002a; Levin, 2002, 2003, for early contributions).[4] Halac (2014) studies a setting in which the value of the players' relationship increases exogenously with its duration, allowing for greater efficiency. In our setting, players' experimentation endogenously shapes the continuation value of their relationship, which may not increase monotonically over time. For instance, while exploration in any domain enhances the players' continuation value, exploitation diminishes it, potentially hindering experimentation in other domains.

A closely related paper to ours is Chassang (2010). In his model, the agent knows which arms are productive and which are not, while the principal, at the outset, cannot differentiate between the two. Without monetary incentives, incentivizing the agent to choose productive arms is accomplished by the threat of firing the agent following failures. This dynamic makes motivating exploration progressively expensive as more productive arms are identified. Should the relationship endure, it ultimately enters an "exploitation" phase and its value stops growing. In our model, the players are symmetrically informed about their environment, and the presence of transferable utility–apt for modeling firms–removes the need for inefficient on-path punishments. Yet, it generates dynamics similar to those observed in collaborations between and within firms (see Section 6.2).[5]

Finally, we contribute to the literature on gradualism in collaborations. Watson (1999, 2002) examine a setting in which players are uncertain regarding their counterpart' intentions—to either collaborate genuinely or take advantage of the other. They begin with low cooperation to mitigate the losses from defection. As the players become more optimistic, the collaboration grows. Collaborations involving trustworthy players achieve optimal cooperation, while those with untrustworthy players eventually fail. In our setting, the scope of players' experimentation can expand or contract over time due to the evolving continuation value of the relationship. Moreover, the two settings make opposite predictions about how the discount factor affects players' incentives to "start small.' In our setting, a higher discount factor reduces this need, whereas in the frameworks analyzed by Watson (1999, 2002) and the broader dynamic screening literature (e.g., Ely and Välimäki, 2003; Acharya and Ortner, 2022), a higher discount factor increases it, as separation becomes harder to achieve.

[4]Also at the intersection of the bandit and the relational contracting literatures, Urgun (2021) examines a scenario where a principal interacts with multiple agents whose publicly-observable types depend on the contracting history.

[5]Introducing money in Chassang (2010), where information asymmetry plays a central role, would make the value of the relationship constant on path. We further discuss Chassang (2010) in Footnote 14 and Section 6.2. For a setting similar to Chassang (2010) but with imperfect transfers and uncertainty about the value of the relationship, see Venables (2013). For work on experimentation in principal-agent settings with commitment, see Halac et al. (2016) and Ide (2024).

2 The Setup

Two players, with a discount factor $\delta < 1$ and zero per-period outside options, have the opportunity to interact over multiple time periods $t = 1, 2, \ldots$. Their interaction spans m exogenously fixed domains—such as distinct geographical markets or product categories in a buyer-supplier relationship—where each domain j contains a countably infinite set of projects $\mathcal{P}_j$. The union of all these sets forms the total set of projects, denoted as $\mathcal{P} = \cup_j \mathcal{P}_j$, where each project within $\mathcal{P}$ is indexed by p. In each period t, and for each domain j, each player $i = 1, 2$ chooses up to one project from the set $\mathcal{P}_j$. The finite set of projects chosen by player i in period t is denoted by P_i^t. The players cooperate on the set of projects $\mathbf{P}^t = P_1^t \cap P_2^t$, following a unanimity rule, and cannot work individually on projects not included in $\mathbf{P}^t$, as both players possess indispensable and complementary assets or skills. The cardinality of this set, $|\mathbf{P}^t| \leq m$, is referred to as the scope of the players' experimentation in period t.

Each project in $\mathbf{P}^t$ costs $c > 0$ for each player and has initially unknown time-invariant value $v_p \in \mathcal{R}$, which is publicly observed after the first cooperation. We assume that for each project, a single player receives the entire value v_p of the project.[6] The identity of any project's beneficiary is, however, initially unknown and we denote it by $x_p \in \{1, 2\}$. Both v_p and x_p are each i.i.d. across projects and domains, making all domains ex ante identical. We denote by $\alpha \in \left[\frac{1}{2}, 1\right]$ the probability that $x_p = 1$, implying that player 2 receives v_p with probability $1 - \alpha$.

We say that a project is being "explored" when cooperated on for the first time and "exploited" when cooperated on in both the current period and at least one prior period. There are no intertemporal restrictions on project availability.

We make two assumptions on the distribution of project values. First, we assume that the distribution of v_p admits a continuous density with a convex support equal to $\mathcal{R}^+$. Next, we assume $\mathbb{E}(v_p) \geq 2c$. These assumptions ensure that the first-best experimentation policy will be unique and non-empty.[7]

Further, the players exchange money twice during each period. At the beginning of each period t, the players make discretionary transfers to each other, where $w_{i,-i}^t \in \mathcal{R}^+$ denotes such a transfer from player i to player $-i$. At the end of each period t, players again make discretionary transfers to each other, where $b_{i,-i}^t \in \mathcal{R}^+$ denotes such a transfer from player i to player $-i$.[8] Finally, player i's period t payoff is equal to:

$$\pi_i^t = w_{-i,i}^t - w_{i,-i}^t + b_{-i,i}^t - b_{i,-i}^t + \sum_{p \in \mathbf{P}^t} \left(v_p 1_{x_p = i} - c \right), \text{ where } i \in \{1, 2\}, \tag{2.1}$$

and where $1_{x_p=i} = 1$ if $x_p = i$ and otherwise is equal to zero.

[6] Results hold as long as, for each project, one player values it above c and the other below it.

[7] Assuming an unbounded support also simplifies some technical aspects of the proofs. Further, $\mathbb{E}(v_p) < 2c$ could make no experimentation optimal in the first-best for low discount factors, unnecessarily complicating our analysis of the second-best policy where the discount factor is key.

[8] We incorporate the option of monetary transfers both before and after the players' project choices, although removing either would not qualitatively affect our results. Without transfers at the beginning of each period, surplus might no longer be fully redistributed across the players without affecting incentives. Without transfers at the end of each period, incentives for the current period would rely on transfers from the subsequent period, complicating the proofs.

We conclude the model's description by stating the timing of the stage game. Both players simultaneously choose their discretionary transfers $w_{i,-i}^t$. Next, both players simultaneously make their project choices P_i^t. For each project $p \in \mathbf{P}^t$, the players incur c and observe its beneficiary x_p and its value v_p, and player x_p pockets v_p. Finally, both players simultaneously choose their discretionary transfers $b_{i,-i}^t$.

Relational Contracts. A relational contract is a complete plan for the relationship. Let $h^t = (\mathbf{w}^1, \mathbf{P}^1, \mathbf{v}^1, \mathbf{x}^1, \mathbf{b}^1, \ldots, \ldots, \mathbf{w}^{t-1}, \mathbf{P}^{t-1}, \mathbf{v}^{t-1}, \mathbf{x}^{t-1}, \mathbf{b}^{t-1})$ denote the history up to date t and $\mathcal{H}^t$ the set of possible date t histories, where boldface lowercase letters indicate vectors. Then, for each date t and every history $h^t \in \mathcal{H}^t$, a relational contract describes: (i) the $\mathbf{w}^t$ transfers; (ii) the set of projects $\mathbf{P}^t(\mathbf{w}^t)$ as a function of $\mathbf{w}^t$; and (iii) the $\mathbf{b}^t(\mathbf{w}^t, \mathbf{P}^t, \mathbf{v}^t, \mathbf{x}^t)$ transfers as a function of $\mathbf{w}^t$, $\mathbf{P}^t$, and the realizations of $\mathbf{v}^t$ and $\mathbf{x}^t$. A relational contract is self-enforcing if it constitutes a Subgame Perfect Equilibrium of the repeated game. Within this class, we focus on equilibria that maximize joint surplus. Restricting to pure strategy equilibria is without loss of optimality since (i) mixing on transfers only increases the maximal transfers players can promise and (ii) mixing on projects leads to limited scope that can be replicated by being idle in some domains. In the event of a deviation in some period t, the players respond (i) by choosing $P_i^t = \emptyset$ and $b_{i,-i}^t = 0$ if these choices have not been made yet and (ii) by permanently breaking off their relationship (i.e., reverting to the worst equilibrium of the stage game from the next period onward). This punishment is without loss of optimality as it occurs out-of-equilibrium (c.f. Abreu, 1986).[9] Throughout, a relational contract is defined as "non-empty" if $\Pr(\sum_t |\mathbf{P}^t| > 0) > 0$.

3 First-Best Experimentation

We characterize the optimal experimentation policy for a benchmark where a single decision maker, "player 0," maximizes the sum of the payoffs of both players. This optimal experimentation policy is identical to the one we would obtain if we modified the model described in Section 2 so that the projects always benefit both players equally. The proof of the following proposition closely follows Bergemann and Välimäki (2001) and is provided in the Appendix, along with proofs for all other statements omitted from the main text.

Proposition 1 (**First-Best Experimentation Policy**)
For each domain j and period t, player 0 adopts the following experimentation policy: if a previously-explored project p has the highest value and $v_p \geq v^0(\delta)$, exploit it; If no previously-explored project has a value exceeding $v^0(\delta)$, explore a new project. The threshold $v^0(\delta)$ is increasing in δ.

Player 0 treats each domain separately and identically, given the additive separability of payoffs across projects and domains, as well as the ex ante identical nature of domains. The threshold v^0 arises from player 0's decision in each domain to either exploit the best project found thus far or explore a new project in search of a superior one. Furthermore, exploitation

[9] Alternatively, players could maintain the equilibrium but allocate all surplus to the non-deviator. This provides identical incentives and, being Pareto optimal, is less prone to renegotiation.

is permanent because player 0 does not acquire new information when exploiting a project. Likewise, given the infinite supply of ex ante identical projects in every domain, player 0 never chooses to exploit a project he chose not to exploit in the past. Finally, as the discount factor increases, the value of exploration increases, which explains the comparative statics result for v^0.

In summary, the first-best policy maximizes experimentation scope, with exploration/exploitation in each domain following a fixed, time-invariant threshold. We now analyze the model from Section 2, identifying when these features break down and the resulting dynamics.

4 Main Analysis

In Section 4.1, we characterize the class of optimal relational contracts on which the analysis focuses and establish a necessary and sufficient condition for an experimentation policy to be implementable by an optimal relational contract. In Section 4.2, we provide the conditions under which the players can implement the first-best policy. In Section 4.3, we characterize key properties of the optimal experimentation policy when they are unable to implement the first-best policy.

4.1 Optimal Experimentation Policies: Implementability

In our setting, surplus-maximizing relational contracts depend on the players' beliefs about the projects, which we denote by $\mu^t(h^t) := \{\Delta(v_p, x_p)|h^t\}_{p\in\mathcal{P}}$. We show that there exist surplus-maximizing relational contracts that condition on h^t only through $\mu^t(h^t)$. Moreover, restricting attention to relational contracts specifying the same continuation equilibrium following any two on-path histories h_1^t and $h_2^{t'}$ leading to the same beliefs μ is without loss of optimality, since the only history-dependent outcome that alters the set of continuation equilibria are the players' beliefs μ^t. Furthermore, the continuation equilibria prescribed by such surplus-maximizing relational contracts are also surplus-maximizing; otherwise, non-surplus-maximizing continuation equilibria could be replaced with surplus-maximizing ones, with appropriate transfers to maintain incentives. We refer to such relational contracts as optimal. The following proposition formalizes this characterization and provides a necessary and sufficient condition for an experimentation policy $\hat{\mathbf{P}} : \{\Delta(v_p, x_p)\}_{p\in\mathcal{P}} \to \mathcal{P}$ to be implementable by an optimal relational contract.

Proposition 2 (**Optimal Relational Contracts**)

- *For any surplus-maximizing relational contract, there exists an alternative surplus-equivalent relational contract such that (i) for all t and for all on-path histories $h^t \in \mathcal{H}^t$, the continuation equilibrium is surplus maximizing, and (ii) for any two on-path histories h_1^t and $h_2^{t'}$, if $\mu^t(h_1^t) = \mu^{t'}\left(h_2^{t'}\right)$, then the relational contract specifies the same continuation equilibrium following these histories.*

- *There exists an optimal relational contract that implements an experimentation policy*

$\hat{\mathbf{P}}(\cdot)$ *if and only if the following inequality holds for all on-path* $h^t \in \mathcal{H}^t$*:*

$$\sum_{p\in\hat{\mathbf{P}}(\mu^t)} \sum_{i=1}^{2} \max\left(0, c - \mathbb{E}\left(v_p 1_{x_p=i}|\mu^t\right)\right) \leq \mathcal{C}\left(\mu^t\right), \tag{2.2}$$

where $\mathcal{C}\left(\mu^t\right)$ *("the continuation value") is the expected net present value of the players' joint surplus starting in* $t+1$ *given* $\hat{\mathbf{P}}(\cdot)$ *and* μ^t*.*

The proof of this proposition extends the work of Levin (2003). In our setting, despite the stochastic nature of the players' continuation value, we show that considering its expectation is sufficient to characterize the experimentation policies that can be implemented by a relational contract.

The intuition for the first statement was provided above the proposition. Next, recall that the main tension faced by the players is that the experimentation policy which maximizes their joint surplus involves the selection of projects that do not benefit both players. Inequality (2.2) states that for an optimal relational contract to implement an experimentation policy everywhere on path, the continuation value induced by this policy must exceed the total reneging temptation across players and projects in all periods and histories. In turn, the total reneging temptation is the sum across players and projects of a project's reneging temptation to a player, which is either zero if the project generates a positive net expected gain, or equal to the magnitude of the net expected loss. The sum is across projects because, for any beliefs μ, each player can deviate by selecting any subset of $\hat{\mathbf{P}}(\mu)$. This condition is necessary for the relational contract to constitute an equilibrium. In the proof, we show that the presence of money also ensures sufficiency.

The proposition implies that characterizing the optimal relational contract reduces to determining the players' optimal experimentation policy, subject to Inequality (2.2) holding along the equilibrium path. This simplification arises because all transfers cancel out in both the joint surplus expression and the right-hand side of (2.2). Building on this observation, we now state the corresponding optimization problem.

The optimal experimentation policy in any given period depends only on the values of the most valuable projects identified in each of the m domains, denoted by $\hat{v}_1, \ldots, \hat{v}_m$, where $\hat{v}_j := 0$ if no projects have been explored in domain j. Players never exploit a project with a lower value than another, as doing so would reduce their joint payoff and make Inequality (2.2) (weakly) tighter. Thus, tracking $\hat{\mathbf{v}} := (\hat{v}_1, \ldots, \hat{v}_m)$ is sufficient to represent players' beliefs about the projects. For each j, they choose one of three actions: remain idle ($a_j = 0$), explore a new project ($a_j = 1$), or exploit the highest-valued known project ($a_j = 2$). The experimentation policy is then determined by solving the following Bellman equation, where $B(\hat{\mathbf{v}})$ represents the players' joint surplus:

$$B(\hat{\mathbf{v}}) = \max_{\mathbf{a}\in\{0,1,2\}^m} \left\{ \sum_{j=1}^{m} \left[1_{a_j=1}\mathbb{E}(v_p - 2c) + 1_{a_j=2}(\hat{v}_j - 2c)\right] + \mathcal{C}(\mathbf{a}, \hat{\mathbf{v}}) \right\} \tag{2.3}$$

$$\text{subject to:} \quad \sum_{j=1}^{m} \left[1_{a_j=1}c + 1_{a_j=2}\max\{0, c - (1-\alpha)\mathbb{E}(v_p)\}\right] \leq \mathcal{C}(\mathbf{a}, \hat{\mathbf{v}}). \tag{2.4}$$

Notably, Inequality (2.4) aggregates incentives across domains, introducing interdependencies. The implications of these interdependencies for players' experimentation will be the focus of our analysis. Moreover, they prevent an analytical characterization of the optimal policy, as we explain below.

Further, we caution against the following intuition. While the players' joint surplus, $B(\cdot)$, increases over time and the continuation value for a fixed policy, $\mathcal{C}(\mathbf{a},\cdot)$, also grows, the equilibrium continuation value, $\mathcal{C}(\mathbf{a},\cdot)$, is *not* necessarily monotonic. This non-monotonicity arises even under the first-best policy described in Proposition 1. For instance, $\mathcal{C}(a(v^0+\epsilon), v^0+\epsilon) < \mathcal{C}(a(0),0)$, because after identifying a project with a value slightly above v^0, player 0 becomes nearly indifferent between exploiting the current project and continuing to explore. This implies that the continuation value associated with exploration exceeds that of exploitation. The non-monotonic nature of the continuation value further complicates the analysis, as Inequality (2.4) does not necessarily relax over time.

4.1.1 Challenges in Characterizing Optimal Experimentation

As discussed in Section 1.1, our setting does not admit a Gittins Index characterization. More generally, any characterization of the optimal experimentation policy is generally infeasible. First, the choice set is discrete, which precludes the use of continuous optimization methods. Second, due to Inequality (2.4), this multi-dimensional optimization problem cannot be decomposed into m independent optimization problems. As a result, the curse of dimensionality arises for $m > 1$ due to two interrelated reasons. First, even for $m = 2$, the choice set in any given period t consists of 9 options (or 5, under symmetry), and this number grows exponentially with m. Second, determining whether a given choice is feasible and optimal requires knowledge of $B(\hat{\mathbf{v}}')$ for all $\hat{\mathbf{v}}' \geq \hat{\mathbf{v}}$, and subsequently, computing its respective integral over all possible future values of $\hat{\mathbf{v}}'$ for each choice $\mathbf{a}$ to evaluate $\mathbf{C}(\mathbf{a},\hat{\mathbf{v}})$. If the support of v_p were discrete with cardinality n, the problem could, in principle, be solved using "backward induction" on the Bellman equation. However, this approach is analytically feasible only when both n and m are very small (in Online Appendix A we provide a characterization for the $n = m = 2$ case).[10]

4.2 Implementability of First-Best Experimentation

We provide necessary and sufficient conditions on the values $\hat{v}_1,\ldots,\hat{v}_m$ under which the players can implement the first-best experimentation policy described in Proposition 1 in the current and in all subsequent periods. We refer to this outcome as "implementing the first-best experimentation policy." As we will show, there may exist a period $t' > t$ such that the players can implement the first best in period t' and all subsequent periods, but not in the earlier period t.

Inequality (2.2) implies that there exists a threshold $\tilde{v}$, equal to $c\,(1+\delta)\,/\delta$, which corresponds to the minimum project value required for a project's exploitation to be sustainable

[10] Note that meaningful exploration/exploitation decisions require the support of v_p to have a cardinality strictly greater than 2. For this reason, we assume a continuous support, which also facilitates the presentation of some of our results. However, it follows from Lemma 1 in the Appendix that none of our results rely on continuous supports.

in equilibrium when there is only one domain of cooperation ($m = 1$). Using this threshold $\tilde{v}$, we now provide the conditions on $\hat{v}_1, \ldots, \hat{v}_m$ under which the players can implement the first-best experimentation policy, which entails exploiting a project if and only if its value is at least v^0.

Proposition 3 (**Nec. and Suff. Condition for First-Best Experimentation**)
In any optimal relational contract and for any period t, the players implement the first-best experimentation policy for all $t' \geq t$ if and only if:

$$h(\hat{v}_1, \ldots, \hat{v}_m) := \frac{1}{m}\sum_{j=1}^{m} \max\{\hat{v}_j, v^0\} \geq \tilde{v} := c\frac{1+\delta}{\delta}. \tag{2.5}$$

As a result, there exists a threshold $\delta^0 < 1$ such that the players implement the first-best experimentation policy from period 1 onward if and only if $\delta \geq \delta^0$.

When Inequality (2.5) is satisfied, the continuation value of the relationship is sufficiently high to enable the implementation of the first-best experimentation policy. Because the players can pool relational incentives across domains, the condition requires that the *average* across domains of the maximum between the value of the most valuable project found in each domain and the threshold v^0 must exceed the threshold $\tilde{v}$. The function $h(\hat{v}_1, \ldots, \hat{v}_m)$ is not the arithmetic mean of the values $\hat{v}_1, \ldots, \hat{v}_m$ for two reasons: (i) under the first-best policy, players explore rather than exploit projects with values lower than v^0, and (ii) exploration contributes to the players' continuation value. Furthermore, the condition $v^0 \geq \tilde{v}$ is both necessary and sufficient for Inequality (2.5) to hold from period 1 onwards. The function $v^0(\delta) - \tilde{v}(\delta)$ exhibits a single-crossing property in δ, implying the existence of a threshold δ^0.[11]

Proposition 3 allows us to give necessary and sufficient conditions under which the players cease all exploration and transition to exploiting the most valuable project discovered in each domain, provided that they are already implementing the first-best experimentation policy. We refer to this outcome as "permanent exploitation."

Corollary 1 (**Nec. and Suff. Condition for Permanent Exploitation**)
In any optimal relational contract, the players permanently exploit projects with values $\hat{v}_1, \ldots, \hat{v}_m$ if and only if $\hat{v}_j \geq v^0$ for all j and the average of $\hat{v}_1, \ldots, \hat{v}_m$ exceeds $\tilde{v}$.

Proof of Corollary 1. Proposition 3 establishes that these conditions are jointly sufficient. Fixing $\hat{\mathbf{v}}$, the continuation value associated with permanent exploitation of $\hat{\mathbf{v}}$ is weakly lower than the continuation value under the first-best policy at $\hat{\mathbf{v}}$. Hence, if the players are able to permanently exploit $\hat{\mathbf{v}}$, they can also implement the first-best experimentation policy. This implies that these conditions are not only sufficient but also jointly necessary. □

The conditions stated in Corollary 1 imply that, in expectation, the players achieve the permanent exploitation outcome weakly later than if they could follow the first-best experimentation policy from period 1 onward. This delay relative to the first-best is strictly

[11]Proposition 1 establishes that player 0's threshold, $v^0(\delta)$, monotonically increases in δ, while the definition of $\tilde{v}$ implies that $\tilde{v}(\delta)$ monotonically decreases in δ.

positive when $\delta < \delta^0$. In fact, as we will show in Proposition 6, permanent exploitation in all domains of cooperation is not even guaranteed to occur.

We conclude by noting that the conditions listed in Corollary 1 fully characterize the players' optimal experimentation policy for the second natural benchmark case in our analysis: a single-domain collaboration. When there is only one domain (and the optimal relational contract is non-empty), the players face a simple decision in each period: either to exploit the best project found thus far or to explore a new project. The exploitation threshold in this setting is time-invariant, as the players' continuation value depends solely on the value of the best project in this single domain.

Corollary 2 (**Single-Domain Experimentation Benchmark**)
When $m = 1$, there exists a threshold $\delta^* < \delta^0$ such that the optimal relational contract is non-empty if and only if $\delta \geq \delta^*$. Furthermore, in any non-empty optimal relational contract, there exists a threshold $v^*(\delta) = \max\{\tilde{v}(\delta), v^0(\delta)\}$ such that the players explore projects until they find a project p with an associated value $v_p \geq v^*$. Once they find such a project, the players exploit it in all subsequent periods.

In this subsection, we have provided the conditions on the best projects found in each domain under which the players implement the first-best experimentation policy. We have also shown that, if δ is not sufficiently high, the players will initially be unable to implement the first-best policy. We now proceed to characterize key properties of the players' experimentation policy in the periods that precede an eventual transition to the first-best policy when collaboration spans multiple domains.

4.3 Second-Best Experimentation

We now analyze the players' optimal experimentation policy when they cannot implement the first-best policy in the current period. We refer to experimentation in this region as "second-best experimentation." A non-empty region where the second-best policy is relevant (i.e., $\delta \geq \delta^*$) but the first-best policy is not implementable (i.e., $\delta < \delta^0$) follows from Corollary 2 and is further examined in Section 5.1. This analysis focuses on the case where the maximal potential scope of experimentation, m, is strictly greater than 1 (for the case $m = 1$, see Corollary 2).

The players' exploration and exploitation decisions within their active domains of collaboration are inherently intertwined with their choices of which domains to engage in. To disentangle these dynamics, we analyze them separately: Section 4.3.1 focuses on exploration and exploitation, keeping scope decisions in the background, while Section 4.3.2 reverses the focus.

4.3.1 The Dynamics of Exploration-Exploitation Decisions

Under the first-best policy, each domain is treated independently and identically, with a time-invariant threshold for project exploitation. This time-invariance ensures that once a project is exploited or deemed unworthy of exploitation, the decision is permanent. For collaborative experimentation, players aggregate incentives across all domains, with domains being treated neither identically nor independently. We show that this observation implies that the criterion used to determine project exploitation is dynamic. As a result, the players may exploit a project temporarily, and further, they may recall a project they previously chose not to exploit.

Proposition 4 (**Temporary Exploitation and Recall of Projects**)
When the players cannot implement the first-best experimentation policy in period 1 *and the optimal experimentation policy is non-empty (i.e., when* $\delta \in [\delta^*, \delta^0)$*), then with strictly positive probability for any* $m > 1$*, at least one of the following occurs:*

1. *The players choose to exploit a project in period* t*, but later decide not to exploit the same project in some period* $t' > t$*.*

2. *The players choose not to exploit a project in period* t*, but later decide to exploit the same project in some period* $t' > t$*.*

We provide intuition for why these two seemingly suboptimal behaviors are optimal by examining two specific examples with $m = 2$. The proof establishes that these behaviors necessarily occur with strictly positive probability.

The first statement can be understood by considering the following scenario. Suppose the values of the best projects in domains 1 and 2 satisfy $\hat{v}_1 \geq \hat{v}_2$. Further, assume that both values are sufficiently large for the players' scope of experimentation to be maximal, but not large enough to enable them to implement the first-best policy. If $\hat{v}_1$ is particularly high, the players will choose to exploit the project in domain 1 and explore in domain 2. Now, imagine that the exploration in domain 2 uncovers a project with a value slightly higher than $\hat{v}_1$. In this case, the players find themselves in a situation similar to the previous period, but with the roles of the domains reversed. They will now choose to exploit the newly discovered project in domain 2 and explore in domain 1. In Section 5.2, we simulate the optimal experimentation policy for a parameterized example to further illustrate and develop intuition about the emergence of this behavior.

To understand the intuition behind the second statement, consider a scenario where the discount factor δ is small enough to prevent the exploitation of projects with values only slightly above the threshold v^0. Suppose the players' scope of experimentation is maximal, which occurs, for instance, when $\alpha = 1/2$.[12] If period 1 explorations yield two projects with values just above v^0, the players must explore again in the next period. However, if a newly explored project has a sufficiently high value, it can raise the continuation value of their relationship, potentially enabling first-best experimentation. In this case, they may optimally revert to exploiting a period 1 project despite initially choosing to explore further.

Temporary project exploitation or project recall are common in experimentation settings, and can arise due to various factors, including the presence of a finite number of projects or project characteristics that may not be fully revealed immediately. Our analysis shows that strategic interactions alone can also drive these behaviors.

[12] When $\alpha = 1/2$, exploration occurs in each domain in the static equilibrium, so any optimal relational contract implements an experimentation policy with maximal scope throughout.

4.3.2 The Dynamics of the Scope of Experimentation

Proposition 3 established a threshold δ^0, such that when $\delta \geq \delta^0$, players implement the first-best policy starting in period 1, maintaining maximal scope. We now examine the dynamics of the players' scope of experimentation when $\delta \in [\delta^*, \delta^0)$ and show that scope is not always maximal along the equilibrium path. To focus on the relevant case, we assume $(1-\alpha)\mathbb{E}(v_p) < c$, requiring player 1 to incentivize player 2 to explore. If instead $(1-\alpha)\mathbb{E}(v_p) \geq c$, project exploration is a static equilibrium, and optimal experimentation always maintains maximal scope.

We define a non-empty experimentation policy as "initially maximal" if $|\mathbf{P}^1| = m$, "initially limited" if $|\mathbf{P}^1| < m$, "terminally maximal" if $\lim |\mathbf{P}^t| = m$, and "terminally limited" if $\lim |\mathbf{P}^t| < m$. An initially maximal policy is always preferred over an initially limited one whenever both are implementable, as exploring all domains provides immediate benefits ($\mathbb{E}(v_p) \geq 2c$) and maximizes the continuation value of the relationship. The key question, then, is whether an initially limited policy can be implemented when an initially maximal one cannot. Intuitively, starting with a limited number of domains and allowing for future expansion may be more sustainable, as (i) it reduces early reneging temptation while maintaining a high continuation value due to these potential future scope expansions, and (ii) finding valuable projects in early domains can enable both their exploitation and the exploration of additional domains. We show that this intuition holds when the maximum potential scope of experimentation m exceeds a threshold, but may fail below it.

To build intuition, we present the period-1 version of Inequality (2.2) for a specific initially limited policy, where players explore projects in domain 1 during period 1, and the corresponding inequality for the initially maximal policy, respectively:

$$c \leq \delta \int B(\hat{v}_1, 0, \ldots, 0|\delta) d\hat{v}_1, \tag{2.6}$$

$$m \cdot c \leq \delta \int B(\hat{v}_1, \ldots, \hat{v}_m|\delta) d\hat{v}_1, \ldots, d\hat{v}_m, \tag{2.7}$$

where $B(\cdot)$ was defined in Equation (2.3), and where we explicitly highlight the relationship between $B(\cdot)$ and δ, as this will play a key role in the intuition below. We focus solely on period 1, remaining agnostic about the long-term dynamics of both policies. We note that the right-hand side of (2.7) increases with δ, indicating the existence of a cutoff $\bar{\delta}(m) \in (0, 1)$ below which this constraint is violated. Therefore, the question is whether (2.6) holds for $\delta < \bar{\delta}(m)$.

We proceed under the (incorrect) assumption that $B(\cdot|\delta)$ is continuous with respect to δ.[13] Under this assumption, and using (2.6) and (2.7), the initially limited policy outlined above is optimal when δ is just below $\bar{\delta}$ if and only if:

$$\int B\left(\hat{v}_1, 0, \ldots, 0|\bar{\delta}(m)\right) d\hat{v}_1 > \frac{1}{m} \int B\left(\hat{v}_1, \ldots, \hat{v}_m|\bar{\delta}(m)\right) d\hat{v}_1, \ldots, d\hat{v}_m. \tag{2.8}$$

The right-hand side of this inequality represents the average surplus per domain from period 2 onward under the initially maximal policy, and, intuitively, is bounded above by that of

[13] $B(\cdot \mid \delta)$ is not continuous with respect to δ because optimal experimentation is not.

a single domain under the first-best policy. Conversely, the left-hand side represents the total surplus across domains from period 2 onward under the initially limited policy. By monotonicity of the Bellman equation, the left-hand side of Inequality (2.8) is bounded below by:

$$B\left(\mathbf{0}|\bar{\delta}(m)\right) = m\mathbb{E}\left(v_p - 2c\right) + \mathcal{C}\left(\mathbf{0}\right) \geq m\mathbb{E}\left(v_p - 2c\right) + m\left(c - (1-\alpha)\mathbb{E}\left(v_p\right)\right), \tag{2.9}$$

where the last step follows from Inequality (2.2). Since this lower bound diverges with m, Inequality (2.8) holds for sufficiently large m, implying that an initially limited policy has a lower critical discount factor than the initially maximal one. We formalize this argument in the Appendix, accounting for the potential discontinuity of $B(\cdot|\delta)$ in δ.

In contrast, for small values of m, an initially limited experimentation policy may or may not be easier to implement than an initially maximal one. As shown in Proposition 4, exploration and exploitation decisions are optimally co-determined across domains. Thus, delaying exploration in domain j not only reduces its associated surplus but may also lower the surplus in all other domains. In the Appendix, we show that, when m is small, this advantage of initially maximal policies can outweigh the benefits of initially limited policies discussed above. To demonstrate this, we construct a distribution of project values that yields significant advantages of conducting multiple explorations in parallel, making initially limited policies suboptimal for all discount factors. These intuitions are consolidated in the following proposition.

Proposition 5 (**Initial Scope of Experimentation**)
Suppose $(1-\alpha)\mathbb{E}(v_p) < c$ *and* $m > 1$. *Two thresholds* $\delta^* \leq \bar{\delta} < \delta^0$ *exist such that:*

1. *If* $\delta \geq \bar{\delta}$, *any optimal relational contract is such that the scope of experimentation is initially maximal.*
2. *If* $\delta \in [\delta^*, \bar{\delta})$, *any optimal relational contract is such that the scope of experimentation is initially limited.*
3. *If* $\delta < \delta^*$, *the scope of experimentation is equal to zero in all periods.*

Further, denote $m^* := \sup_{m\geq 2}\{m : \bar{\delta} = \delta^*\}$. *An initially limited experimentation policy is optimal for intermediate discount factors for large* m *(i.e.,* $m^* < \infty$*), but may never be optimal for small* m *(i.e.,* $m^* > 2$ *may occur).*

The previous proposition established results on the players' initial scope of experimentation but did not address its long-term dynamics. We now present findings on their terminal scope. Any non-empty experimentation policy—whether initially limited or initially maximal—has a strictly positive probability of becoming terminally maximal, as players may always, by chance, identify a project valuable enough to sustain the first-best policy indefinitely. Moreover, as discussed above, the optimality of initially limited policies relies crucially on the prospect of sufficiently likely subsequent scope expansions. These observations raise a broader question: is experimentation scope guaranteed to be maximal—and therefore efficient— asymptotically?

Proposition 6 (**Terminal Scope of Experimentation**)
The following statements hold:

1. *There exist optimal experimentation policies that are both initially limited and, with strictly positive probability, terminally limited.*

2. *There exist optimal experimentation policies that are both initially maximal and, with strictly positive probability, terminally limited.*

The reason why the players' scope of experimentation may be terminally limited on path can be understood by considering a vector of project values, $\hat{\mathbf{v}}$, and a subset of domains $s \subset \{1, \ldots, m\}$ for which:

[label=)]players can only permanently exploit projects in s due to insufficient continuation value; exploring any domain $j \in \{1, \ldots, m\} \setminus s$ requires foregoing exploitation in one or more domains in s due to insufficient continuation value; and players prefer exploiting all projects in s over delaying some exploitations to explore additional domains.

To prove the first statement of the proposition (respectively, the second statement), in the Appendix we show that a), b), and c) hold simultaneously under an initially limited (respectively, initially maximal) policy. Intuitively, these dynamics arise only when $\hat{\mathbf{v}}$ is high enough for a) and c) to hold but low enough for b) to be satisfied.[14]

In this subsection, we analyzed the dynamics of the players' scope under second-best experimentation. We showed that for intermediate discount factors and large maximal potential scope m, the players find it optimal to begin with limited scope, an approach made credible by the possibility of many subsequent scope expansions created by the discovery of valuable projects in the early domains of cooperation. Because the discovery of such projects is path-dependent, the players may end with a permanently limited and thus inefficient scope of experimentation.

5 Further Analysis and Extensions

This section extends our analysis in three directions. First, we examine how the maximum potential scope of experimentation influences its feasibility and profitability. Second, we analyze a concrete example to graphically illustrate some of the key dynamics of the model. Finally, we explore several simple extensions in which the domains of cooperation are not identical or independent.

[14]The fact that terminally limited scope may arise for intermediate values of $\hat{\mathbf{v}}$—and consequently for intermediate values of the relationship—is reminiscent of Chassang (2010)'s result, where exploration may cease when some but not all productive actions have been "revealed," leaving the value of the relationship in an intermediate range. Despite the differences in setting, the core intuition is similar: conducting additional exploration requires halting the exploitation of an existing project. In our setting, the newly explored project cannot be exploited in the current period due to b). In Chassang's setting, the absence of transferable utility means that exploring an additional action may require terminating the relationship, thereby sacrificing some future exploitation. The difficulty in computing the endogenous loss from these forgone exploitations in closed form is precisely what hinders analytical characterizations in both settings.

5.1 Comparative Statics of Scope

The maximum potential scope of experimentation, m, can vary significantly depending on the application. When firms pool resources, some pairings may yield numerous cooperation opportunities, while others result in fewer viable collaborative areas, depending on the complementarity of their assets. In this subsection, we analyze how variations in m affect the profitability and sustainability of experimentation.

Before proceeding, we revisit Bernheim and Whinston (1990)'s analysis of scope, in stationary environments without learning dynamics. First, for a scaling factor $k \geq 1$, when scaling the scope of interaction by k, players can maintain the same per-domain average payoffs by replicating the original equilibrium k times independently. Second, when domains are identical, pooling incentives across domains cannot improve the players' per-domain average payoffs. However, if domains are asymmetric, players may gain from doing so and, hence, greater scope may be beneficial.

Let $\tilde{\pi}(m) := \pi(m)/m$ denote the average joint surplus per domain of the collaboration. Recall that $\delta^*(m)$ represents the minimum discount factor for which the optimal relational contract is non-empty. For a scaling factor $k \geq 1$, the following weak inequalities follow from Bernheim and Whinston (1990): $\tilde{\pi}(mk) \geq \tilde{\pi}(m)$ and $\delta^*(mk) \leq \delta^*(m)$.[15] In our setting, we can provide necessary and sufficient conditions for these inequalities to hold strictly, due to the dynamics stemming from the players' exploration of projects. Specifically, $0 < \delta^*(m \cdot k) < \delta^*(m)$ for $k > 1$ if $(1-\alpha)\mathbb{E}(v_p) < c$ and otherwise $\delta^*(m \cdot k) = 0$ regardless of k. When $(1-\alpha)\mathbb{E}(v_p) < c$, the optimal relational contract will be empty for low discount factors. In these instances, scaling up m will strictly decrease δ^*. To see why, note that if the players were to implement k independent and concurrent collaborations, each with an identical experimentation policy, the threshold $\delta^*(m \cdot k)$ would be independent of k. However, this approach would be inefficient as it only leverages relational interdependencies within segmented multi-domain experimentation policies. Therefore, the players could sustain a non-empty relational contract for lower discount factors by leveraging interdependencies across all $m \cdot k$ domains. By an identical reasoning, $\tilde{\pi}(m \cdot k) > \tilde{\pi}(m)$ whenever the second-best experimentation policy is non-empty.

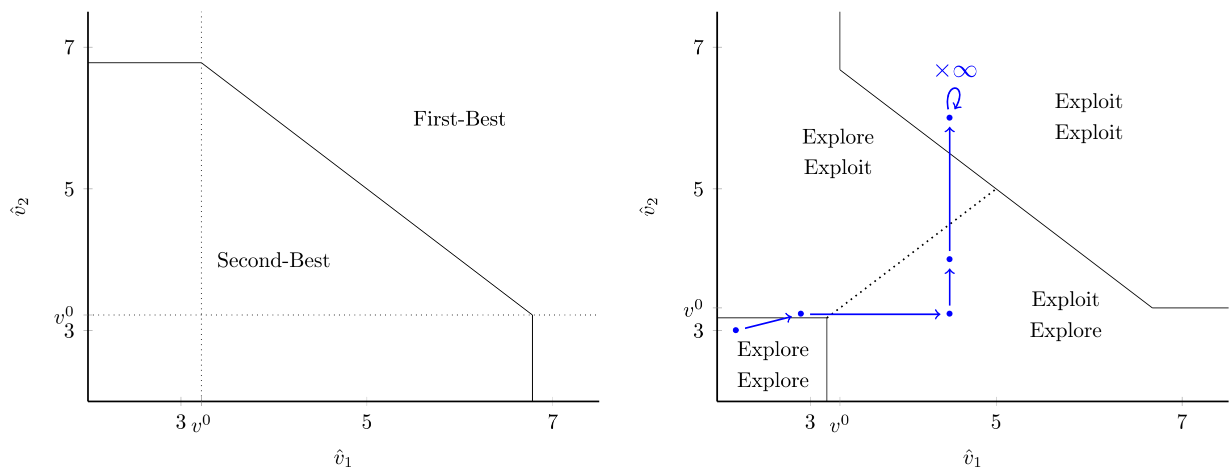
5.2 Multi-Project Collaborations: A Graphical Illustration

We analyze an example with specific parameter values. We set $c = 1$ and $\delta = 1/3$. Furthermore, we consider a symmetric relationship by setting $\alpha = 1/2$. The players can cooperate in two domains ($m = 2$). Finally, the project values v_p are drawn from a shifted exponential distribution with a rate parameter $\lambda = 1/2$, i.e., $v_p \sim 1 + \text{Exp}(1/2)$. Under this distribution, $\mathbb{E}(v_p) = 3$. The players' scope of experimentation is always maximal since $\alpha\mathbb{E}(v_p) - c = (1-\alpha)\mathbb{E}(v_p) - c > 0$, making exploration preferable to inactivity. Further, the continuation value $\mathcal{C}(\hat{v}_1, \hat{v}_2)$ is weakly greater than 1 for all $\hat{v}_1$ and $\hat{v}_2$, as players can always explore two new projects per period, yielding a payoff of $\mathbb{E}(v_p) - 2c = 1$ per project and a

[15] $\tilde{\pi}(m)$ is not necessarily monotone in m. For example, it may depend on the parity of m—pooling incentives across two domains could enable a relatively efficient experimentation policy, yet leave insufficient slack to improve efficiency in a third domain (as seen in the distribution used to prove Statement 2 of Proposition 6).

continuation value $\mathcal{C}(\hat{v}_1, \hat{v}_2)$ also equal to 1. As a result, if Inequality (2.5) does not hold, players either: (i) exploit one project while exploring another, or (ii) explore two projects simultaneously.

Figure 2.1(a). The figure depicts the first-best policy stated in Proposition 1. The vertical and horizontal black dotted lines represent the time-invariant threshold v^0 for domains 1 and 2, respectively. In both domains, projects with values above this threshold are permanently exploited, while those below are never exploited. Further, the solid black line in the figure divides the project value space into two distinct regions. This line represents the set of $(\hat{v}_1, \hat{v}_2)$ values satisfying $h(\hat{v}_1, \hat{v}_2) = \tilde{v}$, a condition stated in Proposition 3. To the northeast of this line, in the region labeled "First-Best," the players can implement the first-best experimentation policy. In contrast, to the southwest of the line, in the region labeled "Second-Best," the players can exploit at most one project at a time. The horizontal segment represents where $\hat{v}_1 < v^0$, so project 1 is never exploited under the first-best policy, and implementation depends solely on $\hat{v}_2$. Symmetrically, the vertical segment shows where $\hat{v}_2 < v^0$, with implementation depending only on $\hat{v}_1$. The downward-sloping segment captures instances where both $\hat{v}_1$ and $\hat{v}_2$ exceed v^0. Here, increasing one project's value allows decreasing the other's while maintaining sufficient continuation value for first-best policy implementation.



((a)) Feasible Region for First-Best Experimentation

((b)) Project Exploitation and Sample Path

Figure 2.1: Optimal Multi-Project Experimentation

In the figure, we assume $c = 1$, $m = 2$, $\delta = 1\,/\,3$, and $v_p \sim 1 + \text{Exp}(1\,/\,2)$. $\hat{v}_1$ and $\hat{v}_2$ denote the values of the best projects discovered in domains 1 and 2, respectively. The left figure plots (i) the threshold v^0 for switching from exploration to exploitation in the first-best and (ii) the set of $\hat{v}_1$ and $\hat{v}_2$ values satisfying $h(\hat{v}_1, \hat{v}_2) = \tilde{v}$ in solid black. The right figure divides the project value space into four regions, determined by the exploitation or non-exploitation of each project. The top mention indicates the decision for the project with value $\hat{v}_1$, while the bottom mention shows the decision for the project with value $\hat{v}_2$. In Blue, we plot one realization of a sample path.

Figure 2.1(b). The project value space is divided into four regions, determined by the exploitation or non-exploitation (in favor of exploration) of each project. The top mention indicates the decision for the project with value $\hat{v}_1$, while the bottom mention shows the decision for the project with value $\hat{v}_2$. It follows from Figure 2.1(a) that both projects are chosen for exploitation when in the "First-Best" region and $\hat{v}_1, \hat{v}_2 \geq v^0$. Outside of this region, the players can choose one project for exploitation at most. One can prove that there exists a threshold, v', on the value of the best of the two projects such that, below this threshold, the players choose to explore two new projects rather than exploiting the best of the two projects. We observe that the threshold v' is lower than v^0, indicating that players may opt to exploit a project even when they are certain to not permanently exploit it in the future.[16]

Figure 2.1(b) also presents a sample path illustrating the evolution of realized project values over time, depicted in blue. In the early phase where the players are exploring two projects, both $\hat{v}_1$ and $\hat{v}_2$ weakly increase over time. In the phase where the players exploit a project in domain j, $\hat{v}_j$ remains constant, while $\hat{v}_{-j}$ weakly increases over time. Finally, in the phase where the players exploit both projects, $\hat{v}_1, \hat{v}_2$ stay constant because exploitation is permanent. Arrows are used to signify changes in project values when a more valuable project is identified, while self-loops indicate situations where more valuable projects are either not discovered or not pursued. The path shown in the figure includes temporary exploitation in domain 2 (of a project guaranteed to be not permanently exploited), as discussed in Proposition 4.

5.3 Beyond Independent and Identical Domains

Our main analysis assumed identical and independent collaboration domains. In practice, firms often collaborate across domains with diverse characteristics and technological interdependencies. This reality raises the question: Which domains, if any, should be prioritized when initiating collaboration? The Online Appendix explores three natural scenarios that address these questions and formulate predictions. We briefly summarize these extensions here.

When to explore risky domains?
Our main analysis, by assuming an infinite number of independent and identically distributed projects, effectively eliminated risk considerations. However, collaborating parties often face uncertainty about their collaboration's potential value, with varying degrees of uncertainty across cooperation domains. For instance, a buyer-supplier collaboration might involve both incremental improvements to an existing product and the development of a radically new—and thus potentially unprofitable—project. To capture these features, we modify a two-domain version of our framework by supposing that that the first domain is exactly as in the main model, while the other contains a single project with either low or high value. We show that even when immediate cooperation across both domains is feasible, players may choose to postpone exploring the risky domain 2 project. This delay continues until a sufficiently valuable project is discovered in domain 1. Such a gradual approach safeguards the collaboration against complete dissolution should the radical innovation fail.

[16]The threshold v' presented in the figure is computed using numerical integrals and approximate solutions to the Bellman equation. The result that v' can be lower than v^0 can be proven analytically.

Can "win-win" projects serve as stepping-stones?
In the main analysis, we assumed that each project's benefits accrue to only one player. However, the model can be extended to reflect more nuanced real-world scenarios. Collaborating parties often engage in both "win-win" projects yielding mutual benefits and projects that disproportionately advantage certain participants. In modeling these scenarios, this extension assumes two domains with distinct benefit structures. In one domain, projects yield equal benefits to both players.[17] The other domain follows the main analysis, where project benefits accrue exclusively to one player. We show that optimal experimentation is initially limited for low values of the discount factor and that the domain with symmetric projects is explored first.

How do technological interdependencies influence gradualism?
In the third extension, we introduce positive correlation between project values across domains, such that discovering a valuable project in one domain immediately reveals a project of equal value in the other. This assumption reflects how success in one area can enhance opportunities in another (e.g., mRNA technology's wide applicability across medical conditions). Absent incentive issues, players would optimally explore both domains concurrently to expedite valuable project discovery. With asymmetric benefits, an initially limited approach is strictly optimal for intermediate discount factors. These findings suggest initially limited approaches are more likely to be optimal in R&D environments with stronger cross-domain knowledge spillovers.

6 Applied Insights

This section connects our theoretical analysis to two key literatures: buyer-supplier relationships and persistent productivity differences across firms.

6.1 Buyer-Supplier Collaborations

The economics literature on buyer-supplier relationships has predominantly examined issues such as vertical integration in the presence of relationship-specific investments (Williamson, 1975; Grossman and Hart, 1986a; Hart and Moore, 1990), optimal contracts under externalities or agency issues (see references in Tirole, 1988, Chapter 4), and, more recently, relational contracts for supplier allocation (Board, 2011; Andrews and Barron, 2016). While these studies justifiably assume predetermined gains from trade to address their specific objectives, our research explores a complementary direction: scenarios requiring collaborative experimentation to determine the gains from trade, often across multiple products or markets.

Our model formalizes the process of collaborative experimentation in buyer-supplier relationships through several key elements. The parameter m represents the number of product categories or market geographies. Both firms make non-contractible investments of c for experimentation. These investments are observable to both parties but not verifiable by third parties, hence not contractible. The innovation process involves both firms, each

[17] The Online Appendix includes another extension in which domains differ in the probability α, with qualitatively similar results.

possessing complementary and indispensable expertise or resources. Even after the exploration phase, when parties agree on an input or service to exploit, non-contractible investments (also c) remain essential. These include efforts such as worker training and marketing. The distribution of benefits is asymmetric because final product proceeds accrue to the buyer (high α), who compensates the supplier through either the upfront transfer w or the bonus b.

Our theoretical analysis both draws from and contributes to an extensive body of case-study literature on buyer-supplier dynamics. This literature emphasizes experimentation and trust as critical factors for successful collaborations, particularly in contexts where benefits are asymmetrically distributed. A McKinsey report highlights this asymmetry of benefits: "Some collaborations promise equal benefits for both parties. [...] In other cases, however, the collaboration might create as much value overall but the benefit could fall more to one partner than to the other" (Benavides et al., 2012). This asymmetry underscores the central role of trust, given the inherent limitations of formal contracts. Doney and Cannon (1997) distinguish between two types of trust: "benevolence" trust (belief in a partner's genuine desire to collaborate) and "credibility" trust (expectation that a partner will fulfill promises due to self-interest). Our analysis primarily focuses on credibility trust, operating under the assumption that both parties desire collaboration. Consequently, in this section we emphasize work that similarly concentrates on credibility trust. The concept of benevolence trust, while important, corresponds more directly to the analyses by Watson (1999, 2002), which we discuss in Section 1.1.

Dwyer et al. (1987) highlight the dynamic nature of buyer-supplier relationships, emphasizing the central role of relational contracts. They describe an initial "search and trial phase" that evolves into an "expansion phase," characterized by increased risk-taking and deeper mutual dependence. As they note, "The rudiments of trust and joint satisfactions established in the exploration stage now lead to increased risk taking within the dyad. Consequently, the range and depth of mutual dependence increase." A senior executive from a Toyota supplier similarly described their relationship with Toyota: "We started by making one component, and as we improved, [Toyota] rewarded us with orders for more components" (Liker and Choi, 2004). The common pattern of these relationships starting small before expanding is consistent with our findings, particularly Proposition 5, which shows the potential optimality of gradual expansion in collaborative scope. It also supports our extension in Section 5.1, which examines the strategic delay of high-risk ventures in these relationships.

Building on Dwyer et al. (1987), Vanpoucke et al. (2014) corroborate both the prevalence of gradualism and the occurrence of extended experimentation periods in buyer-supplier relationships. These phenomena are driven by the parties' need to establish credibility in the context of relational contracts. As one CEO in their study noted, "We use contracts, but not everything, certainly in the long run, can be put in contracts." Their case study of soybean product development, where partners took a decade to initiate integration and build sufficient credibility, illustrates this phenomenon. This evidence is consistent with our analysis, particularly Corollary 1, which predicts that collaborating firms must engage in prolonged experimentation in order to identify joint projects of sufficient value to sustain the subsequent exploitation phase. Furthermore, Vanpoucke et al. (2014) emphasize the strong path dependence of relationship dynamics, observing that "events, rather than time," define relationship development stages. Their case studies consistently reveal that successes in initial cooperation domains typically drive further joint collaborations. This observation supports our theoretical model, where increases in scope are driven by discrete "events" that change the players' continuation value from the collaboration, rather than the mere passage of time.

Lastly, our analysis, particularly Proposition 6, showed that the long-term scope of a collaboration is determined during the initial phases, with early outcomes influencing the trajectory and ultimate extent of the partnership. This finding is corroborated by the existing literature. Dwyer et al. (1987) characterize the early exploration phase in buyer-supplier relationships as "very fragile," highlighting the critical nature of these initial interactions. Benavides et al. (2012) provide a concrete example of this fragility, describing a case where an early collaboration attempt between a retailer and manufacturer yielded somewhat disappointing results. While their relationship did not terminate entirely, Benavides et al. (2012) suggest that this initial setback was the primary reason their partnership did not expand further.

6.2 Persistent Performance Differences

While much of our focus has been on interactions between firms, our model serves as a valuable lens for examining employer-employee dynamics. One can conceptualize one party in our model as the employer and the other as the employee, where, for instance, benefits consistently accrue to the employer. Furthermore, the different domains of collaboration can be seen as various dimensions of the production improvement process.

With this interpretation in mind, our work also contributes to the literature on persistent performance differences among seemingly similar enterprises (see Syverson, 2011; Gibbons and Henderson, 2013, and references therein). Numerous empirical studies have documented enduring disparities in firm performance across a range of industries, with these gaps proving surprisingly robust against plausible explanations such as market competition or local geographical and demand conditions, while being strongly associated with managerial practices (c.f. Bloom and Van Reenen, 2007). According to Gibbons and Henderson (2013), and the body of evidence they review, variations in managerial practices, because of their reliance on relational contracts, are key in creating productivity disparities across firms. We adapt for our purposes their categorization of explanations: (i) managers might either be unaware of their poor performance, or, even if aware, believe that the best practices from other firms are not suitable for their context; (ii) managers are aware of their poor performance and are able to seek superior managerial practices suitable to their context, but opt not to; and (iii) managers are "striving mightily" to adopt superior practices but face hurdles during the implementation phase.

The first explanation underscores information barriers, prompting questions about why such information does not diffuse more readily (c.f. Bloom et al., 2013; Atkin et al., 2017).The second explanation is consistent with the framework developed by Chassang (2010) and discussed in Section 1.1, in which players are informed about the existence of more efficient practices but choose not to pursue them to preserve their relationship. Our analysis in Section 4.3 provides a complementary rationalization of explanation (ii) by showing that the long-run scope of collaboration may be inefficiently limited. When players transition from exploration to exploitation in one domain, they may lose the ability to cooperate in other domains, potentially resulting in limited scope.

Unlike other models we know, our model also offers insight into explanation (iii) presented by Gibbons and Henderson (2013). Consider two organizations with identical characteristics implementing ex-ante identical experimentation policies, operating under a discount factor where the scope of experimentation is initially limited. Their paths diverge if one organization discovers a highly valuable practice early on, thus expanding its scope, while the other does not. The second organization, still attempting to achieve any success, appears to be "striving mightily" to match the first organization's performance. However, identifying superior practices is time-intensive. The second organization cannot increase its scope until it finds a sufficiently valuable practice, potentially leading to a persistent performance gap.

7 Concluding Remarks

This paper presents a framework for analyzing the dynamics of multi-domain collaborative experimentation in scenarios where benefits are unevenly distributed among participants and any experimentation policy must be self-enforcing. Our model yields three key insights. First, when the initial relationship value is low, the collaborating parties do not treat each domain of experimentation independently and they engage in extended exploration phases. Second, cross-domain relational interdependence in optimal experimentation leads to seemingly counterintuitive exploration/exploitation decisions, including prolonged exploitation of ultimately discontinued projects or revival of previously abandoned ones. Third, experimentation often progresses gradually, with parties initially exploring some domains and potentially expanding to others based on initial success, and exploration of all domains is not guaranteed.

While our primary focus is on buyer-supplier dynamics and firm-level productivity, our framework extends to political economy settings involving multi-domain collaboration. In federal systems, central governments use fiscal transfers to incentivize subnational policy experimentation (c.f. Callander and Harstad, 2015; Wang and Yang, forthcoming). Our analysis highlights substantial path dependence in policy implementation and suggests a priori unexpected spillovers across policy domains. Moreover, political and economic unions like the European Union facilitate cross-domain collaboration through shared resources and structural funds, with members retaining the option to exit. Consistent with our analysis, the formation of the EU involved gradual step-by-step integration, initially prioritizing mutually beneficial projects before expanding to more ambitious policies with unevenly distributed costs and benefits (see, e.g., Spolaore, 2015, and references therein, as well as Section 5).

Future work could extend the current framework in several directions. For example, relaxing the assumption of identically and independently distributed project benefits within domains could help address questions related to directed innovation strategies and differentiate between radical and incremental innovation (c.f. Callander, 2011; Garfagnini and Strulovici, 2016; Callander and Matouschek, 2019). Further, we assumed that both players' cooperation was necessary for exploration and exploitation, keeping their outside options independent of experimentation. Future research could explore scenarios where players' outside options evolve based on their experimentation history, examining how this additional interdependence affects joint experimentation dynamics. Finally, introducing asymmetric roles in the collaboration presents another natural extension. One could model a scenario where exploration requires only one player (e.g., an R&D unit), while exploitation needs a different player (e.g., a

Sales unit). This approach would enable analysis of cooperation dynamics in contexts where exploration and exploitation efforts are disentangled (see Krieger et al., 2019; Lizzeri et al., 2024, for qualitative and theoretical treatments, respectively).

8 Appendix A

3. *Proof of Proposition 1.* Following a reasoning almost identical to that in Bergemann and Välimäki (2001), player 0 treats each domain independently and identically and never recalls a project because $|\mathcal{P}_j| = \infty \ \forall j$. Therefore, the optimal policy conditions only on the project with the highest value amongst all previously explored projects, whose value we denote $\hat{v}$. The Bellman Equation for player 0 is:

$$B^0(\hat{v}) = \max_{\text{explore, exploit } \hat{v}} \left\{ \mathbb{E}\left(v'\right) - 2c + \delta \mathbb{E}\left(B^0\left(\max\left(\hat{v}, v'\right)\right)\right), \hat{v} - 2c + \delta B^0\left(\hat{v}\right) \right\}. \tag{2.10}$$

The first term in the maximum operator corresponds to the player's expected surplus when exploring one more project and the second term is their surplus when exploiting the project with value $\hat{v}$. Next, there exists a threshold v^0, wherein the players explore if $\hat{v} < v^0$ and exploit if $\hat{v} \geq v^0$. Further, Blackwell's Sufficient Conditions imply that there exists a unique solution to the Bellman Equation, and hence the threshold rule dictated by v^0 is a solution. This threshold is determined by:

$$\frac{1}{1-\delta}(v^0 - 2c) = \mathbb{E}(v_p - 2c) + \frac{\delta}{1-\delta}\mathbb{E}(\max\{v, v^0\} - 2c), \tag{2.11}$$

where standard comparative statics arguments imply that v^0 is increasing in δ. □

Proof of Proposition 2. Recall that after a deviation in period t, players set $P_i^t = \emptyset$ and $b_{i,-i}^t = 0$ if not already chosen. In subsequent periods, they revert to the static equilibrium with zero transfers and no selected projects.

The proof proceeds in four steps: (i) we show that it is without loss of optimality to restrict attention to relational contracts that are surplus-maximizing following every on-path history h^t; (ii) we provide a necessary and sufficient condition for the existence of a relational contract that implements a given experimentation policy $\hat{\mathbf{P}}(\cdot)$; (iii) we show that this condition is independent of the division of surplus between the players; and (iv) we show that, for any two histories that generate the same beliefs, selecting the same continuation equilibrium is without loss of optimality.

Step 1 We show that it is without loss of optimality to restrict attention to relational contracts that are surplus-maximizing following every on-path history h^t. To see this, suppose that there exists an on-path history h^t such that the continuation equilibrium starting in period t, denoted by e^1, has lower total surplus than an alternative continuation equilibrium e^2. Thus, if we define $\mathcal{C}_i^k$ to be the continuation value to player i in equilibrium e^k, then $\sum_i \mathcal{C}_i^1 < \sum_i \mathcal{C}_i^2$. For the rest of Step 1, we omit the superscript $t-1$ in our notation, as we are solely concentrating on period $t-1$ objects.

Let us modify the players' relational contract such that play in and after period t is dictated by e^2 and the period $t-1$ $b_{i,j}(\cdot)$ transfers associated with history h^t (and, thus, corresponding to a specific realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$) are adjusted so that: (i) player 2's expected payoff following the realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$ is the same as under the original equilibrium and (ii) player 1's expected payoff following the realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$ increases by $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1$. Specifically, take the vector of transfers $\mathbf{b}_1 = (b^1_{1,2}, b^1_{2,1})$ associated with the original equilibrium and create a new vector of transfers $\mathbf{b}_2 = (b^2_{1,2}, b^2_{2,1})$ such that:

$$\mathcal{C}_1^2 + b^2_{2,1} - b^2_{1,2} > \mathcal{C}_1^1 + b^1_{2,1} - b^1_{1,2}, \tag{2.12}$$

$$\mathcal{C}_2^2 + b^2_{1,2} - b^2_{2,1} = \mathcal{C}_2^1 + b^1_{1,2} - b^1_{2,1}. \tag{2.13}$$

Because $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1 > 0$, finding payments that satisfy $b^2_{1,2} \leq \mathcal{C}_1^2$ and $b^2_{2,1} \leq \mathcal{C}_2^2$ is always feasible.

Note that these changes have no impact on player 1's choices of actions made in any period $t' \leq t-1$ because all actions are observable, and hence choosing a different action from the proposed equilibrium would be labeled a defection. If defections were deterred in the original equilibrium, which had a strictly smaller continuation value for player 1, then they are also deterred in the new equilibrium. The same logic applies to player 2 since they obtain the same expected payoff in period $t-1$ (compared to the original equilibrium), and thus also have the same continuation values in all periods $t' < t-1$. Finally, note that surplus from a date 0 perspective is strictly higher under the new equilibrium.

Step 2 We show that there exists a relational contract that implements an experimentation policy $\hat{\mathbf{P}}(\cdot)$ if and only if the following inequality holds for all t and for all histories $h^t \in \mathcal{H}^t$:

$$\sum_{p \in \hat{\mathbf{P}}^t} \sum_{i=1,2} \max\Big(0, c - \mathbb{E}(x_p \cdot v_p | h^t)\Big) \leq \mathcal{C}(h^t), \tag{2.14}$$

where $\mathcal{C}(h^t)$ is the continuation value.

To show that (2.14) is a necessary and sufficient condition, consider a set of transfers $b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t) \geq 0$ to be paid on path given a vector of realized values $\mathbf{x}^t, \mathbf{v}^t$.

Given an equilibrium experimentation policy $\mathbf{P}^t$, note that it is without loss of generality to assume that $P_1^t = P_2^t = \mathbf{P}^t$. Thus, for each player and for each $p \in \mathbf{P}^t$, the player must weakly prefer to include p in P_i^t, rather than excluding it. Let $\sigma_i(\mathbf{x}^t, \mathbf{v}^t)$ denote player i's share of $\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)$ as a function of $\mathbf{x}^t, \mathbf{v}^t$. Hence, the condition for selecting $\mathbf{P}^t$ is:

$$\begin{aligned}\sum_{p \in \mathbf{P}^t} \max\left(c - \mathbb{E}(x_p v_p | h^t), 0\right) \leq \mathbb{E}\Big(& b_{-i,i}(\mathbf{x}^t, \mathbf{v}^t) - b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t) \\ & + \sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)\Big), \quad \forall i,\end{aligned} \tag{2.15}$$

$$b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t) \leq \sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t\mathbf{x}^t, \sqcup\mathbf{v}^t), \quad \forall \mathbf{v}^t, \forall i. \tag{2.16}$$

Expectations are taken over the project valuations realizations $\mathbf{x}^t, \mathbf{v}^t$ and $h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t$ denotes the players' updated beliefs after observing $\mathbf{x}^t, \mathbf{v}^t$.[18] The first expression states that the

[18]The history also includes the project selections, and both the upfront and end-of-period transfers. However, for notational convenience we only include the realized valuations as every other object can be inferred on path from the realized valuations.

promised transfers and the expected share of the total continuation value must be enough to prevent a player from shirking on any subset of the projects. The second expression states that the each player is willing to pay the other player the necessary transfer.

To show necessity: Note that since Equation (2.15) must hold for a fixed i, the inequality also holds summing over all i. Further, all transfers cancel out when summing over i. Finally, by definition, $\mathbb{E}(\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)) = \mathcal{C}(h^t)$. Hence, we are left with Equation (2.14).

To show sufficiency: We will show this result in two substeps.

SubStep 1: We show it is necessary and sufficient to replace Equation (2.16) by its expectation. This new expression is as follows:

$$\mathbb{E}(b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t)) \leq \mathbb{E}\Big(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)\Big) \quad \forall i. \tag{2.17}$$

We first show that if there is a solution to Equations (2.17) and (2.15), then there exists a solution to Equations (2.16) and (2.15).

Take a set of transfers $b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t)$ that satisfy Equations (2.17) and (2.15). Define:

$$\begin{aligned} b'_{i,-i}(\mathbf{x}^t, \mathbf{v}^t) =& \sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t) \\ &- \Bigg(\mathbb{E}\Big(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t) - b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t)\Big)\Bigg). \end{aligned} \tag{2.18}$$

Since Equation (2.17) holds, the term in the expectation of Equation (2.18) is positive and thus Equation (2.16) holds for all realizations of $\mathbf{x}^t, \mathbf{v}^t$ under the set of transfers $b'_{i,-i}(\mathbf{x}^t, \mathbf{v}^t)$. Finally, $\mathbb{E}(b'_{i,-i}(\mathbf{x}^t, \mathbf{v}^t)) = \mathbb{E}(b_{i,-i}(\mathbf{x}^t, \mathbf{v}^t))$ so Equation (2.17) continues to hold.

SubStep 2: Using substep 1, it suffices to show that Equation (2.14) implies a solution to Equations (2.15) and (2.17). To simplify all the notation with expectations, Equation (2.15) can be re-expressed as:

$$\beta_i - \gamma_i \leq (\tilde{b}_{-i,i} - \tilde{b}_{i,-i}), \tag{2.19}$$

where $\tilde{b}_{i,-i}$ is the expected transfer from i to $-i$, $\beta_i = \sum_{p \in \mathbf{P}^t} \max\big(0, c - \mathbb{E}(x_p v_p | h^t)\big)$, and $\gamma_i = \mathbb{E}(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t))$. Equation (2.17) can thus be re-written as:

$$\tilde{b}_{i,-i} \leq \gamma_i. \tag{2.20}$$

Rearranging Equation (2.14) implies $\sum_i (\beta_i - \gamma_i) \leq 0$. One can now show that $\tilde{b}_{i,-i} = \max(0, \beta_{-i} - \gamma_{-i})$ satisfies Equation (2.20). Further, Equation (2.19) holds because:

$$\beta_i - \gamma_i \leq \max(0, \beta_i - \gamma_i) - \max(0, \beta_{-i} - \gamma_{-i}) \tag{2.21}$$

$$\iff \max(0, \beta_{-i} - \gamma_{-i}) - \min(0, \gamma_i - \beta_i) \leq 0 \tag{2.22}$$

$$\Longleftarrow \sum_i (\beta_i - \gamma_i) \leq 0, \tag{2.23}$$

where the final step follows from noting that $\beta_1 - \gamma_1$ and $\beta_2 - \gamma_2$ cannot both be positive and analyzing the remaining three cases based on the signs of $\beta_i - \gamma_i$.

Finally, Equation (2.20) reduces to

$$\max(0, \beta_{-i} - \gamma_{-i}) \leq \gamma_i \iff \beta_{-i} - \gamma_{-i} \leq \gamma_i \tag{2.24}$$

$$\iff \sum_i (\beta_i - \gamma_i) \leq 0, \tag{2.25}$$

where the final implication is due to β_i being weakly positive.

Step 3: We show that any relational contract that implements a given experimentation policy can be replaced by an alternative relational contract that implements the same experimentation policy and yields no surplus to player 2. First, note that the way the players share their continuation value does not affect Equation (2) from the main text. Hence, for any period t where player 2's expected payoff is positive, $w_{2,1}$ can be increased until player 2's expected payoff is zero. Player 2 is willing to make this transfer because not doing so would be seen as a deviation, resulting in a payoff of 0 for player 2.

Step 4: We now show that, for any two histories h_1^t and $h_2^{t'}$ that generate the same beliefs μ, selecting the same continuation equilibrium is without loss of optimality. Take a relational contract r that is surplus-maximizing at all on-path histories and has two histories h_1^t and $h_2^{t'}$ prescribing different (surplus-maximizing) continuation equilibria under the same beliefs μ. Recall from Step 3 that one can consider relational contracts in which player 2 obtains an expected payoff equal to 0 in every period. In this case, since the two continuation equilibria are both optimal and both give all the surplus to player 1, switching from one continuation equilibrium to the other does not change the players' incentives as both prescribe the exact same payoffs to the players. Hence, when focusing on relational contracts that specify the same continuation equilibrium following histories that induce the same beliefs, one can replace $\mathcal{C}(h^t)$ with $\mathcal{C}(\mu^t)$. □

Proof of Proposition 3. When the players have identified projects with values $\hat{v}_1, \ldots, \hat{v}_m$ at history h, the condition for the players being able to replicate the first-best experimentation policy in all subsequent periods is that, for all histories h' occurring after h and with associated project values $\hat{v}'_1, \ldots, \hat{v}'_m$, the players exploit $\hat{v}'_j$ if and only if $\hat{v}'_j \geq v^0$. This condition is as follows:

$$c \sum_{j=1}^{m} 1_{\hat{v}'_j \geq v^0} + \max\{0, c - (1-\alpha)\mathbb{E}(v_p)\} \sum_{j=1}^{m} 1_{\hat{v}'_j < v^0} \leq \sum_{j=1}^{m} \mathcal{C}^0(\hat{v}'_j), \tag{2.26}$$

$\forall (\hat{v}'_1, \ldots, \hat{v}'_m) \geq (\hat{v}_1, \ldots, \hat{v}_m)$, which corresponds to (2.2) when the players implement the first-best policy and where $\mathcal{C}^0(\hat{v}'_j)$ denotes the continuation value associated with domain j under the first-best policy. Note that $\mathcal{C}^0(\hat{v}'_j)$ (i) is constant below v^0, (ii) is such that $\lim_{x \uparrow v^0} \mathcal{C}^0(x) > \lim_{x \downarrow v^0} \mathcal{C}^0(x)$ and (iii) is increasing above v^0. Given such properties, setting $\hat{v}'_j = \max\{\hat{v}_j, v^0\}$ both minimizes the right-hand side and maximizes the left-hand side of (2.26). Thus, an equivalent condition is:

$$m \cdot c \leq \delta \Big(\sum_{j=1}^{m} \frac{1}{1-\delta} (\max\{\hat{v}_j, v^0\} - 2c) \Big). \tag{2.27}$$

Finally, the existence of a threshold δ^0 was proven in the text. □

Proof of Corollary 2. The characterization of $v^*(\delta)$ follows from Corollary 1. The existence of δ^* follows an identical argument to that made in Proposition 5. Suppose $\delta < \delta^0$ and consider the policy described in the corollary. By definition of $\tilde{v}$, it suffices to check that the policy is implementable in period 1 (i.e., satisfies Inequality (2.4)), which is most binding when $\alpha = 1$ (henceforth assumed). At $\delta = \delta^0$, this constraint is as follows: $c \leq \mathcal{C}^0(\text{explore})$, where $\mathcal{C}^0(\text{explore})$ is defined by:

$$\frac{v^0 - 2c}{1-\delta} = \mathbb{E}(v_p - 2c) + \mathcal{C}^0(\text{explore}) \implies \mathcal{C}^0(\text{explore}) > \frac{\delta}{1-\delta}(v^0 - 2c). \tag{2.28}$$

δ^0 is defined by: $c = \frac{\delta}{1-\delta}(v^0 - 2c)$. Therefore, for δ slightly below δ^0, the policy in period 1 satisfies Inequality (2.4) if and only if the continuation value is continuous at δ^0, which holds since $\tilde{v}$ is continuous with respect to δ. □

Proof of Proposition 4. Denote $t(p) = \inf_t\{t : p \in \mathbf{P}^t\}$. By contradiction, $\forall p \in \mathcal{P}$, either (i) $p \in \mathbf{P}^t \; \forall \, t > t(p)$ or (ii) $p \notin \mathbf{P}^t \; \forall \, t > t(p)$. Further, by monotonicity, for each domain j, there exists a threshold $v_j^*(\hat{\mathbf{v}}_{-j})$ such that the players exploit a project with value $\hat{v}_j$ if and only if $\hat{v}_j \geq v_j^*(\hat{\mathbf{v}}_{-j})$, where $\hat{\mathbf{v}}_{-j}$ denotes the values of the best projects found in the remaining domains.

Note that $v_j^*(\cdot)$ is weakly increasing in each of its arguments; otherwise, with positive probability, statement 2 of the proposition would be satisfied. Further, for a sufficiently large $\hat{\mathbf{v}}_{-j}$, the first-best experimentation policy is implementable (Proposition 3), implying that $v_j^*(\hat{\mathbf{v}}_{-j}) \leq v^0$. Therefore, with positive probability, the players permanently exploit only projects with value weakly less than v^0, which would imply (i) $v^0 \geq \tilde{v}$ and, thus, (ii) that $\delta \geq \delta^0$. □

The following lemma will aid in proving the next two propositions. Let F denote a distribution of v_p with finite support. Consider a sequence of continuous approximations F_n such that $F_n \leq_{\text{F.O.S.D}} F_{n-1}$, $F_n \geq_{\text{F.O.S.D}} F \; \forall n$, and $F_n \to F$. Define the optimal experimentation policy $\mathbf{a}(\cdot)$ as *strict* if the following conditions hold for all $\hat{\mathbf{v}}$: (i) $\mathbf{a}(\hat{\mathbf{v}})$ satisfies Inequality (2.4) strictly, (ii) if $\mathbf{a}'(\hat{\mathbf{v}})$ is preferred to $\mathbf{a}(\hat{\mathbf{v}})$, then $\mathbf{a}'(\hat{\mathbf{v}})$ fails Inequality (2.4) strictly, and (iii) if $\mathbf{a}'(\hat{\mathbf{v}})$ satisfies Inequality (2.4), the players strictly prefer $\mathbf{a}(\hat{\mathbf{v}})$ over $\mathbf{a}'(\hat{\mathbf{v}})$. Let $B^n(\cdot)$ denote the associated Bellman equation with F^n, and $\mathbf{a}^n(\cdot)$ the corresponding optimal experimentation policy.

Lemma 1 (Discretization)
For any $\hat{\mathbf{v}} \in F^m$, if the optimal experimentation policy is strict, then (i) $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$ and (ii) for all $\hat{\mathbf{v}} \in F^m$, $\mathbf{a}^n(\hat{\mathbf{v}}) \to \mathbf{a}(\hat{\mathbf{v}})$.

Proof of Lemma 1. First note that (i) $\implies$ (ii). By contradiction, suppose $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$ for all $\hat{\mathbf{v}}$ but there exists a $\hat{\mathbf{v}}^*$ such that $\mathbf{a}^n(\hat{\mathbf{v}}^*) \not\to \mathbf{a}(\hat{\mathbf{v}}^*)$. As $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$ for all $\hat{\mathbf{v}}$, then $\mathcal{C}^n(\hat{\mathbf{v}}^*, \mathbf{a}) \to \mathcal{C}(\hat{\mathbf{v}}^*, \mathbf{a})$ uniformly with respect to $\mathbf{a}$ (as $\mathbf{a}$ belongs to a finite set). However, given that the preference at $\hat{\mathbf{v}}^*$ is strict, we must have $\mathbf{a}^n(\hat{\mathbf{v}}) \to \mathbf{a}(\hat{\mathbf{v}})$, a contradiction.

Let us now prove (i) by contradiction. Note that $\hat{\mathbf{v}}$ has a lattice structure. If (i) fails, there exists a $\hat{\mathbf{v}}^*$ such that $B^n(\hat{\mathbf{v}}^*) \not\to B(\hat{\mathbf{v}}^*)$ but $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$ for any $\hat{\mathbf{v}} > \hat{\mathbf{v}}^*$.

Claim 1: Denote by $\bar{v} = \sup \text{Support}\{v_p\}$, then $\hat{\mathbf{v}}^*$ cannot correspond to $\bar{v}, \ldots, \bar{v}$. Note that $\bar{v} > v^0$. Therefore, there exists an n^* for which, if $n > n^*$, the players permanently exploit in all domains when $\hat{\mathbf{v}} = \bar{v}, \ldots, \bar{v}$. Claim 1 follows by noting that the net-present value of this policy is identical under $B^n(\cdot)$ and $B(\cdot)$.

Let us now consider $\hat{\mathbf{v}}^* \neq \bar{v}, \ldots, \bar{v}$. By assumption, $\forall \hat{\mathbf{v}} > \hat{\mathbf{v}}^*$ $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$. Further, as $B^n(\cdot)$ is decreasing and bounded below by $B(\cdot)$, $B^n(\hat{\mathbf{v}}^*) \not\to B(\hat{\mathbf{v}}^*) \implies B(\hat{\mathbf{v}}^*) < \lim B^n(\hat{\mathbf{v}}^*)$. Next, as $B^n(\hat{\mathbf{v}}) \to B(\hat{\mathbf{v}})$ $\forall \hat{\mathbf{v}} > \hat{\mathbf{v}}^*$, then $\mathcal{C}^n(\hat{\mathbf{v}}^*, \mathbf{a}) \to \mathcal{C}(\hat{\mathbf{v}}^*, \mathbf{a})$. As $B(\hat{\mathbf{v}}^*) < \lim B^n(\hat{\mathbf{v}}^*)$, then $\lim \mathbf{a}^n(\hat{\mathbf{v}}^*) \neq \mathbf{a}(\hat{\mathbf{v}}^*)$. Given that the continuation value converges, $\mathbf{a}^n(\hat{\mathbf{v}}^*)$ must converge, implying $\mathbf{a}(\hat{\mathbf{v}}^*) \neq \mathbf{a}'(\hat{\mathbf{v}}^*) := \lim \mathbf{a}^n(\hat{\mathbf{v}})$. Therefore, $\mathbf{a}'(\hat{\mathbf{v}}^*)$ must be strictly preferred to $\mathbf{a}(\hat{\mathbf{v}}^*)$. Given our definition of "strict," $\mathbf{a}'(\hat{\mathbf{v}}^*)$ must strictly fail Inequality (2.3) for $B(\hat{\mathbf{v}}^*)$. However, this leads to a contradiction because the continuation value has been proven to converge, implying that $\mathbf{a}'(\hat{\mathbf{v}}^*)$ cannot satisfy Inequality (2.3) as $n \to \infty$. □

Proof of Proposition 5. We first prove the existence of δ^*. Suppose $\delta_1 < \delta_2$ and, by contradiction, that the optimal experimentation policy is non-empty for δ_1 but empty for δ_2. The optimal experimentation policy for δ_1 yields strictly positive surplus and yet cannot be implemented at δ_2. However, holding fixed the policy, the left-hand side of (2.2) is independent of δ and the right-hand side is increasing in δ, implying that the experimentation policy is feasible under δ_2, which is a contradiction. This reasoning implies that a threshold exists. Finally, $\delta^* < 1$ since $\mathcal{C}(\cdot) \to \infty$ as $\delta \to 1$.

We now prove the existence of $\bar{\delta}$. Scope is initially maximal if and only if $m \cdot \max\{0, c - (1-\alpha)\mathbb{E}(v_p)\} \leq \mathcal{C}(\mu^1)$. However, by an identical argument as that in the preceding paragraph, the right-hand side of (2.2) is increasing in δ and the left-hand side of (2.2) is independent of δ. This implies the existence of a threshold on δ. Further, when $\delta \to 1$, $\mathcal{C}(\mu^1) \to \infty$, implying maximal scope. As a result, $\bar{\delta} < 1$.

$\delta^* \leq \bar{\delta}$ because any initially maximal relational contract is non-empty. We now show that this inequality is strict when m is sufficiently large. Inequality (2.2) implies

$$m(c - (1-\alpha)\mathbb{E}(v_p)) \leq B\left(\mathbf{0}|\bar{\delta}(m)\right) - m\mathbb{E}\left(v_p - 2c\right). \tag{2.29}$$

Next, note that $\lim_{\delta \uparrow \bar{\delta}(m)} B\left(\tilde{v}, 0, \ldots, 0|\delta\right) \geq B\left(\mathbf{0}|\bar{\delta}(m)\right)$. We construct a suboptimal, initially limited policy in which, during period 1, players explore a single domain in search of a project with value exceeding $\tilde{v}$. If they find such a project, they expand their scope; otherwise, they terminate their relationship. $\delta^*(m) = \bar{\delta}(m)$ implies that this policy cannot be implemented for $\delta < \bar{\delta}(m)$. Therefore:

$$c - (1-\alpha)\mathbb{E}(v_p) \geq \bar{\delta}(m)\Pr(v_p > \tilde{v})B(\mathbf{0}|\bar{\delta}(m)). \tag{2.30}$$

Combining (2.29) and (2.30) implies:

$$\begin{aligned} &\bar{\delta}(m)\Pr(v_p > \tilde{v})B(\mathbf{0}|\bar{\delta}(m)) \leq \frac{B(\mathbf{0}|\bar{\delta}(m))}{m} - \mathbb{E}(v_p - 2c) \\ \iff\ &B(\mathbf{0}|\bar{\delta}(m))\left(\bar{\delta}(m)\Pr(v_p > \tilde{v}) - \frac{1}{m}\right) \leq -\mathbb{E}(v_p - 2c). \end{aligned} \tag{2.31}$$

However, as shown in the text $B(\mathbf{0}|\bar{\delta}(m))$ diverges to infinity. Hence, $\bar{\delta}(m)\Pr(v_p > \tilde{v}) - \frac{1}{m}$ converges to zero. However (i) $\bar{\delta}(m)$ remains bounded away from zero when $\mathbb{E}(v_p)(1-\alpha) < c$, which has been assumed, and (ii) if $\bar{\delta}(m)$ remains bounded away from zero, then $\tilde{v}$ remains bounded above, implying $\Pr(v_p > \tilde{v})$ remains bounded away from zero, proving that $\bar{\delta}^*(m) < \bar{\delta}(m)$ when m is large.

Finally, we consider a discrete support distribution and leverage Lemma 1 to show that $\delta^* = \bar{\delta}$. We consider a three-point support of benefits: $\{0, \underline{v}, \bar{v}\}$. Further, the experimentation policy described below will be shown to satisfy the definition of "strict" employed in Lemma 1. We also assume $\mathbb{E}(v_p) = 2c$, $\alpha = 1$, and $m = 2$ when listing sufficient inequalities for $\delta^* = \bar{\delta}$ to hold. The following inequalities ensure that there exists a δ such that $|\mathbf{P}^1| = 1$ is not feasible but $|\mathbf{P}^1| = 2$ is:

$$\frac{\underline{v} - 2c}{1-\delta} > \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} \tag{2.32}$$

$$2c < \frac{\delta}{1-\delta}(\bar{v} - 2c) \tag{2.33}$$

$$c > \frac{\delta}{1-\delta}(\underline{v} - 2c) \tag{2.34}$$

$$2c > \frac{\delta}{1-\delta}(\underline{v} - 2c) + \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} \tag{2.35}$$

$$c > \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} + \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}\mathrm{npv}^0 \tag{2.36}$$

$$\begin{aligned} 2c < {} & \frac{2\Pr(\bar{v})(1 - \Pr(\bar{v}))}{1 - \delta\big(1 - 2\Pr(\bar{v})(1 - \Pr(\bar{v}))\big)}\left((\bar{v} - 2c)\frac{\delta}{1-\delta} + \mathrm{npv}^0\right) \\ & + \frac{\Pr(\bar{v})^2}{1 - \delta\big(1 - \Pr(\bar{v})^2\big)}\frac{\delta}{1-\delta}2(\bar{v} - 2c) \end{aligned} \tag{2.37}$$

$$\mathrm{npv}^0 = \frac{\delta}{1-\delta}\frac{\Big(\Pr(\bar{v}) + \Pr(\underline{v})\Big)\left(\frac{\Pr(\bar{v})\bar{v}}{\Pr(\bar{v}) + \Pr(\underline{v})} + \frac{\Pr(\underline{v})\underline{v}}{\Pr(\bar{v}) + \Pr(\underline{v})} - 2c\right)}{1 - \delta(1 - \Pr(\bar{v}) - \Pr(\underline{v}))} \tag{2.38}$$

Inequality (2.32) ensures that $v^0 \leq \underline{v}$. Inequality (2.33) ensures that $\bar{v}, 0$ satisfies Inequality (2.5). Inequality (2.34) ensures that the players are unable to exploit a project worth $\underline{v}$ in isolation and, by extension, cannot jointly exploit two such projects. Inequality (2.35) implies that the players cannot exploit a project worth $\underline{v}$ while exploring in the additional domain. Inequality (2.36) implies that the players would be unable to begin exploring if the players began their exploration on one domain and, upon finding a project worth $\bar{v}$, started exploring the additional domain.[19] The continuation value under this experimentation policy is bounded above by npv^0: the net-present value of a single domain under the first-best experimentation policy, the value of which is stated in (2.38). Finally, Inequality (2.37) is a necessary condition for the initially maximal policy to be feasible at date 1. This condition is necessary, but not sufficient, as the continuation value is computed assuming that if the players discover one project worth $\bar{v}$ before doing so on the other domain, the highest-valued project on the other domain is zero. We use Mathematica to show that these constraints jointly hold strictly.[20] □

[19] Further, it will never be optimal to explore the additional domain upon finding $\underline{v}$, since such a project cannot be exploited.

[20] Code available upon request.

Proof of Proposition 6. We first show that our notion of terminal scope is well defined. If there exists a period t for which the players conduct no explorations, then $\mathbf{P}^{t'} = \mathbf{P}^t$ for all $t' \geq t$. By contradiction, if there exists an equilibrium path where $\liminf |\mathbf{P}^t| < \limsup |\mathbf{P}^t|$, the players must explore at least one project in each period t. However, for each exploration, with positive probability the players discover a project with value exceeding $m\tilde{v}$, implying that the first-best policy is implementable in all subsequent periods. As a result, $\liminf |\mathbf{P}^t| = \limsup |\mathbf{P}^t|$. This argument also shows that terminal scope equals m with positive probability.

Statement 1: We consider $v_p \in \{0, \underline{v}, \bar{v}\}$, $m = 2$, and $\alpha = 1$. Further, the experimentation policy outlined below will satisfy the definition of "strictness" employed in Lemma 1, implying these results will hold true for continuous approximations. These inequalities ensure the existence of a feasible experimentation policy where the scope of experimentation reaches its maximum of 2 with interior probability, while ensuring no other feasible policy yields a higher joint surplus. We list all the inequalities and comment on each one separately below.

$$2c < \frac{\delta}{1-\delta}(\bar{v} - 2c) + \mathcal{C}^0(\text{explore}) \tag{2.39}$$

$$c < \frac{\delta}{1-\delta}(\underline{v} - 2c) \tag{2.40}$$

$$2c > \frac{\delta}{1-\delta}(\underline{v} - 2c) + \mathcal{C}^0(\text{explore}) \tag{2.41}$$

$$2c > \mathcal{C}^0(\text{explore}) + \frac{\delta}{1-\delta}\Big(\Pr(\bar{v})\bar{v} + (1 - \Pr(\bar{v}))\underline{v} - 2c\Big) \tag{2.42}$$

$$\begin{aligned} \frac{\underline{v} - 2c}{1-\delta} > v := \mathbb{E}(v_p - 2c) + \frac{\delta}{1-\delta}\Big(\Pr(\bar{v})(\bar{v} - 2c) + \Pr(\underline{v})(\underline{v} - 2c)\Big) \\ + \Big(\Pr(\bar{v}) + \Pr(\underline{v})\Big)\frac{\delta}{1-\delta}(\underline{v} - 2c) + \Big(1 - \Pr(\bar{v}) + \Pr(\underline{v})\Big)\delta v \end{aligned} \tag{2.43}$$

$$c \leq \mathcal{C}^0(\text{explore}) + \frac{\delta \Pr(\bar{v})\mathcal{C}^0(\text{explore})}{1 - \delta(1 - \Pr(\bar{v}))} \tag{2.44}$$

Inequality (2.39) implies that $\{\bar{v}, 0\}$ satisfy Equation (2.5), where $\mathcal{C}^0(\text{explore})$ was defined in Equation (2.28). Inequality (2.40) ensures that the players are able to exploit a project worth $\underline{v}$ in isolation. Inequality (2.41) ensures that $|\mathbf{P}^t| < 2$ while exploiting the project worth $\underline{v}$ (when the best project found so far on the other domain has value 0). This inequality uses $\mathcal{C}^0(\text{explore})$ as an upper-bound. These statements imply that if the players ever reach a point with a project worth $\underline{v}$, they either exploit the project, explore a project on the other domain while maintaining a scope of 1, or conduct 2 explorations. Inequality (2.42) ensures that conducting two explorations is not feasible because the upper-bounds associated with the continuation value for the new domain and the domain with a project with value $\underline{v}$ is provided by the first-best policy. Next, Inequality (2.43) ensures that the players prefer to exploit the project worth $\underline{v}$ as opposed to exploring the domain where the best project is worth 0 until Equation (2.5) holds and then subsequently implementing the first-best policy. These constraints imply that $|\mathbf{P}^t| = 1$ if the best projects are worth $\underline{v}, 0$. Finally, Inequality (2.44) ensures that this experimentation policy is feasible. One can check that these constraints, along with (i) $\mathbb{E}(v_p) \geq 2c$ and (ii) $v^0 \leq \underline{v}$, hold jointly.[21]

[21] The code can be provided upon request.

Statement 2: We prove the result for a discrete support distribution, as Lemma 1 can be used to extend the result to a convex support distribution. We consider $m = 3$ and a trinary support distribution, $v_p \in \{0, \underline{v}, \bar{v}\}$, where $0 < \underline{v} < \bar{v}$.[22] Throughout, let $\tilde{c} := c - (1-\alpha)\mathbb{E}(v_p) > 0$. Let $\bar{p}, \underline{p}$ correspond to the probability that $v_p = \bar{v}$, $\underline{v}$, respectively. For the subsequent argument, consider $\bar{p}$ to be arbitrarily small, in a sense we will make precise below. Suppose

$$2c = \frac{\delta}{1-\delta}(\bar{v} + \underline{v} - 4c), \tag{2.45}$$

implying that the players can jointly exploit projects worth $\bar{v}$ and $\underline{v}$, but could not permanently exploit a project worth $\bar{v}$ and two projects worth $\underline{v}$. As a result, for any $\tilde{c} > 0$, when $\hat{\mathbf{v}} = \bar{v}, \underline{v}, 0$, the players either (i) permanently exploit the two projects and conduct no additional explorations or (ii) implement a policy involving exploration whose associated surplus is bounded below that of exploiting the project with value $\bar{v}$ and exploring in the other two domains until finding a project with value $\bar{v}$ and subsequently implementing the first best. As the payoff bound of (ii) tends to zero as $\bar{p}$ goes to zero, there exists $\bar{p}^* > 0$ such that for $\bar{p} < \bar{p}^*$, the players choose (i). Hence, it suffices to show that Equation (2.45), $\bar{p} < \bar{p}^*$, and $|P^1| = m$ may jointly hold. Because the first two of these three conditions are independent of α, they are also independent of $\tilde{c}$. Therefore, holding fixed all remaining parameter values, the continuation value at date one of conducting three explorations is bounded below by $\Pr(v_p = \bar{v})\frac{\delta}{1-\delta}(\bar{v} - 2c)$. As a result, if $\tilde{c}$ is sufficiently small, the players' initial scope will be maximal, thereby completing the proof. □

9 Appendix B: Online

9.1 Complete Characterization with Binary Support

Let $m = 2$ and consider a binary support distribution: $v_p \in \{\underline{v}, \bar{v}\}$ where $\bar{v} > \max\{\underline{v}, \tilde{v}\}$. We say a project is "suitable for exploitation" if and only if it has value $\bar{v}$. One can check that if $\alpha = \frac{1}{2}$ the players can always implement the first-best. For conciseness, we therefore assume $\alpha = 1$.

Proposition A1 *When $m = 2$ there are three possible non-empty experimentation policies which are optimal for a non-empty region of parameter values:*

1. ***First Best:*** *The players' scope of experimentation is maximal in all periods and, in each domain, the players explore projects until finding one suitable for exploitation.*

2. ***Immediate Exploitation:*** *The players explore projects in one domain and, upon finding a project suitable for exploitation, they permanently exploit said project and immediately begin the exploration of projects in the other domain until identifying one suitable for exploitation.*

[22] Unlike Proposition 5, considering a trinary support distribution and $m = 2$ fails.

3. ***Delayed Exploitation:*** *The players explore projects in one domain and, upon finding a project suitable for exploitation, they idle on that domain and explore projects on the other domain until finding a project suitable for exploitation. Subsequently the players permanently exploit the two projects they have found which are suitable for exploitation.*

The players always prefer (1) to (2) and (2) to (3). For this reason, this proposition fully characterizes the optimal experimentation policy, since given any set of parameter values one can check which of these three policies is feasible.

Proof. The proof proceeds in two steps. First, we argue that no other experimentation policy may be optimal. Next, we use backwards induction to verify that these three policies are indeed optimal for a non-empty region of parameter values.

By symmetry, we can consider the experimentation policy as a function of the number of projects suitable for exploitation, n. We first prove two results: scope is increasing with respect to n and scope is weakly larger than n.

To prove the former, note that any project selection rule that the players can implement when they have identified $n-1$ projects suitable for exploitation is also implementable when they have identified n such projects. This occurs because the continuation value of their relationship is strictly increasing in the number of projects suitable for exploitation. Additionally, exploiting a project suitable for exploitation results in a positive net present value. Hence, the scope of the players' relationship can never shrink on path.

The proof for the second statement notes that (i) by assumption, the players can always exploit n projects suitable for exploitation, (ii) the continuation value following exploration is weakly higher than that of exploitation, and (iii) the players maximize scope because $\mathbb{E}(v_p) \geq 2c$. Combining these observations implies that the players' scope of experimentation always exceeds n.

Next, note that the players will neither exploit a project which is not suitable for exploitation (exploration provides higher current period profits and relaxes the implementability constraint), nor explore on a domain for which they have discovered a project suitable for exploitation (exploitation delivers higher current period profits, and both exploration and exploitation lead to an identical implementability constraint). Combining these observations with the result that scope cannot decrease implies that, if the players begin with maximal scope, they can implement the first best. Therefore, it suffices to analyze experimentation policies where $|\mathbf{P}^1| = 1$, and to determine what the experimentation strategy dictates when $n = 1$. If when $n = 1$ the players maintain a scope of 1 and exploit this project, then the players never utilize the second domain. This cannot be optimal: by the logic of Bernheim and Whinston (1990), the players could instead implement the first best. If the players expand their scope, then this is necessarily the second policy in the proposition. If the players maintain a scope of 1, then this is the third policy in the proposition.

For the second step of the proposition, note that the first-best is optimal as $\delta \to 1$. Finally, one can show that both (2) and (3) are optimal by computation: for instance, upon setting $\underline{v} = 0, \bar{v} = 4, c = 1, \frac{\delta}{1-\delta}(\bar{v} - 2c) = 1.01c$, there will exist two different values of $\Pr(v_p = \bar{v})$ for which (2) and (3) are optimal for an open set of parameter values around these points.[23]

[23]One could also instead prove that (2) is always implementable for a strictly larger region of discount factors than (1) to show that (2) is optimal.

□

9.2 Extensions

In this section we provide the proofs for the extensions outlined in Section 5.3 of the text.

9.2.1 Risky Collaborations

We make the following simplifications: $\alpha = 1$, $\mathbb{E}(v_p) = 2c$, and $m = 2$. Let the distribution of benefits on domain 1 be as in the main analysis. On domain 2 there exists only one project with value $v_r \in \{0, v\}$, where $\Pr(v_r = v) = q$ and $v \geq 2\tilde{v}$, ensuring that upon discovery of a project with value v the players can implement the first-best policy on domain 1.

Proposition A2 *For an open set of parameter values: (i) the players explore the domain 2 project after having first discovered a project with value exceeding $\tilde{v}$ on domain 1, (ii) despite being able to explore both domains in period 1.*

Proof of Proposition A2. We now provide the necessary conditions for the statement in the proposition to hold:

$$c \leq q\frac{\delta}{1-\delta}(v - 2c) \tag{2.46}$$

$$c > \frac{\Pr(v_p \geq \tilde{v})\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})}\frac{1}{1-\delta}\mathbb{E}(v_p - 2c | v_p \geq \tilde{v}) \tag{2.47}$$

$$c < \frac{\Pr(v_p \geq \tilde{v})\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})}\Big(\frac{1}{1-\delta}\mathbb{E}(v_p - 2c | v_p \geq \tilde{v}) + qv - 2c + q\frac{\delta}{1-\delta}(v - 2c)\Big) \tag{2.48}$$

These conditions state, respectively, that domain 2 could be started on its own, that domain 1 cannot be started on its own (or, equivalently, with knowledge that the domain 2 project is worth zero), and that domain 1 can be started when anticipating that domain 2 is started after discovering a project with value at least $\tilde{v}$. These conditions imply that the optimal experimentation policy takes one of two forms: (1) Explore projects on both domains, and if the risky project fails, then permanently idle on domain 1 if a project with value $\tilde{v}$ is not discovered in the first period. If the risky project has value v, then the experimentation policy becomes efficient on domain 1. (2) Explore domain 1 and only explore domain 2 once finding a project with value at least $\tilde{v}$. As policies (1) and (2) are both feasible, the players choose the policy with a higher net present value. Exploring both domains implies a surplus of:

$$\begin{aligned} &qv - 2c + q\frac{\delta}{1-\delta}(v - 2c) \\ &+ (1-q)\frac{\delta}{1-\delta}\mathbb{E}((v_p - 2c)1_{v_p \geq \tilde{v}}) + q\frac{\Pr(v_p > v^0)\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})}\frac{1}{1-\delta}\mathbb{E}(v_p - 2c | v_p \geq v^0), \end{aligned} \tag{2.49}$$

where the first line corresponds to the surplus from domain 2 and the second line corresponds to the surplus from domain 1. Further, the surplus from exploring domain 1 and later exploring domain 2 is exactly Equation (2.48) as the expected surplus in period 1 is equal to zero. Let us hold fixed the surplus from domain 2 and δ, but let us consider varying

q while v remains fixed to maintain the same expected surplus. When $q = 1$, there is no uncertainty and the surplus from exploring both domains at once can be shown to be strictly greater than delaying. When domain 2 is explored after domain 1, changing q but holding the expected surplus from domain 2 fixed has no effect on the total expected surplus. In contrast, when exploring both domains at once, a decrease in q increases the probability that domain 1 is never explored. As q continues to decrease, the surplus from exploring both domains at once converges to the expected surplus from domain 2 plus the probability that the first drawn project in domain 1 exceeds $\tilde{v}$, multiplied by the profitability of exploiting such a project. In contrast, the surplus from a sequential approach converges to the surplus obtained with certainty in domain 1, plus the expected profitability of domain 2 when explored after identifying a suitable project in domain 1. One can see that the players may delay domain 2, even if it is more profitable than domain 1, so as to prevent permanently idling on domain 1. □

9.2.2 Symmetric Benefits

Let us first consider a case where in one domain the benefits are symmetric but in the other domain benefits are asymmetric. Namely, for domain 1 all projects equally benefit both parties but domain 2 is as in our main analysis.

Proposition A3 *Let δ^* and $\bar{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \bar{\delta}$. In other words, for intermediate discount factors, the scope of experimentation is initially limited on path, with scope increasing with strictly positive probability along the equilibrium path.*

Proof of Proposition A3. Observe that $\delta^* = 0$ because an experimentation policy consisting of (i) exploration in the symmetric benefits domain and (ii) expanding scope once a project exceeding $\tilde{v}$ is discovered in the first domain is always feasible. Further, $\bar{\delta} > 0$ because $\mathcal{C}(\cdot) \to 0$ as $\delta \to 0$, implying that the players cannot explore the asymmetric benefits domain in period 1. □

Let us now consider the case where α, the probability that the benefits go to player 1, differs across the domains. Denote by α_j, the corresponding probability for domain j. For simplicity, let us assume $\alpha_1 = 1/2$ and $\alpha_2 = 1$.

Proposition A4 *Let δ^* and $\bar{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \bar{\delta}$. In other words, for intermediate discount factors, the scope of experimentation is initially limited on path, with scope increasing with strictly positive probability along the equilibrium path.*

Proof of Proposition A4. Observe that $\delta^* = 0$ because an experimentation policy consisting of (i) exploration in domain 1 and (ii) expanding scope to domain 2 once a project exceeding $2\tilde{v}$ is discovered in domain 1 is always feasible. Further, $\bar{\delta} > 0$ because $\mathcal{C}(\cdot) \to 0$ as $\delta \to 0$, implying that the players cannot explore projects in domain 2 in period 1. □

9.2.3 Technological Interdependencies

In this extension, we assume that domains are perfectly technologically dependent. Formally, assume there is a bijection between the projects in the two domains, such that for any two projects paired by the bijection, the project values are equal. Such an assumption implies that any project value found in a given domain can be used to produce an identical value on the second domain. Further, for simplicity, we will again restrict attention to $\alpha = 1$ and $m = 2$, but the forces naturally extend.

Proposition A5 *Let δ^* and $\bar{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \bar{\delta}$. Further, scope is terminally maximal with probability 1 for any non-empty experimentation.*

Proof of Proposition A5. The latter statement follows directly from technological interdependencies and Bernheim and Whinston (1990): if the players permanently exploit k projects with value $\hat{v}$, then the players can exploit m projects with value $\hat{v}$.

In this analysis, Proposition 1 continues to hold and the optimal experimentation policy only conditions on the best project found across all domains: $\hat{v}$. Denote by $B(\hat{v})$ the bellman equation corresponding to the players' strategy. Therefore, $\bar{\delta}$ is defined as

$$2c = \bar{\delta}\mathbb{E}_{\hat{v}_1,\hat{v}_2}(B(\max\{\hat{v}_1,\hat{v}_2\})|\bar{\delta}). \tag{2.50}$$

To prove, $\delta^* < \bar{\delta}$, it suffices to show:

$$c < \bar{\delta}\lim_{\delta\uparrow\bar{\delta}}\mathbb{E}_{\hat{v}_1}(B(\hat{v}_1))|\delta) \iff 2\lim_{\delta\uparrow\bar{\delta}}\mathbb{E}_{\hat{v}_1}(B(\hat{v})|\delta) > \mathbb{E}_{\hat{v}_1,\hat{v}_2}(B(\max\{\hat{v}_1,\hat{v}_2\})|\bar{\delta}). \tag{2.51}$$

$B(\cdot|\bar{\delta})$ is characterized by two thresholds, $\underline{\hat{v}}, \bar{\hat{v}}$: for $\hat{v} \geq \bar{\hat{v}}$ the players exploit two projects with value $\hat{v}$, for $\hat{v} \in [\underline{\hat{v}}, \bar{\hat{v}})$ the players exploit one project with value $\hat{v}$ and explore a new project, and for $\hat{v} < \underline{\hat{v}}$ the players explore two new projects.

Consider an alternative experimentation policy with $\epsilon > 0$: for $\hat{v} \geq \bar{\hat{v}} + \epsilon$ the players exploit two projects with value $\hat{v}$, for $\hat{v} \in [\underline{\hat{v}} + \epsilon, \bar{\hat{v}} + \epsilon)$ the players exploit one project with value v and explore a new project, and for $\hat{v} < \underline{\hat{v}} + \epsilon$ the players explore one new project. As $\epsilon > 0$, this experimentation policy is implementable for $\hat{v} > \underline{\hat{v}} + \epsilon$ as $\delta \uparrow \bar{\delta}$. Hence, it suffices to show that this policy is implementable at period 1.

Under this policy, $\lim_{\delta\uparrow\bar{\delta}} B(0|\delta) > \frac{1}{2}B(0|\bar{\delta})$. Therefore, the expected continuation value of this policy at date 1 when $\delta \uparrow \bar{\delta}$ is strictly greater than half the expected continuation value of the optimal policy when $\delta = \bar{\delta}$. This assertion is exactly the claim we needed to show Equation (2.51) holds.

□

Chapter 3

Who Versus When: Designing Decision Processes in Organizations

Roi Orzach[1]

1 Introduction

The organizational economics literature often analyzes who should make decisions within and between organizations (Grossman and Hart, 1986b; Aghion and Tirole, 1997; Baker et al., 2002b). Less focus has been given to the order in which decisions should be made. For example, consider a firm that sells complementary products, each of which is produced and marketed by a different unit. Because of private information, it may be optimal for each of these units to set its own prices. When should the units set prices concurrently and when should they set prices sequentially? Many theoretical models cannot answer this question because they take the timing of decision-making as fixed and, often, simultaneous.

This question of whether firms should use sequential or concurrent decision-making extends beyond the pricing example given above and is of extensive study in the management literature, such as the comparison between "concurrent engineering" versus "sequential engineering." This literature makes two broad predictions. First, sequential engineering is optimal when the organization is involved in breakthrough innovation (cf. Valle and Vázquez-Bustelo, 2009). Second, if a unit cannot make a decision before another (e.g., manufacturing cannot start before design, even though they could proceed simultaneously) and its dependence on the preceding unit is sufficiently high, then sequential decision-making is optimal (cf. Terwiesch et al., 2002). Section 4 analyzes how the predictions of Valle and Vázquez-Bustelo (2009) and Terwiesch et al. (2002) can be interpreted within the model described below.

In this paper, I enrich the coordinated-adaptation model described in Alonso et al. (2008) and Rantakari (2008), endogenizing the timing of decisions. There are two self-interested units and a surplus-maximizing headquarters. A unit's profit depends on how well aligned

[1]Department of Economics, MIT, 50 Memorial Drive, Cambridge MA 02142 (e-mail: orzach@mit.edu). Department of Economics, Massachusetts Institute of Technology; orzach@mit.edu, 50 Memorial Dr, Cambridge, MA 02139
I am very thankful for helpful conversations with Charles Angelucci, Glenn Ellison, Robert Gibbons, Heikki Rantakari, Kramer Quist, the editor (Julian Wright), three anonymous referees, and seminar participants at the MIT Organizational Economics Lunch. All remaining errors are my own.

its decision is to their privately-known local condition (adaptation) and to the other unit's decision (coordination). Before making decisions, the units engage in strategic communication about their local conditions. Each unit may differentially weight coordination and adaptation. Alonso et al. (2008) and Rantakari (2008) assume concurrent decision-making and analyze when decentralization (units make both decisions), centralization (headquarters makes both decisions), or partial centralization (headquarters makes one decision and a unit makes the other) is optimal. These papers note that decentralization allows each unit's decision to adapt to its true state of the world, but each unit makes its decision without internalizing the effect on the other unit. In contrast, centralization requires the headquarters to solicit information from the units, which results in coarse communication, but the headquarters then makes decisions reflecting the socially efficient amount of coordination given this coarse information.

I address three unanswered questions. The first two questions assume decision rights are inalienable to the units, i.e., each unit must make its own decision. First, when should decentralized decisions be made sequentially rather than concurrently? Second, if sequential decentralization dominates concurrent decentralization, which unit should make its decision first? The third question allows decision rights to be alienable: How does decentralization compare to centralization and partial centralization when sequential decisions are permitted and decision rights can be re-allocated across units and headquarters?

To answer these questions, let us first consider the trade-offs under decentralization. By staggering the decisions, the unit that moves second can perfectly observe the decision of the first mover. This is in contrast to concurrent decision-making, where units convey information to each other by communicating only. 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 unit that moves second observes the decision made by the unit that moves first, but the first mover now has an incentive to over-adapt to their local state. Analyzing when the 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 concluded decentralization was never optimal when the need for coordination was relatively high. Instead, my model finds that decentralization with sequential decision-making can be optimal in such instances. This finding implies that centralization may be over-prescribed as a solution to coordination-adaptation problems when ignoring sequential decision-making.

An extensive literature exists on sequential versus concurrent contributions to public goods and analyzes which timing 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) shows 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.[2] In fact, Varian (1994) shows that in a continuous public goods setting, weakly smaller contributions will occur when they happen sequentially. However,

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

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 has been realized for their partner. This uncertainty implies that the intuition from these public-goods papers cannot be immediately extended into the setting of an organization. With private information, information quality increases under a sequential structure because the follower can perfectly observe the decision of the leader; this observability plays a new role in the organizational models, not present in the public-goods models.

As mentioned before, 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 the allocation of decision rights.[3] None of these papers gives predictions as to when decisions should be made concurrently or sequentially. In an extension, Alonso et al. (2008) compares sequential decentralization to centralization. However, since timing is not the focus of their paper, they do not compare concurrent decentralization to sequential decentralization. Two open questions remain. First, how does concurrent decentralization compare to sequential decentralization? Second, how does the optimal form of decentralization compare to centralization or partial centralization? To answer these questions, I first analyze which timing is optimal, conditional on the firm 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) considers symmetric units, the issue of which unit should go first when decisions are sequential does not arise. Finally, Rantakari (2013) shows that decentralized firms do better in volatile environments, but Rantakari (2013) does not give insights into whether the decentralization should be sequential or concurrent, which is the focus my analysis. This paper shows that multi-unit organizations with asymmetric volatilities prefer sequential decision-making.

The paper that 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 whether 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 whereby concurrent decision-making is always preferable. However, Lewis and Mistree (1997) get closest to the idea of decision-making timings within an organization.

Further, this paper connects to a literature in Industrial Organization on pricing behavior within firms. For instance, Hortaçsu et al. (2024) analyzes the interactions between various sub-units in the pricing decision of an airline and documents 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 logic through a competitive search market, this process can also be explained by 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.

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

However, the coordination-adaptation intuition offers an alternative explanation: for example, the budget unit may have a large gain from moving after the flagship unit as coordinating on as many features as possible drives down R&D costs.

The paper is organized as follows. Section 2 describes the model. 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 under public information: namely, the firm would always have units decide concurrently if there were 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 analyzing these benchmarks, I calculate the total payoffs of the organization for both sequential and concurrent decision-making when there exists private information but the units are able to communicate.

Section 4 compares the losses across these various timings and 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 decide second. In this section, I return to the predictions from Valle and Vázquez-Bustelo (2009) and Terwiesch et al. (2002) and show how to interpret them through the model's implications.

Section 5 shows that the region of parameter values in which partial centralization or centralization are optimal remains qualitatively unchanged by introducing sequential decentralization. However, in the analysis where sequential decision-making is not permitted if decentralization is optimal, the optimal form of decentralization is either directional authority (one unit making both decisions) or concurrent decision-making. By introducing sequential decentralization, both of these forms of decentralization become sub-optimal. Finally, Section 6 concludes, 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 $\theta_i \sim U\left[-\bar{\theta}_i, \bar{\theta}_i\right]$.[4] I define the volatility of unit i to be the variance of its local condition, σ_i^2. Unit i wishes to match 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:

$$\mathcal{L}_i(d_i, d_{-i}, \theta_i) = (1 - r_i)(\theta_i - d_i)^2 + r_i(d_{-i} - d_i)^2. \tag{3.1}$$

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

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

adaptation.[5] 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).[6] 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 5, I describe how the analysis changes if headquarters is allowed to make the decisions.[7]

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 concurrently.

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 3.1.

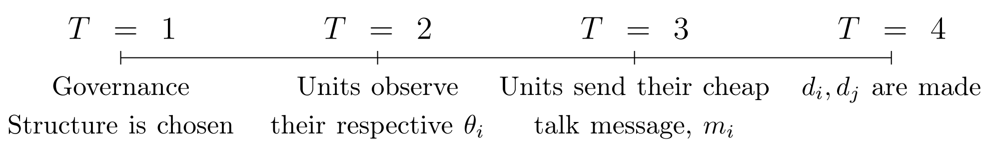


Figure 3.1: Timing of the Game

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 that the units' local conditions are public information, and then assuming that 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

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

[6]Alonso et al. (2008) and Rantakari (2008) allow for the headquarters to weight the units asymmetrically. Figure 3.3, by analyzing a headquarters that cares only about one unit, sheds light on how asymmetric weights alter the analysis.

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

computations can be found in Sections 7.4 and 7.5 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 excessively adapt to the first mover's 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 decisions 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. As a result, since the follower is the unit that coordinates more, 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 backward 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 also without loss.[8] 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 *The equilibrium decisions under sequential decision-making are given by*

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

[8] It is equivalent to assume that $m_1 = \emptyset$, 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.

whereas under concurrent decision-making they are given by

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

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 $\mathbb{E}(\theta_2|m_2)$ 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 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.[9] Given this limit partition, I define information loss about state i to be $c_i = \text{Var}(\theta_i - \mathbb{E}(\theta_i|m_i))$, where Var 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.[10] 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 I utilize the solutions provided in Rantakari (2008). The next lemma calculates the information loss for both governance structures.

Lemma 2 *Equilibrium information loss, c_i, about state i under a given decision timing is equal to $\sigma_i^2 \frac{1}{3\phi_{i,timing}+4}$ 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 $\phi_{1,\ sequential} = \infty$.*

[9]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 7.1.

[10]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 one can define the information loss to be zero, since the (observed) decision of the leader perfectly reveals the state.

The term σ_i^2 represents the variance of the state being communicated, and the next term is the percentage of the total information lost due to imperfect communication.[11] 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 the expected total loss for the leader and follower under sequential decision-making.[12]

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

$$\mathcal{L}_{Leader} = (1 - r_1)\gamma_1\Big(\sigma_1^2 + \sigma_2^2\Big) + \gamma_1\beta_1 c_2. \tag{3.2}$$

Similarly, the equilibrium expected total loss for the follower is

$$\mathcal{L}_{Follower} = (1 - r_2)r_2(1 - \gamma_1)^2\Big(\sigma_1^2 + \sigma_2^2\Big) - (1 - r_2)r_2(\gamma_1^2 - 2\gamma_1)c_2. \tag{3.3}$$

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 the information loss due to strategic communication.[13] Note that whereas under sequential decision-making there is only an information loss about the follower's state, with concurrent decision-making there is an 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:*

$$\begin{aligned}\mathcal{L}_{concurrent} &= (\sigma_1^2 + \sigma_2^2)\frac{(1 - r_1)\beta_1}{(1 - r_1 r_2)^2} + \sigma_1^2 r_1(1 - r_1)\frac{(1 - r_1 r_2)^2 - (1 - r_1)^2}{(1 - r_1 r_2)^2}V\left(\phi_1^{sim}\right)\\ &+ \sigma_2^2 r_1(1 - r_2)^2\frac{(1 - r_1 r_2)^2 - (1 - r_1)}{(1 - r_1 r_2)^2}V\left(\phi_2^{sim}\right)\\ \textit{where } V(\phi) &= \frac{1}{4 + 3\phi}, \phi_i^{sim} = \frac{\beta_i}{1 - r_i}, \gamma_i = \frac{\beta_i}{1 - r_i + \beta_i}, \textit{ and } \beta_i = r_i(1 - r_j)^2.\end{aligned}$$

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

[11] Given the uniform assumption on the local states, $\sigma_i^2 = \frac{\bar{\theta_i}^2}{3}$.

[12] Recall that the headquarters' loss is the sum of the losses for the two units.

[13] 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.

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

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

2. *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 the one who would not have coordinated much anyway.[14]

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

1. *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.*

2. *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* $r_1 > r_2$.

3. *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* $\sigma_1^2 > \sigma_2^2$.

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.[15] 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. This over-adaptation implies the leader is less responsive to the message of the follower 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. To see this, note that a unit who solely cares about adaptation is indifferent about which action the other unit takes. Given this indifference, the unit can perfectly communicate their local condition.

[14] 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.

[15] If the payoffs to a unit included a $(d_i - \theta_{-i})^2$ term, then perfect transmission about the leader's state yields an additional benefit. To be consistent with the literature, I do not include such a term.

This causes the 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 the aggregate information loss.[16] The next section reports on whether the effects stated 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 when the decision rights are assumed to be inalienable (Section 5 relaxes this assumption). In the next subsection, I analyze the empirical evidence on optimal decision-making timing in light of this model.

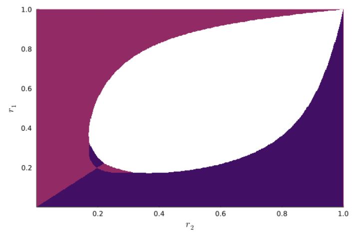
4.1 Relative Performance

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.[17] In Figure 3.2, I compare both versions of sequential decision-making with concurrent decision-making. Figure 3.2a shows the optimal governance structure when both units have equal volatilities.[18] Figure 3.2b 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 to unit 2 leading, and White to concurrent decision-making.

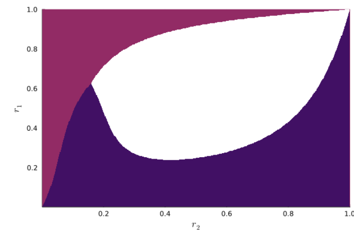
[16] This result was shown in Alonso et al. (2008) for the case of symmetric units.

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

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



((a)) Equal volatility across units.



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

Figure 3.2: Decentralization and Centralization

The color denotes the optimal timing with Purple corresponding to unit 1 leading, Pink to unit 2 leading, and White to concurrent decision-making. The optimal timing refers to 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.

One should note that, when r_1 and r_2 are of a similar magnitude and high, the optimal 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. The informational effect implies an improved information flow from the leader to a follower. Given that the follower puts higher weight on coordination, the improved information allows the follower to better coordinate with the leader, thus decreasing the losses for both parties.

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 such weakening of communication 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.

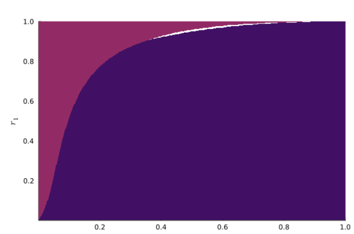
For most of Figure 3.2, the dominant force is that described in Proposition 1: conditional on information, letting the unit which places a higher priority on coordination move second increases coordination. However, in the region with symmetric and intermediate values of r_i and r_j, the dominant force is that described in Proposition 2: having the unit which places a higher priority on coordination move first decreases communication. By Proposition 2, this force is strongest with equally weighted units. Further, when r_i and r_j are either both small/both large, communication is poor/informative irrespective of timing, respectively. Hence, the effect in Proposition 2 is strongest for intermediate and symmetric coordination weights. This region is discussed further in Section 7.2 of the Appendix.

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 result occurs 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 stated in Proposition 2 of increased information transmission becomes stronger than the effect described in Proposition 1 of decreased coordination.

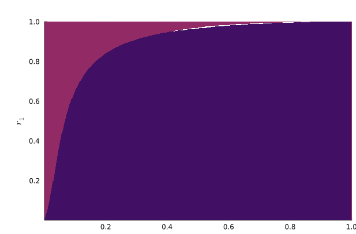
While Figure 3.2 analyzed how to minimize joint losses, Figure 3.3 analyzes the timing that minimizes the loss for a single unit. Without loss of generality, let this be unit 1. Thus far the model has assumed that firm surplus is the sum of the two units' total losses, however some units may receive greater weights than others. For example, Apple may weight their phone unit more than the headphones unit. As the weight the firm places on unit 1 increases, Figure 3.2 will continuously form into Figure 3.3.

When analyzing the total loss of a specific unit, a unit that cares sufficiently about coordination would in fact prefer to move second to ensure that they are able to coordinate with the decision of the leader. For instance, a unit that only values coordination can achieve zero losses by simply observing (and matching) the decision of the first-mover. Meanwhile, moving first requires unit i to predict d_j based on the (imperfect) strategic communication about θ_i. 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. Moreover, a unit only prefers concurrent decision-making for a small region of parameter values, as seen by the small White regions in Figure 3.3a and 3.3b. This preference occurs for intermediate values of coordination, because concurrent decision-making balances the two forces of more adaptation as a leader but better coordination as a follower.

Additionally, as seen in subfigure b, 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, as a leader, can choose an extreme decision (a high d_1) in response to an extreme state (a high θ_1) with the knowledge that it is in unit 2's best interest to coordinate on such a state. 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.[19]



((a)) Equal volatility across units.



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

Figure 3.3: Decentralization and Centralization for Firm 1

The color denotes the optimal timing with Purple corresponding to unit 1 leading, Pink to unit 2 leading, and White to concurrent decision-making. The optimal timing refers to 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 White region is small, but can be seen around the point $(r_1, r_2) = (.9, .5)$.

[19]Formally, 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.

However, as one can see when 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 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 discussed in the introduction and show how to interpret it through this model.

Terwiesch et al. (2002) analyzes the coordination between multiple units engaged in the production of a vehicle's climate control system. For decisions at this level of specificity, the ability for decisions to be made by headquarters (i.e., centralization) is not discussed in the text, suggesting either that it is suboptimal, or, more likely, unfeasible. We can therefore safely disregard the additional governance structures that will be discussed in Section 5. Further, many of their examples have a natural "follower"-unit who can only act either concurrently or after the "leader"-unit. Thus, their focus is when should the organization be concurrent or sequential (where only one of the two sequential structures is feasible). Figure 3.5 in their analysis notes that when the level of dependence from the follower to the leader's information is high (respectively, low), the optimal structure is sequential (respectively, concurrent). In my model, a high dependence corresponds to r_2 being large, and one can see that fixing r_1, there exists a threshold as a function of r_1, such that the optimal structure is sequential if and only if r_2 exceeds that threshold. The authors note that such intuitions extend beyond a vehicle's climate control system, and to multi-unit decision-making more generally.

The authors have additional predictions based on the level of ambiguity in the environment. Ambiguity is not formally defined in my analysis, but one could argue that a consequence of ambiguity as described in Terwiesch et al. (2002) is an inability for units to communicate before making their decisions (i.e., ambiguity between the units is resolved only after the decision is made). Terwiesch et al. (2002) notes that, "Concurrency is appropriate when the design is far enough advanced to largely eliminate ambiguity." Reassuringly, these predictions are born out in a benchmark version of my analysis as discussed in Remark 1, when the units cannot engage in cheap talk communication. In this benchmark, I find that concurrency is never optimal as no information is exchanged, however sequential allows for the follower to coordinate on the decision of the leader.

Further, Valle and Vázquez-Bustelo (2009) uses survey evidence to show that firms involved in breakthrough innovation are more likely to utilize sequential decision-making.[20] In terms of the model, one unit is the sales unit and the other is the research unit, where the

[20]Takeuchi and Nonaka (1986) argues that "the traditional sequential or 'relay race' approach to product development ... may conflict with goals of maximum speed and flexibility" and detail organizations which have benefited from a switch to concurrent decision-making (Takeuchi and Nonaka:137). However, Takeuchi and Nonaka (1986) also writes that "[concurrent decision-making] may not apply to breakthrough projects that require a revolutionary innovation. This limitation may be particularly true in biotechnology or chemistry," where sequential decision-making would be optimal (Takeuchi and Nonaka:145).

research unit is strongly motivated by the discovery of valuable working chemicals rather than catering to the whims of the sales department. Further, while the sales unit may prefer to sell in less competitive markets or ones where 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).[21] In the model, these situations correspond to $r_{\text{research}} \simeq 0$ and $r_{\text{sales}} \simeq 1$. As Figure 3.2 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 breakthrough R&D. As suggested by Figure 3.2, it will often be optimal for the research unit to move first when $\sigma^2_{\text{research}} > \sigma^2_{\text{sales}}$. Note that the finding that the research unit moves first is not driven by technological assumptions; for instance, early-stage startups may secure sales before completing their engineering. All in all, the predictions of this theoretical analysis are borne out in the qualitative analysis of Valle and Vázquez-Bustelo (2009).

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.[22] 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. Throughout this example, I presume that decision rights cannot move across the units (as will be done in Section 5). In this example the only way to move decision rights across the units would involve integrating Chrysler with their suppliers. As discussed in Dyer (1996), this was not feasible, "Japanese manufacturers like Toyota and Nissan typically own 20% to 50% of the equity of their largest suppliers; Chrysler does not and could not take similar stakes... Chrysler has a much larger group of suppliers, and few of its most important suppliers depend on it for a majority of their sales" (Dyer:11).

Dyer (1996) describes how, prior to the 1990s, Chrysler frequently changed suppliers, leaving both types of parties hesitant to make relationship-specific investments.[23] The lack of these investments, in turn, caused neither party's profits to be too dependent on their partner's decision. In other words, 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 3.2, 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

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

[22] 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.

[23] 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.

"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 becomes optimal.

The new governance structure between Chrysler and their suppliers improved performance, which Dyer (1996) attributed, in part, to their improved communication. Whereas, 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. (2024) shows that information from units who choose their pricing first is not fully adapted to by units who choose their 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 that the firm produces both a budget car, with low margins, and a luxury car, with high margins. By allowing the luxury unit to choose prices after the budget unit, the firm prevents the budget unit from cannibalizing the high margins of the luxury unit. Such a dynamic can be framed as the luxury unit having a higher coordination weight and thus, as in Figure 3.2, moving second.

5 Alternative Governance Structures

Thus far in the analysis, the only role of headquarters has been to decide the optimal timing of decisions by the units. This approach was driven by the fact that, in the motivating examples of Dyer (1996), Terwiesch et al. (2002) and Valle and Vázquez-Bustelo (2009), decision rights were effectively inalienable.[24] However, it is a natural question to ask whether for those organizations where decision rights can be moved more flexibly, if the instances in

[24]Moreover, even when decision rights are transferable, governance structures such as centralization may impose other costs. For instance, Aguilar and Bhambri (1983) notes 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.

which I find sequential decentralization to outperform concurrent decentralization in fact correspond to parameter values such that the decision rights should be alienated. When decision rights are alienable across the organization, many governance structures are possible. I focus my analysis to those considered in Alonso et al. (2008) and Rantakari (2008): centralization, directional authority, and partial centralization.

I will show in Figure 3.4 that when concurrent decision-making under decentralization (hereafter, concurrent decentralization) improves upon sequential decision-making under decentralization (hereafter, sequential decentralization), it is dominated by either centralization or partial centralization. In contrast, sequential decentralization will be optimal within the broader set of governance choices. These two statements imply that sequential decision-making is the structure that should be used to make the centralization vs decentralization comparison. Finally, I show that allowing for sequential decentralization reduces the need for asymmetric decision structures such as directional authority (where both decisions are made by the same unit) and has little effect on the need for partial centralization (where only one decision is centralized while the other unit retains control).

All the results for centralization and partial centralization are taken from Rantakari (2008) which analyzes an identical setting. Under centralization, both units simultaneously send messages m_1 and m_2. Conditional on the messages, headquarters chooses d_1 and d_2. The intuition for the solution is that conditional on beliefs about θ_1 and θ_2 (given the messages m_1 and m_2), headquarters minimizes their loss and sets d_1 and d_2 as follows

$$d_i = \frac{(r_i^2 - 1)\,\mathbb{E}(\theta_i|m_i) + (r_j - 1)\mathbb{E}(\theta_j|m_j)(r_i + r_j)}{r_i^2 + r_i(r_j - 1) + r_j^2 - r_j - 1}. \tag{3.4}$$

Given this decision rule, Rantakari (2008) solves for the communication equilibrium between the two units and the headquarters. Finally, given the decision rule and the communication equilibrium, one can calculate total losses. The fundamental trade-off between centralization and concurrent decentralization is that centralization allows the headquarters to internalize the coordination externalities. However, under centralization d_i is a function of $\mathbb{E}(\theta_i|m_i)$, as opposed to under decentralization where d_i is a function of θ_i.

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.[25] 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, suppose both decision rights are given to unit 1 (which Rantakari (2008) refers to as directional authority). This

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

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. Meanwhile, under centralization, since communication is imperfect, the headquarters must choose a decision for unit 1 based on imprecise communication. Since unit 1 has a high weight on adaptation, this imprecision leads to improper adaptation. Rantakari (2008) shows that giving both decision rights to unit 1 improves upon centralization in these instances.

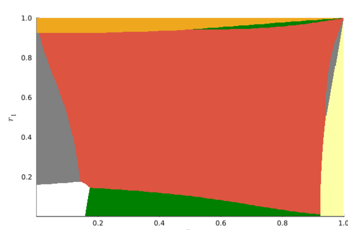
Finally, Rantakari (2008) analyzes partial centralization. This structure involves headquarters making d_i and d_j being made by unit j. Such a choice is a middle ground between centralization and decentralization. The calculations for partial centralization are in Rantakari (2008), but are summarized below for the case in which $i = 2$. Rantakari (2008) solves the game by backward induction. Conditional on messages m_1, m_2 and the posterior beliefs about θ_1, θ_2 one can solve

$$d_1 = \theta_1 \frac{(-1 + r_1)(r_1 + 2r_2)}{r_1^2 + r_1(-1 + r_2) - 2r_2} + \mathbb{E}(\theta_2|m_2) \frac{-r_1 r_2}{r_1^2 + r_1(-1 + r_2) - 2r_2} \tag{3.5}$$

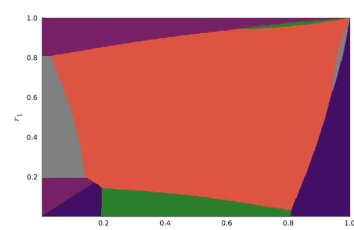
$$d_2 = \mathbb{E}(\theta_1|m_1) \frac{(-1 + r_1)(r_1 + r_2)}{r_1^2 + r_1(-1 + r_2) - 2r_2} + \mathbb{E}(\theta_2|m_2) \frac{-r_2}{r_1^2 + r_1(-1 + r_2) - 2r_2}. \tag{3.6}$$

Given this decision rule, Rantakari (2008) solves for the communication equilibrium between the two units and the headquarters. Finally, given the decision rule and the communication equilibrium, one can calculate total losses. Partial centralization offers a middle ground between centralizing both decisions and decentralizing both decisions.

Below I compare the optimality of these five governance choices to the three considered in Section 4.1. In subfigure (a) the optimal governance choice is shown when not allowing for sequential decentralization. Subfigure (b) conducts an identical analysis but allows for sequential decentralization.[26]



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



((b)) Allowing sequential decision-making.

Figure 3.4: Optimal Governance with Sequential Decision-Making

In both figures Dark Orange Corresponds to centralization, Purple unit 1 leads under sequential decentralization, Pink unit 2 leads under sequential decentralization, White corresponds to concurrent decentralization, Green is partial centralization with unit 2 retaining control, Gray is partial centralization with unit 1 retaining control, Light Orange corresponds to directional authority with unit 2 retaining control, and Yellow corresponds to directional authority with unit 1 retaining control. 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) does not allow for sequential decision-making (b) allows for sequential decision-making.

[26]Subfigure (a) corresponds to the analysis done in Rantakari (2008) and I thank Heikki Rantakari for providing the Mathematica files corresponding to his analysis.

One can see from Figure 3.4 that the regions of parameter values for which centralization and partial centralization are optimal remain qualitatively similar. However, the decentralized governance structures in Figure 3.4a, directional authority and concurrent decentralization, are no longer optimal; instead, sequential decentralization is optimal.[27] The intuition that concurrent decentralization is no longer optimal is that it was optimal in Figure 3.4a when both units' weights on coordination were low. However, in such instances, communication quality is low under concurrent decentralization, but perfect communication always occurs from the leader to the follower under sequential decentralization. As seen in Figure 3.2a, low weights on coordination is precisely when sequential decentralization is preferred.

Further, to understand why directional authority is not optimal, let us return 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. When unit 1 moves first, unit 1 can steer unit 2's decision by knowing unit 2 will coordinate with unit 1. This steering allows unit 1 to adapt to their own state. However, unlike directional authority, where unit 1 controls unit 2's decision, sequential decision-making allows unit 2 to adapt their decision to their own local conditions.

This new insight regarding directional authority suggests a possible re-interpretation of Rantakari (2008)'s original example for directional authority: Microsoft integrating the ITPD (their browser unit at the time) into the operating system unit. From the perspective of Rantakari (2008), such a decision was optimal as the browser unit's weight on coordination was high and the operating system unit's weight on coordination was low. However, as it turned out, such a decision came at a large loss to the browser unit, and Rantakari (2008) notes "Microsoft's lack of focus on the Internet has left vacant various market opportunities that might provide a future threat to Microsoft." In contrast, my model predicts sequential decentralization with the operating system unit moving before the browser unit would still give the operating system unit autonomy, while allowing the browser unit to continue adapting to their conditions. In other words, my analysis suggests that sequential decentralization may have improved upon directional authority in this instance.

Finally, in Appendix Section 7.3, I consider *sequential* centralization in which the timing is: unit i sends message m_i, upon observing m_i the headquarters chooses d_i.[28] Further, after observing d_i, unit j sends m_j, and finally the headquarters chooses d_j. Such a choice never improves upon concurrent centralization due to two inefficiencies. First, d_i is made irrespective of θ_j. Second, I show that communication about θ_j is worsened, implying there is an inefficiency on d_j.

[27] 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, one reason why concurrent decentralization may be seen is that centralization is not feasible.

[28] The calculations for sequential centralization are available upon request from the author.

6 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 the follower cares only about the leader's decision, not the leader'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 concurrent decision-making is better when the preferences for coordination between two units are both sufficiently high and similar. Finally, even when allowing for a broader class of governance choices, such as centralization, sequential decision-making under decentralization continues to be optimal when the weights on coordination are low or sufficiently asymmetric.

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, since when a unit values coordination more they have a reduced incentive to preempt the decision of their partner. 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 θ_2, unit 2 chooses

$$d_2 = (1 - r_2)\theta_2 + r_2 d_1. \tag{3.7}$$

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

$$U_2(\theta_2, d_1) = (1 - r_2)r_2(d_1 - \theta_2)^2. \tag{3.8}$$

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

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

Substituting in equation 3.7 and the definition of β yields

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

The minimizer of the above is also the minimizer to

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

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 (3.8), this generates the following indifference condition:

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

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 $\text{Var}(\theta_i - \mathbb{E}(\theta_i|m_i)) = \frac{1}{3}\frac{1}{3\phi+4}$. Note that, in my case, $\phi = \frac{\beta_1}{1-r_1}$.

□

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

$$\begin{aligned}
&\mathbb{E}\Big((1-r_1)(d_1-\theta_1)^2 + r_1((1-r_2)\theta_2 + r_2 d_1 - d_1)^2|i\Big) \\
&= (1-r_1)\mathbb{E}((d_1-\theta_1)^2|i) + r_1(1-r_2)^2\big(\mathbb{E}((\theta_2-d_1)^2|i)\big) \\
&\text{recall } d_1 = (1-\gamma_1)\theta_1 + \gamma_1\bar{m}_i \\
&= (1-r_1)\mathbb{E}((-\gamma_1\theta_1+\gamma_1\bar{m}_i)^2|i) + r_1(1-r_2)^2\Big(\mathbb{E}((\theta_2-(1-\gamma_1)\theta_1-\gamma_1\bar{m}_i)^2|i)\Big) \\
&= (1-r_1)\gamma_1^2\mathbb{E}((\theta_1-\bar{m}_i)^2|i) + r_1(1-r_2)^2\Big((1-\gamma_1)^2\sigma_1^2 + \mathbb{E}(\theta_2^2|i) + \gamma_1^2(\bar{m}_i)^2 - 2\gamma_1(\bar{m}_i)^2\Big) \quad (3.9) \\
&= (1-r_1)\gamma_1^2(\mathbb{E}(\theta_1^2|i) + \bar{m}_i{}^2) + \beta_1\Big((1-\gamma_1)^2\sigma_1^2 + \mathbb{E}(\theta_2^2|i) + (\gamma_1^2-2\gamma)(\bar{m}_i)^2\Big).
\end{aligned}$$

Taking the expectation across intervals yields

$$\begin{aligned}
&\sum_{i=1}^{N} P(i)\Bigg((1-r_1)\gamma_1^2(\mathbb{E}(\theta_1^2|i) + \bar{m}_i{}^2) + \beta_1\Big((1-\gamma_1)^2\sigma_1^2 + \mathbb{E}(\theta_2^2|i) + (\gamma_1^2-2\gamma)(\bar{m}_i)^2\Big)\Bigg) \\
&= (1-r_1)\gamma_1^2(\sigma_1^2 + \mathbb{E}(\bar{m_i^2})) + \beta_1\Big((1-\gamma_1)^2\sigma_1^2 + \sigma_2^2 + (\gamma_1^2-2\gamma)\mathbb{E}(\bar{m}_i{}^2)\Big)\Big).
\end{aligned}$$

One can re-arrange the above in terms of loss from miscoordination and loss from imperfect communication as follows where $c_i := \sigma_i^2 - \mathbb{E}(\bar{m_i}^2)$ is the loss from communication:

$$\begin{aligned}
&= (1-r_1)\gamma_1^2(\sigma_1^2+\sigma_2^2-c_2) + \beta_1\Big((1-\gamma_1)^2\sigma_1^2+\sigma_2^2+(\gamma_1^2-2\gamma)(\sigma_2^2-c_2)\Big)\\
&= \Big(\sigma_2^2+\sigma_1^2\Big)\Big((1-r_1)\gamma_1^2+\beta_1(1-\gamma_1)^2\Big) - c_2\Big((1-r_1)\gamma_1^2+\beta_1(\gamma_1^2-2\gamma)\Big)\\
&= (1-r_1)\gamma_1\Big(\sigma_2^2+\sigma_1^2\Big)+\gamma_1\beta_1 c_2.
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}((d_1(k(\theta))-\theta)^2|i) &= \mathbb{E}(\big((1-\gamma_1)\theta_1+\gamma_1\mu_i-\theta_2\big)^2\big|\, i)\\
&= (1-\gamma_1)^2\sigma_1^2+\mathbb{E}(\big(\theta_2-\gamma_1\mu_i\big)^2\big|\, i)\\
&= (1-\gamma_1)^2\sigma_1^2+\mathbb{E}(\theta_2^2|i)+\gamma_1^2(\bar{m}_i)^2-2\gamma_1(\bar{m}_i)^2.
\end{aligned}$$

Taking the expectation across intervals yields

$$\begin{aligned}
&(1-r_2)r_2\sum_{i=1}^{N}P(i)\Bigg(((1-\gamma_1)^2\sigma_1^2+\mathbb{E}(\theta_2^2|i)+\gamma_1^2(\bar{m}_i)^2-2\gamma_1(\bar{m}_i)^2\Bigg)\\
&= (1-r_2)r_2\Big((1-\gamma_1)^2\sigma_1^2+\sigma_2^2+(\gamma_1^2-2\gamma_1)\mathbb{E}(\bar{m}_i^2)\Big).
\end{aligned}$$

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 $c_i = \sigma_i^2 - \mathbb{E}(\bar{m_i}^2)$, so

$$= (1-r_2)r_2(1-\gamma_1)^2\Big(\sigma_2^2+\sigma_1^2\Big) - (1-r_2)r_2(\gamma_1^2-2\gamma_1)c_2.$$

□

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 2 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:

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

Hence coordination loss is lower 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 1 of Proposition 1:

Meanwhile coordination loss when the units move at the same time is bounded above by $(\frac{(1-r_1)(1-r_2)}{1-r_1r_2})^2\mathbb{E}(\theta_2^2+\theta_1^2|m_1,m_2)$. Comparing this to the sequential structure yields[29]

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

□

Proof of Proposition 2. I begin by proving statement 1.

Statement 1 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 $\sigma_i^2\frac{1}{4+3\frac{\beta_i}{1-r_i}}$. Note that this is always less than $\frac{\sigma_i^2}{4}$. It is without loss to assume $\sigma_2^2 < \sigma_1^2$ 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

$$\begin{aligned}
&\mathcal{L}_{\text{concurrent}} = \frac{\sigma_1^2}{4+3\phi_{\text{concurrent}_1}} + \frac{\sigma_2^2}{4+3\phi_{\text{concurrent}_2}} \overset{?}{>} \frac{\sigma_2^2}{4} \\
\iff &\mathcal{L}_{\text{concurrent}} = \frac{1}{4+3\phi_{\text{concurrent}_1}} + \frac{1}{4+3\phi_{\text{concurrent}_2}} \overset{?}{>} \frac{1}{4} \\
\iff &\frac{\phi_{\text{concurrent}_1}\phi_{\text{concurrent}_2} - \frac{16}{9}}{(\phi_{\text{concurrent}_1}+\frac{4}{3})(\phi_{\text{concurrent}_2}+\frac{4}{3})} \overset{?}{<} 0 \\
\iff &\phi_{\text{concurrent}_1}\phi_{\text{concurrent}_2} \overset{?}{<} \frac{16}{9} \\
\iff &r_1r_2 < \frac{16}{9}.
\end{aligned}$$

Thus the sequential information loss is always bounded below by $\frac{\sigma_2^2}{4}$; meanwhile, the concurrent information loss is bounded above by $\frac{\sigma_2^2}{4}$. The latter statement that there is worse information

[29]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.

transmission about the followers state follows by noting that $\phi_{\text{concurrent}} > \phi_{\text{sequential}}$ and that information loss is monotonic in ϕ.

Statement 2 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 3 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. □

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 3.1, 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 3.1 than it is below since this region is where concurrent decision-making is optimal.

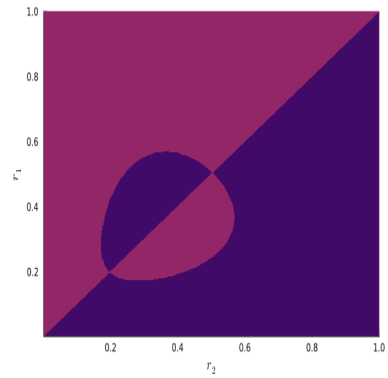


Figure 3.5: Optimal Sequential Decentralization

The color denotes the optimal timing with Purple corresponding to unit 1 leading and Pink unit 2 leading. The optimal timing refers to 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.

7.3 Additional Governance Choices

First, one might wonder why there is no scope for sequential centralization. The most natural analog of sequential decentralization within centralization is unit i sends m_i, the headquarters chooses d_i, then unit j sends m_j, and finally headquarters chooses d_j. Further, sequential communication under concurrent centralization (and other modes of communication) has already been studied in Alonso and Rantakari (2022) in this coordinated-adaptation context where concurrent messages are shown to be the optimal messaging structure. Their result contributes to a larger literature discussing optimal communication from units to headquarters started by Gilligan and Krehbiel (1989) and Krishna and Morgan (2001). Given the result in Alonso and Rantakari (2022), one can show the above decision timing is the only remaining form of centralization for which the units send at most one message.

Sequential centralization never improves upon concurrent centralization for two reasons. First, under sequential centralization d_i is made irrespective of θ_j. This inefficiency may improve the communication from unit i to the headquarters under sequential centralization, as the conflict introduced by θ_j has not been realized yet. However, there is still a net inefficiency. To gain intuition, at the extreme, a decision taken independent of m_i will induce truthful communication, but will never be efficient. Second, communication about θ_j is worsened. Previously, unit j, irrespective of r_i, r_j, could communicate whether θ_j was above or below 0. Now, headquarters always wants d_j closer to d_i than unit j would want, implying that unit j may only be able to communicate whether d_j is above or below θ_j or not. In comparison to sequential decentralization, where the leader still receives communication from the follower before making their decision, this is inconsistent with sequential decision-making under centralization. This is because, if d_1 and d_2 are made after both messages are received, then there is no difference between headquarters making both decisions at the same time or staggering the decisions.

7.4 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 units 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.[30] However, as the weight the headquarters places on unit 1 increases, sequential structures will eventually become optimal because leading is best for unit 1.

7.5 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:

- $\frac{(1-r_1)\beta_1}{(1-r_1)+\beta_1}\sigma_1^2 + \beta_1\sigma_2^2$ when unit 1 leads
- $r_1(1-r_1)\left[\sigma_1^2 + \frac{\alpha_2^2}{(\alpha_2+\beta_2)^2}\sigma_2^2\right]$ when unit 1 follows
- $r_1(1-r_1)\sigma_1^2 + \beta_1\sigma_2^2$ is the loss when unit 1 decides concurrently.

One can see the loss from miscoordination, i.e., the σ_2^2 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 in Figure 3.6 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.

[30] 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.

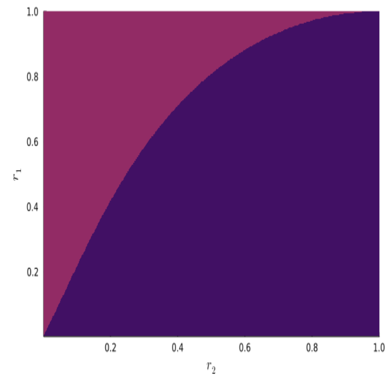


Figure 3.6: Optimal Governance for Unit 1 without Communication

The color denotes the optimal timing where Purple corresponds to unit 1 leads, Pink is unit 2 leads, and the non-existent White concurrent decision-making. The optimal timing refers to the governance structure which minimizes the loss 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, with equal volatilities across the two units.

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 3.7. 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.

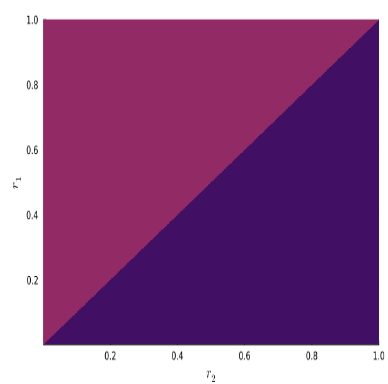


Figure 3.7: Optimal Governance without Communication

The color denotes the optimal timing where Purple corresponds to unit 1 leads, Pink is unit 2 leads, and the non-existent White concurrent decision-making. The optimal timing refers to the governance structure that minimizes the losses of the two units when the units are not allowed to communicate as a function of the coordination motives of the units, r_1 and r_2, with equal volatilities for the two units.

Bibliography

Abreu, Dilip, "Extremal equilibria of oligopolistic supergames," *Journal of Economic Theory*, 1986, *39* (1), 191–225.

Acharya, Avidit and Juan Ortner, "Paths to the Frontier," *American Economic Journal: Microeconomics*, 2022, *14* (1), 39–69.

Admati, Anat R and Motty Perry, "Joint Projects Without Commitment," *The Review of Economic Studies*, 1991, *58* (2), 259–276.

Aghion, Philippe and Jean Tirole, "Formal and real authority in organizations," *Journal of political economy*, 1997, *105* (1), 1–29.

Aguilar, Francis J and Arvind Bhambri, "Johnson & Johnson (A)," *Harvard Bussines School Case*, 1983, (5).

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

Albrecht, James, Axel Anderson, and Susan Vroman, "Search by committee," *Journal of Economic Theory*, 2010, *145* (4), 1386–1407. Search Theory and Applications.

Ali, S Nageeb, "On the role of responsiveness in rational herds," *Economics Letters*, 2018, *163*, 79–82.

Alonso, Ricardo and Heikki Rantakari, "The art of brevity," *Journal of Economic Behavior & Organization*, 2022, *195*, 257–271.

_ , Niko Matouschek, and Wouter Dessein, "Strategic communication: prices versus quantities," *Journal of the European Economic Association*, 2010, *8* (2-3), 365–376.

_ , Wouter Dessein, and Niko Matouschek, "When Does Coordination Require Centralization?," *American Economic Review*, 2008, *98* (1), 145–79.

Andrews, Isaiah and Daniel Barron, "The Allocation of Future Business: Dynamic Relational Contracts with Multiple Agents," *American Economic Review*, September 2016, *106* (9), 2742–59.

Anesi, Vincent and T. Renee Bowen, "Policy Experimentation in Committees: A Case against Veto Rights under Redistributive Constraints," *American Economic Journal: Microeconomics*, August 2021, *13* (3), 124–62.

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan, "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks," *Econometrica*, 2007, *75* (3), 711–756.

Angeli, Deivis, Matt Lowe et al., "Virtue Signals," 2022.

Arieli, Itai and Manuel Mueller-Frank, "Multidimensional social learning," *The Review of Economic Studies*, 2019, *86* (3), 913–940.

Asch, Solomon E, "Effects of group pressure upon the modification and distortion of judgements," *Group dynamics*, 1953.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *The Quarterly Journal of Economics*, 03 2017, *132* (3), 1101–1164.

Austen-Smith, David and Roland G Fryer Jr, "An economic analysis of 'acting white'," *The Quarterly Journal of Economics*, 2005, *120* (2), 551–583.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics*, 1994, *109* (4), 1125–1156.

_ , _ , and _ , "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

_ , _ , and Kevin J Murphy, "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Banerjee, Abhijit and Drew Fudenberg, "Word-of-mouth learning," *Games and economic behavior*, 2004, *46* (1), 1–22.

Banerjee, Abhijit V, "A simple model of herd behavior," *The quarterly journal of economics*, 1992, *107* (3), 797–817.

Benabou, Roland and Jean Tirole, "Laws and Norms," Working Paper, Princeton University 2024.

Benavides, Luis, Versa De Eskinazis, and Daniel Swan, "Six steps to successful supply chain collaboration," *CSCMP's Supply Chain Quarterly*, 2012, *2012* (2), 23–27.

Bergemann, Dirk and Juuso Välimäki, "Stationary multi-choice bandit problems," *Journal of Economic dynamics and Control*, 2001, *25* (10), 1585–1594.

_ and Juuso Välimäki, "Bandit Problems," in Steven Durlauf and Larry Blume, eds., *The New Palgrave Dictionary of Economics*, 2nd ed., Macmillan Press, 2008.

Bernheim, B Douglas, "A theory of conformity," *Journal of political Economy*, 1994, *102* (5), 841–877.

Bernheim, B. Douglas and Michael D. Whinston, "Multimarket contact and collusive behavior," *The RAND Journal of Economics*, 1990, pp. 1–26.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 1992, *100* (5), 992–1026.

_ , _ , Omer Tamuz, and Ivo Welch, "Information cascades and social learning," Technical Report, National Bureau of Economic Research 2021.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, " Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 11 2013, *128* (1), 1–51.

Board, Simon, "Relational Contracts and the Value of Loyalty," *American Economic Review*, December 2011, *101* (7), 3349–67.

Bohren, J Aislinn, "Informational herding with model misspecification," *Journal of Economic Theory*, 2016, *163*, 222–247.

Bolton, Patrick and Christopher Harris, "Strategic experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Braghieri, Luca, "Political correctness, social image, and information transmission," *Work. Pap., Stanford Univ., Stanford, CA*, 2021.

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *The Quarterly Journal of Economics*, 1987, *102* (1), 147–159.

Burguet, Roberto and Xavier Vives, "Social learning and costly information acquisition," *Economic theory*, 2000, *15* (1), 185–205.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott, "Misperceived social norms: Women working outside the home in Saudi Arabia," *American economic review*, 2020, *110* (10), 2997–3029.

_ and David Y Yang, "Misperceptions about others," *Annual Review of Economics*, 2022, *14*, 425–452.

_ , Georgy Egorov, and Robert Jensen, "Cool to be smart or smart to be cool? Understanding peer pressure in education," *The Review of Economic Studies*, 2019, *86* (4), 1487–1526.

_ , _ , and Stefano Fiorin, "From extreme to mainstream: The erosion of social norms," *American economic review*, 2020, *110* (11), 3522–3548.

Callander, Steven, "Searching and Learning by Trial and Error," *American Economic Review*, October 2011, *101* (6), 2277–2308.

_ **and Bård Harstad**, "Experimentation in Federal Systems *," *The Quarterly Journal of Economics*, 04 2015, *130* (2), 951–1002.

_ **and Niko Matouschek**, "The Risk of Failure: Trial and Error Learning and Long-Run Performance," *American Economic Journal: Microeconomics*, February 2019, *11* (1), 44–78.

Canes-Wrone, Brandice, Michael C Herron, and Kenneth W Shotts, "Leadership and pandering: A theory of executive policymaking," *American Journal of Political Science*, 2001, pp. 532–550.

Chan, Jimmy, Alessandro Lizzeri, Wing Suen, and Leeat Yariv, "Deliberating collective decisions," *The Review of Economic Studies*, 2018, *85* (2), 929–963.

Chandrasekhar, Arun G, Benjamin Golub, and He Yang, "Signaling, shame, and silence in social learning," Technical Report, National Bureau of Economic Research 2018.

Chang, Edward H, Katherine L Milkman, Dolly Chugh, and Modupe Akinola, "Diversity thresholds: How social norms, visibility, and scrutiny relate to group composition," *Academy of Management Journal*, 2019, *62* (1), 144–171.

Chassang, Sylvain, "Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts," *American Economic Review*, 2010, *100* (1), 448–65.

Cho, In-Koo and David M Kreps, "Signaling games and stable equilibria," *The Quarterly Journal of Economics*, 1987, *102* (2), 179–221.

Crawford, Vincent P and Joel Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

Dasaratha, Krishna, Benjamin Golub, and Nir Hak, "Learning from neighbours about a changing state," *Review of Economic Studies*, 2023, *90* (5), 2326–2369.

Dessein, Wouter and Tano Santos, "Adaptive organizations," *Journal of Political Economy*, 2006, *114* (5), 956–995.

Doney, Patricia M. and Joseph P. Cannon, "An Examination of the Nature of Trust in Buyer-Seller Relationships," *Journal of Marketing*, 1997, *61* (2), 35–51.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh, "Developing Buyer-Seller Relationships," *Journal of Marketing*, 1987, *51* (2), 11–27.

Dyer, Jeffrey H, "How Chrysler Created an American Keiretsu.," *Harvard Business Review*, 1996, *74* (4), 42–52.

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

Ely, Jeffrey C and Juuso Välimäki, "Bad reputation," *The Quarterly Journal of Economics*, 2003, *118* (3), 785–814.

Fernández-Duque, Mauricio, "The probability of pluralistic ignorance," *Journal of Economic Theory*, 2022, *202*, 105449.

Franzen, Axel and Sebastian Mader, "The power of social influence: A replication and extension of the Asch experiment," *Plos one*, 2023, *18* (11), e0294325.

Frick, Mira, Ryota Iijima, and Yuhta Ishii, "Misinterpreting others and the fragility of social learning," *Econometrica*, 2020, *88* (6), 2281–2328.

Fudenberg, Drew and Jean Tirole, "Perfect Bayesian equilibrium and sequential equilibrium," *journal of Economic Theory*, 1991, *53* (2), 236–260.

Garfagnini, Umberto and Bruno Strulovici, "Social Experimentation with Interdependent and Expanding Technologies," *The Review of Economic Studies*, 02 2016, *83* (4), 1579–1613.

Geiger, Nathaniel and Janet K Swim, "Climate of silence: Pluralistic ignorance as a barrier to climate change discussion," *Journal of Environmental Psychology*, 2016, *47*, 79–90.

Gibbons, Roberts and Rebecca Henderson, "What Do Managers Do? Exploring Persistent Performance Differences Among Seemingly Similar Enterprises," *Handbook of Organizational Economics*, 2013, pp. 342–372.

Gilligan, Thomas W and Keith Krehbiel, "Asymmetric information and legislative rules with a heterogeneous committee," *American journal of political science*, 1989, pp. 459–490.

Giuliano, Paola and Nathan Nunn, "Understanding cultural persistence and change," *The Review of Economic Studies*, 2021, *88* (4), 1541–1581.

Golub, Benjamin and Matthew O Jackson, "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2010, *2* (1), 112–149.

Griffin, Kenneth W and Gilbert J Botvin, "Evidence-based interventions for preventing substance use disorders in adolescents," *Child and Adolescent Psychiatric Clinics*, 2010, *19* (3), 505–526.

Grossman, Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 1986, *94* (4), 691–719.

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

Gulesci, Selim, Sam Jindani, Eliana La Ferrara, David Smerdon, Munshi Sulaiman, and H Young, "A stepping stone approach to norm transitions," 2023.

Halac, Marina, "Relationship building: conflict and project choice over time," *The Journal of Law, Economics, & Organization*, 2014, *30* (4), 683–708.

_ **, Navin Kartik, and Qingmin Liu**, "Optimal contracts for experimentation," *The Review of Economic Studies*, 2016, *83* (3), 1040–1091.

Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 1990, *98* (6), 1119–1158.

Hörner, Johannes, Nicolas Klein, and Sven Rady, "Overcoming Free-Riding in Bandit Games," *The Review of Economic Studies*, 11 2022, *89* (4), 1948 – 1992.

Hortaçsu, Ali, Olivia R Natan, Hayden Parsley, Timothy Schwieg, and Kevin R Williams, "Organizational structure and pricing: Evidence from a large us airline," *The Quarterly Journal of Economics*, 2024, *139* (2), 1149–1199.

Ide, Enrique, "Dual Moral Hazard and The Tyranny of Success," *American Economic Journal: Microeconomics*, 2024.

Kartik, Navin, SangMok Lee, Tianhao Liu, and Daniel Rappoport, "Beyond unbounded beliefs: How preferences and information interplay in social learning," *Econometrica*, 2024, *92* (4), 1033–1062.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68.

Krieger, Joshua Lev, Christopher Stanton, and James Barnett, "Forecasting ClimaCell," Dec 2019.

Krishna, Vijay and John Morgan, "A model of expertise," *The Quarterly Journal of Economics*, 2001, *116* (2), 747–775.

Kuran, Timur and William H Sandholm, "Cultural integration and its discontents," *The Review of Economic Studies*, 2008, *75* (1), 201–228.

Lee, In Ho, "On the convergence of informational cascades," *Journal of Economic theory*, 1993, *61* (2), 395–411.

Levin, Jonathan, "Multilateral contracting and the employment relationship," *The Quarterly Journal of Economics*, 2002, *117* (3), 1075–1103.

_ , "Relational Incentive Contracts," *American Economic Review*, June 2003, *93* (3), 835–857.

Lewis, Kemper and Farrokh Mistree, "Collaborative, Sequential, and Isolated Decisions in Design," in "International Design Engineering Technical Conferences and Computers and Information in Engineering Conference," Vol. 80456 American Society of Mechanical Engineers 1997, p. V003T30A018.

Li, Hongyi and Eric Van den Steen, "Birds of a feather… Enforce social norms? Interactions among culture, norms, and strategy," *Strategy Science*, 2021, *6* (2), 166–189.

Liker, Jeffrey K and Thomas Y Choi, "Building deep supplier relationships," *Harvard business review*, 2004, *82* (12), 104–113.

Lippman, Steven A. and John J. McCall, "The Economics Of Job Seach: A Survey," *Economic Inquiry*, 1976, *14* (2), 155–189.

Liu, Qingmin and Yu Fu Wong, "Strategic Exploration: Pre-emption and Prioritization," *Review of Economic Studies*, 2023, p. rdad084.

Lizzeri, Alessandro, Eran Shmaya, and Leeat Yariv, "Disentangling Exploration from Exploitation," Technical Report, National Bureau of Economic Research 2024.

Loury, Glenn C, "Self-censorship in public discourse: A theory of "political correctness" and related phenomena," *Rationality and Society*, 1994, *6* (4), 428–461.

Macleod, W. Bentley and James Malcomson, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 1989, *57* (2), 447–80.

Manski, Charles F. and Joram Mayshar, "Private Incentives and Social Interactions: Fertility Puzzles in Israel," *Journal of the European Economic Association*, 3 2003, *1* (1), 181–211.

Michaeli, Moti and Daniel Spiro, "Norm conformity across societies," *Journal of public economics*, 2015, *132*, 51–65.

Miller, Dale T, "A century of pluralistic ignorance: what we have learned about its origins, forms, and consequences," *Frontiers in Social Psychology*, 2023, *1*, 1260896.

_ **and Deborah A Prentice**, "Collective errors and errors about the collective," *Personality and Social Psychology Bulletin*, 1994, *20* (5), 541–550.

Moorthy, K Sridhar and Ivan PL Png, "Market segmentation, cannibalization, and the timing of product introductions," *Management science*, 1992, *38* (3), 345–359.

Morris, Stephen, "Political correctness," *Journal of political Economy*, 2001, *109* (2), 231–265.

Ottaviani, Marco and Peter Norman Sørensen, "Reputational cheap talk," *The Rand journal of economics*, 2006, *37* (1), 155–175.

Prentice, Deborah A and Dale T Miller, "Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm.," *Journal of personality and social psychology*, 1993, *64* (2), 243.

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

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

Rappoport, Daniel, "Reputational Delegation," 2022.

Reshidi, Pëllumb, Alessandro Lizzeri, Leeat Yariv, Jimmy Chan, and Wing Suen, "Sequential Sampling by Individuals and Groups: An Experimental Study," American Economic Review: Insights 2025. Forthcoming.

Robbins, Herbert, "Some aspects of the sequential design of experiments," *Bulletin of the American Mathematical Society*, September 1952, *58* (5), 527–535.

Scharfstein, David S and Jeremy C Stein, "Herd behavior and investment," *The American economic review*, 1990, pp. 465–479.

Schroeder, Christine M and Deborah A Prentice, "Exposing pluralistic ignorance to reduce alcohol use among college students 1," *Journal of Applied Social Psychology*, 1998, *28* (23), 2150–2180.

Sheikh, Aziz, GBD 2016 Alcohol Collaborators et al., "Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016," *The Lancet*, 2018, *392* (10152), 1015.

Smith, Lones and Peter Sørensen, "Pathological outcomes of observational learning," *Econometrica*, 2000, *68* (2), 371–398.

_ , Peter Norman Sørensen, and Jianrong Tian, "Informational herding, optimal experimentation, and contrarianism," *The Review of Economic Studies*, 2021, *88* (5), 2527–2554.

Spolaore, Enrico, "The Political Economy of European Integration," NBER Working Paper 21250, National Bureau of Economic Research, June 2015.

Strulovici, Bruno, "Learning while voting: Determinants of collective experimentation," *Econometrica*, 2010, *78* (3), 933–971.

Sunstein, Cass R, *Going to extremes: How like minds unite and divide*, Oxford University Press, 2009.

_ , *How change happens*, Mit Press, 2019.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–65.

Takeuchi, Hirotaka and Ikujiro Nonaka, "The New New Product Development Game," *Harvard business review*, 1986, *64* (1), 137–146.

Terwiesch, Christian, Christoph H Loch, and Arnoud De Meyer, "Exchanging preliminary information in concurrent engineering: Alternative coordination strategies," *Organization Science*, 2002, *13* (4), 402–419.

Tevyaw, Tracy O'Leary, Brian Borsari, Suzanne M Colby, and Peter M Monti, "Peer enhancement of a brief motivational intervention with mandated college students.," *Psychology of Addictive Behaviors*, 2007, *21* (1), 114.

Tirole, Jean, *The Theory of Industrial Organization*, Cambridge, MA: MIT Press, 1988.

_ , "Safe spaces: shelters or tribes ?," 2023. Working Paper.

Urgun, Can, "Restless Contracting," 2021. Working Paper.

Valle, Sandra and Daniel Vázquez-Bustelo, "Concurrent Engineering Performance: Incremental versus Radical Innovation," *International Journal of Production Economics*, 2009, *119* (1), 136–148.

Vanpoucke, Evelyne, Ann Vereecke, and Kenneth K. Boyer, "Triggers and patterns of integration initiatives in successful buyer–supplier relationships," *Journal of Operations Management*, 2014, *32* (1), 15–33.

Varian, Hal R, "Sequential Contributions to Public Goods," *Journal of Public Economics*, 1994, *53* (2), 165–186.

Venables, Sarah, "Essays on organizational economics and the heterogeneity of firm performance." Ph.d. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA 2013.

Wang, Shaoda and David Y. Yang, "Policy Experimentation in China: The Political Economy of Policy Learning," *Journal of Political Economy*, forthcoming.

Watson, Joel, "Starting small and renegotiation," *Journal of Economic Theory*, 1999, *85* (1), 52–90.

_ , "Starting Small and Commitment," *Games and Economic Behavior*, 2002, *38* (1), 176–199.

Weitzman, Martin, "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3), 641–654.

West, Steven L and Keri K O'Neal, "Project DARE outcome effectiveness revisited," *American journal of public health*, 2004, *94* (6), 1027–1029.

Westphal, James D and Michael K Bednar, "Pluralistic ignorance in corporate boards and firms' strategic persistence in response to low firm performance," *Administrative science quarterly*, 2005, *50* (2), 262–298.

Williamson, Oliver E., *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press, 1975.