The Art of Waiting*

Vasundhara Mallick† Ece Teoman‡

November 27, 2022

Click here for the latest version

Abstract

This paper studies delegated project choice without commitment: a principal and an agent have conflicting preferences over which project to implement, and the agent is privately informed about the availability of projects. We consider a dynamic setting in which, until a project is selected, the agent can propose a project, and the principal can accept or reject the proposed project. Importantly, the principal cannot commit to his responses. In this setting, the agent has an incentive to hold back on proposing projects that the principal favors so that the principal approves a project favored by the agent. Nevertheless, the principal achieves his commitment payoff in an equilibrium of the game in the frequent-offer limit. This high payoff equilibrium showcases the art of waiting and contrasts with Coasian logic: by giving proposer power to the agent, the principal makes it credible to reject his dispreferred projects until later in the game giving the agent an incentive to propose principal-preferred projects earlier on. We apply these results to the economics of organization and merger analysis. In particular, these results suggest a pure strategic gain from giving workers the initiative to pursue their desired projects, and to solicit their ideas "bottom-up" rather than issuing "top-down" commands.

*We are grateful to S. Nageeb Ali, Vijay Krishna, and Rohit Lamba for their invaluable guidance and continuous support. We greatly benefited from suggestions from Yu Awaya, Ian Ball, Kalyan Chatterjee, Nima Haghpanah, Berk Idem, Navin Kartik, Andreas Kleiner, R. Vijay Krishna, Wenhao Li, Joshua Mollner, and Ran Shorrer. We thank the participants of the Pennsylvania Economic Theory, Midwest Theory, and Women in Economic Theory conferences for their insightful comments.

†Department of Economics, Pennsylvania State University. Email: vxm266@psu.edu.

‡Department of Economics, Pennsylvania State University. Email: eceteoman@psu.edu.

1 Introduction

This paper considers a principal-agent problem with the following two features: (i) the agent knows what actions or "projects" are feasible and the principal does not, and (ii) the interests of the two parties are not aligned. Such principal-agent problems abound. Consider the interaction between the Board of Directors (principal) of a firm and its CEO (agent). The CEO may be better informed about what actions the firm can undertake, and unlike the Board of Directors, is motivated by empire building. In such cases, the Board of Directors cannot blithely assume that the CEO selects actions purely for shareholder interests. Another example is that of an antitrust authority deciding which mergers to approve: It would approve only those mergers that enhance efficiency or consumer welfare, but firms would like to propose only those mergers that increase industry profits. In such settings, what should the principal do?

These issues have been studied in the literature on *project selection* problems, initiated by the seminal work of Armstrong & Vickers (2010) and Nocke & Whinston (2013). The dominant approach presumes that the principal can commit to which projects he would accept in a one-shot interaction. But in many settings, the principal may be unable to commit, particularly if projects can be proposed across several rounds. If the agent does not propose any project that the principal deems acceptable, the principal may then infer that such projects are infeasible and capitulate. Anticipating this reaction, the agent may then wish to hold back on proposing projects that the principal finds acceptable. How well can the principal do and can he obtain his commitment payoff?

We investigate this question in a dynamic framework. The agent is privately informed about which projects are feasible at time 0. In each round $t \in \{0, 1, 2, \ldots\}$, the agent can propose a project that is feasible or stay silent; if a project is proposed, the principal can accept or reject it. This process continues until a project is accepted—in which case, players obtain payoffs from that selected project—or all feasible projects are rejected, after which all players obtain payoffs from the status quo. We consider the frequent-offer limit of this model, a sequence of games where the period length vanishes.

In such a setting, one may anticipate that the principal would suffer a significant loss of payoffs relative to the (static) commitment benchmark: after all, given the logic sketched above, the principal may capitulate when no acceptable project is proposed.

Moreover, our extensive form endows the agent with *proposer power*: The principal is effectively giving up bargaining or proposer power, so even complete-information intuitions suggest that the principal will do poorly in the dynamic game. Our main result, informally stated, is

> **Theorem**. *In the frequent-offer limit, the principal attains his commitment payoff in an equilibrium of the game.*

The key idea underlying our main result is that endowing the agent with the right to propose circumvents the principal's commitment problem. Our high payoff equilibrium stipulates that the agent and the principal wait for many rounds before respectively proposing and accepting any project that the agent greatly prefers to other projects and the principal disprefers (but prefers to the status quo). We show that this behavior is sequentially rational, even at histories where the principal attributes probability 1 to the agent only having such projects. If the agent proposes such projects earlier than specified, the principal believes that the agent must have other projects at her disposal. Such "punishment through beliefs" incentivizes the agent not to propose such projects earlier than stipulated and as such solves the principal's commitment problem. Because the agent anticipates such delays to get her preferred projects approved, she is willing to propose, as in the commitment benchmark, feasible projects that the principal prefers (and she may disprefer). We observe that it is essential that the agent has proposal power for the commitment problem to be solved. By contrast, if the principal were the one making all the offers, Coasian forces would take over, resulting in him granting full discretion to the agent.

We view this finding to be of more than just theoretical interest. It suggests a gain to organizations from allowing the agents—be it CEOs, managers, or employees—to propose projects rather than issuing "top-down" commands from the principal. One may envision that such "bottom-up" organizational structures emerge to motivate the agent, as implemented projects follow their initiative as proposed in Aghion & Tirole (1997). As highlighted in Dessein (2002), another well-established rationale for delegating decision rights to lower levels of organizational hierarchy is circumventing the loss of information. We offer a complementary channel through which bottom-up structures benefit the principal. Our work also suggests the gains that come to antitrust authorities, venture capital boards, and grant funding agencies to allow the agent to be the one to propose projects flexibly rather than constraining the agent to propose only certain projects.

Apart from our interest in obtaining the static commitment benchmark in equilibria of the game, our paper also studies the optimal delegation protocol in dynamic mechanisms without transfers. We show that the delegation protocol we study—that in which the agent proposes or stays silent, and the principal accepts or rejects—does as well as any delegation protocol under commitment, and always achieves the payoff from the best static stochastic mechanism. When there are only two possible projects, we show that this commitment benchmark is always attained in an equilibrium; with more than two possible projects, some additional assumptions are needed to achieve the commitment benchmark without commitment.

The rest of the paper is organized as follows. We first present the related literature. Section 2 introduces the setup, describes the sequential delegation game, and establishes a commitment benchmark. We present our main result in Section 3 which exhibits the main forces and the intuition behind them in the cleanest way. We establish a commitment benchmark as an upper bound for the principal's payoff and our main result shows that the commitment payoff is always attained in an equilibrium of the game. In Section 4, we include the general analysis for N projects. There, we work in a class of delegation protocols and show that our sequential delegation game is an optimal protocol under commitment. We also extend our main result beyond two projects under some regularity conditions. Section 5 concludes the paper.

1.1 Related Literature

Our paper builds on prior work on delegation in project selection problems. Most notably, Armstrong & Vickers (2010) and Nocke & Whinston (2013) study this problem in a one-shot interaction where the principal commits to a deterministic mechanism. By contrast, we consider a dynamic model where the agent makes proposals, and crucially, the principal cannot commit. Nevertheless, we show that the principal can achieve the payoff he would attain by committing to the best stochastic mechanism of the one-shot setting.[1]

Aghion & Tirole (1997) shows that when the project payoffs depend on the costly effort of the principal and agent, the principal may find it optimal to delegate the

[1] In recent work, Guo & Shmaya (2022) relaxes the restrictions to deterministic policies and allocates the proposal power to the agent but retains the commitment assumption. The optimal mechanism that minimizes the principal's worst-case regret implements the projects with high payoffs for the principal but implements the ones with lower payoffs with interior probabilities.

decision-making authority fully to the agent. This paper highlights delegation as a tool of motivation within organizations. Although the interaction in our setting is different, we offer a complementary reason to delegate in organizations by showing that the principal attains a higher payoff by seeking proposals from the agent.

Dessein (2002) highlights an informational rationale for delegation. He shows that the principal is better off delegating decisions to an informed (but biased) agent rather than attempting to elicit that information through cheap talk; this result, however, obtains only when the interests are mildly misaligned. In our setup, the principal and agent have diametrically opposed preferences over projects, and the principal's gain from letting the agent propose projects emanates from a different source. Crucially, in our analysis, the principal does not fully delegate choices to the agent.

An alternative approach in the project selection literature is to model the interaction as a cheap talk game as in Che, Dessein, & Kartik (2013) and Schneider (2015). Che, Dessein, & Kartik (2013) finds that in the presence of a bias regarding the outside option, the agent tends to propose unconditionally better projects for the principal to secure the approval for implementation. As its dynamic extension, Schneider (2015) finds that the dynamic interaction allows for different equilibrium outcomes. It characterizes a mixing equilibrium where the agent randomizes between pandering and not, and a waiting equilibrium similar to ours in spirit where the agent waits to persuade the principal that the unconditionally worse project has a better payoff realization. In contrast to our paper, the waiting equilibrium there results in inefficiency, and the mixing equilibrium fares better than the waiting.

Our paper also relates to the literature on mechanism design with hard evidence, starting with Green & Laffont (1986) and Bull & Watson (2007). In these papers, the setting is endowed with an evidentiary structure under which different mechanisms are compared and Revelation Principles are established. We start with a type dependence of evidentiary actions in delegation protocols and establish an evidentiary structure. Static mechanisms with this evidentiary structure serve as an upper bound for any outcome that can be implemented with any mechanism with this particular type dependence of actions. The Revelation Principle from these papers does not hold directly in ours because the agents's action also restricts what the principal can choose, and evidentiary actions can be taken at multiple nodes, which these papers do not allow. The paper closest to our mechanism design analysis is Deneckere & Severinov (2008) which has a Revelation Principle that allows for evidentiary actions at multiple nodes. However, this result does not follow directly in our setting either. If the agent is the proposer, the principal cannot implement a project she does not propose, so evidentiary actions have significance beyond providing verifiable information.

The possibility of attaining the commitment payoff without any commitment power is also present in other contexts. In Glazer & Rubinstein (2001) and Glazer & Rubinstein (2004), the design of optimal debate mechanisms is studied in a setting with an uninformed principal and privately informed agent. They also show that the commitment payoff can be attained in equilibrium, but the setting is quite different. Ali, Kartik, & Kleiner (2022) shows that the commitment payoff can be achieved in a sequential bargaining game between a Proposer and Vetoer, however, the force that enables it is the single-peaked preferences of Vetoer. By contrast, in our model, commitment payoffs are achieved through the combination of the informed party making offers and the uninformed party being able to "punish through beliefs". Gerardi, Hörner, & Maestri (2014) shows the possibility of achieving commitment payoff without commitment in a bilateral trade setting. Our setup is different in several aspects; the most important being that we have no transfers, so incentives need to be provided purely through acceptance and rejection of proposals.

2 The Model

A principal (he) and an agent (she) jointly choose a project to implement. A project is a pair of payoffs $(\alpha, \pi) \in \mathbb{R}^2_{++}$, where α is the agent's payoff and π is the principal's. There are two *possible* projects, denoted by $\mathcal{N} = \{g, b\}$: a good project g with payoffs (α_g, π_g) and a bad project b with payoffs (α_b, π_b). The players are expected utility maximizers and while any project is preferred to the status quo by both, they have conflicting preferences over the projects: $\pi_g > \pi_b > 0$ and $\alpha_b > \alpha_g > 0$.

The challenge is that only the agent knows which projects are *available*: her *type* represents the set of available projects. In this case, the agent has four possible types:

- $E = \varnothing$, the *empty* type with no available projects;
- $G = \{g\}$, the *good* type with only the good project available;
- $B = \{b\}$, the *bad* type with only the bad project available;
- $M = \{g, b\}$, the *mixed* type with both projects available.

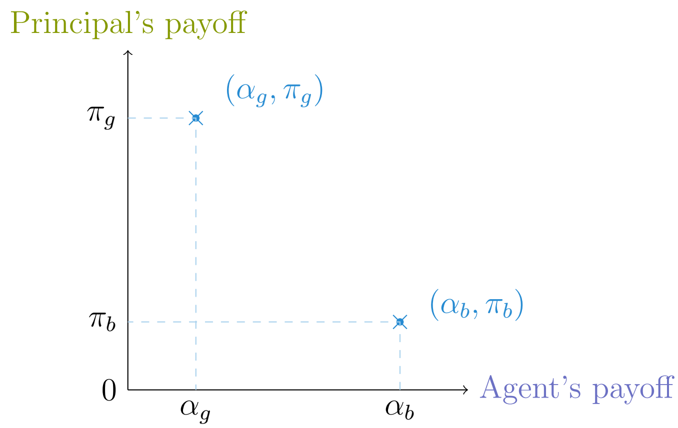

Figure 1: The project space with the principal-preferred good project (α_g, π_g) and the agent-preferred bad project (α_b, π_b).

The agent's type is drawn from $\mathcal{S} \equiv 2^{\mathcal{N}}$ according to the probability distribution $\mu : \mathcal{S} \to [0, 1]$. To simplify notation, we refer to $\mu(S_i)$ as μ_i whenever there is no ambiguity.

The sequential delegation game proceeds as follows. In each period $t = 0, 1, 2, \ldots$, The agent can propose a single available project or stay silent. Following the proposal of a project, the principal can *accept* or *reject*. If the principal rejects a proposal, both players get a payoff of 0 in that period and the game proceeds to the next period. If the principal accepts the proposal of project i at time t, the game ends: the principal's payoff is $\delta^t \pi_i$ and the agent's is $\delta^t \alpha_i$ where $\delta \in (0, 1)$ is the common discount factor. We often focus on the case of $\delta \to 1$ and interpret it as the frequent-offer limit of the game.

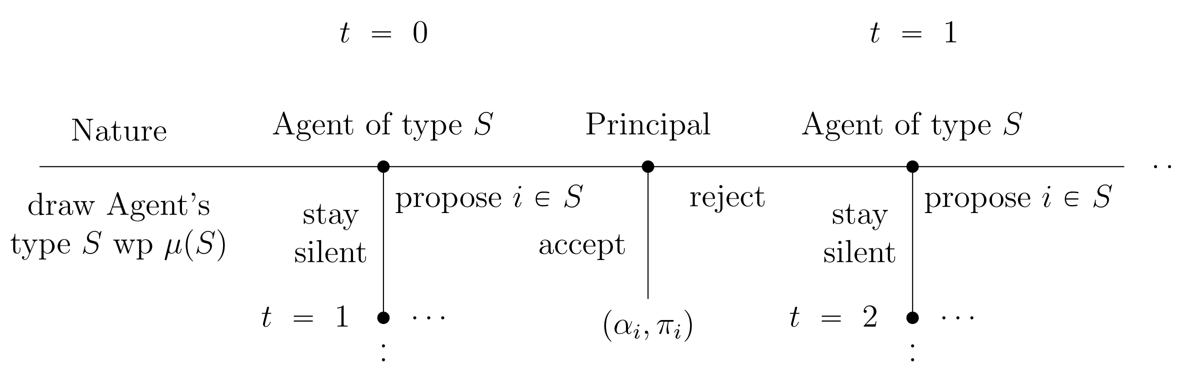

Figure 2: Timeline of the sequential delegation game.

A history in the sequential delegation game is a sequence of proposals that have been rejected. The principal's strategy maps each history and current proposal to an acceptance probability. The agent's strategy maps each history to a probability distribution over available projects and silence. Our equilibrium concept is Perfect Bayesian Equilibrium; both players play sequentially rationally and the principal's beliefs about the agent's type are updated according to Bayes' rule whenever possible.

Although we view this model to speak to a number of applications, let us discuss here a specific application that serves as a running example throughout the paper. Consider the interaction, within a firm, between the Board of Directors and its CEO. The Board of Directors (the principal) would prefer the CEO to undertake a highly profitable project that brings less benefit to the CEO. In contrast, the CEO favors a less profitable project that expands her influence and allows her to pursue empire-building plans. The payoffs of these projects are known, and each is preferred to the status quo. But the CEO is privately informed about which projects are feasible and she makes proposals to the Board. The Board chooses whether to accept a proposal but cannot commit to its responses.

3 The Benefits of Giving Up Control

This section presents our main result. We first establish a commitment benchmark to find an upper bound to what the principal could achieve with commitment power. Then our main result shows that the commitment benchmark is attainable in an equilibrium of the sequential delegation game where the principal lacks commitment power. Finally, we contrast our main result to the case where the principal makes proposals and argue that attaining the commitment payoff may not be possible in that case.

3.1 Commitment Benchmark

As a commitment benchmark, we define a class of static, stochastic mechanisms with type-dependent message spaces to serve as an upper bound and we refer to them as mechanisms hereafter. In Section 4.1, we prove that this class of mechanisms is indeed an upper bound to what the principal can achieve if he were able to commit to a strategy in the sequential delegation game.[2]

The set of messages that a type S of the agent can send is defined to be $M(S) = 2^S$, so each type can report any subset of her available projects as a message in the mechanism. A mechanism is a tuple (M, q), where $M = \Pi_{S \subseteq \mathcal{N}} M(S)$ is the set of all possible messages and q is the outcome function. The outcome function q maps any report $S \in M$ to a probability distribution over $S \cup \varnothing$ where $\varnothing$ represents the null project. In this mechanism, q_{Si} represents the probability of implementing project i when the type S is reported for each project $i \in S$ and for each type $S \in \mathcal{S}$. We include $\varnothing$ to accommodate the possibility of obtaining the status quo when a report is made. We define a mechanism to be incentive compatible (IC) if no type finds it optimal to report a strict subset.

A feasible allocation in this mechanism is $q_{Si} \in [0, 1]$ for each $i \in S$ such that $\sum_{i \in S} q_{Si} \leqslant 1$ for a given type S. The principal-optimal mechanism maximizes the principal's payoff by choosing implementation probabilities for projects in each type, subject to incentive compatibility constraints. We have the following IC conditions: each type is able to report itself and the empty type; additionally, the mixed type can also report the good or bad type. The principal-optimal mechanism solves the following problem:

$$\max_{q_{Gg}, q_{Bb}, q_{Mg}, q_{Mb} \in [0,1]} \quad q_{Gg}\mu_g\pi_g + q_{Bb}\mu_b\pi_b + q_{Mg}\mu_m\pi_g + q_{Mb}\mu_m\pi_b$$

$$\text{subject to} \qquad q_{Sg}\alpha_g + q_{Sb}\alpha_b \geqslant q_{S'g}\alpha_g + q_{S'b}\alpha_b \qquad \forall S, S' : S' \subseteq S \quad (IC_{SS'})$$

where q_{Si} is the probability of implementing project i when type S is reported and $\text{IC}_{SS'}$ denotes the IC constraint for type S to not report type $S' \subseteq S$.

Note that $1 - q_{Gg}$, $1 - q_{Bb}$, and $1 - q_{Mg} - q_{Mb}$ represent the probabilities of obtaining the status quo. We denote the principal's maximum payoff by $v(\mu, \mathcal{N})$. Below, we make a few simplifying observations about this problem and briefly discuss the reasoning behind each.

[2] We also find that committing to strategies within the game is equivalent to committing to an incentive compatible mechanism in this particular class of mechanisms. We define the class of mechanisms formally and establish the equivalence in Section 4. There we also look at commitment in a general class of extensive form games and show that commitment payoff in our game constitutes an upper bound for what can be achieved in any other game with commitment power.

Observation. No type finds it profitable to report the empty type, so IC_{ge}, IC_{be}, and IC_{me} are all redundant.

As the empty type has no projects available, the mechanism does not implement any project from this type and she gets 0 payoff. Since each project has strictly positive payoffs for both players, each type of the agent gets a weakly higher payoff by reporting her own type compared to reporting the empty type. Hence, the IC constraints for the other types to not report the empty type, IC_{ge}, IC_{be}, and IC_{me}, are all redundant.

Observation. In any optimal mechanism, we must have $q^*_{Mg} + q^*_{Mb} = 1$.

When the mixed type is reported, the probabilities of implementing projects must sum up to 1 in the optimal mechanism. Suppose otherwise, $q^*_{Mg} + q^*_{Mb} < 1$. Then, we can increase either q^*_{Mg} or q^*_{Mb} and have new implementation probabilities $(q^{**}_{Mg}, q^{**}_{Mb}) > (q^*_{Mg}, q^*_{Mb})$. It must be the case that the IC constraints for the mixed type still hold as

$$q^{**}_{Mg}\alpha_g + q^{**}_{Mb}\alpha_b > q^*_{Mg}\alpha_g + q^*_{Mb}\alpha_b \geqslant q_{Gg}\alpha_g$$

$$q^{**}_{Mg}\alpha_g + q^{**}_{Mb}\alpha_b > q^*_{Mg}\alpha_g + q^*_{Mb}\alpha_b \geqslant q_{Bb}\alpha_b.$$

Moreover, the principal obtains a strictly higher payoff as well. Then, $q^*_{Mg} + q^*_{Mb} < 1$ cannot be part of the optimal mechanism.

Observation. The incentive compatibility constraint for the mixed type to not report the good type, IC_{Mg}, is redundant.

We know from the previous observation that when the mixed type is reported, the probabilities of implementing projects sum up to 1. It implies that the payoff of the mixed type will be at least α_g and at most α_b. On the other hand, the good type's payoff is at least 0 and at most α_g as $q_{Gg} \in [0, 1]$. We can then conclude that the mixed type's payoff is always weakly higher than the good type's, and IC_{Mg}, is redundant.

Observation. In any optimal mechanism, we must have $q_{Gg} = 1$.

As the IC constraint for the mixed type to not report the good type is redundant and there is no other type that can report the good type, the optimal mechanism must always implement the good project from the good type with certainty.

The problem now reduces to whether or not the principal wants to extract the good project from the mixed type. Extracting the good project from the mixed type comes at a cost: the bad project from the bad type needs to be implemented with some interior probability so that IC_{Mb} holds and the mixed type does not imitate the bad type. Next, we define two forms the principal-optimal mechanism can take in the case of two possible projects.

Definition. *The optimal pooling mechanism implements the bad project from the agent's bad and mixed types:* $q^*_{Gg} = 1, q^*_{Bb} = 1, q^*_{Mg} = 0, q^*_{Mb} = 1$.

The optimal separating mechanism implements the good project from the mixed type and the bad project from the bad type with an interior probability: $q^*_{Gg} = 1, q^*_{Bb} = \frac{\alpha_g}{\alpha_b}, q^*_{Mg} = 1, q^*_{Mb} = 0$.

If the bad project is implemented from both bad and mixed types, we call this optimal mechanism *pooling* as the mixed type is pooled with the bad. On the other hand, if the good project is implemented from the mixed type whereas the bad project is implemented from the bad type with an interior probability and we refer to this optimal mechanism as *separating* as the mixed type is separated from the bad in this case.

Whether the principal finds it worthwhile to get the good project from the mixed type depends on the parameters. More specifically, it depends on the gains and losses from separating the mixed type from the bad compared to pooling her with the bad. The gains from separating compared to pooling is obtaining a higher payoff from the mixed type $(\pi_g - \pi_b)\mu_m$; and the losses from separating compared to pooling is foregoing some of the payoff from the bad type $\left(1 - \frac{\alpha_g}{\alpha_b}\right)\pi_b\mu_b$. The optimal mechanism takes the separating form if the gains are higher than the losses, and the pooling form otherwise.

We can represent the comparison between the gains and losses in terms of the relative likelihood of the types. Let $\lambda = \frac{\mu_m}{\mu_b}$ be the likelihood ratio of mixed type compared to bad type. Separation is optimal whenever $\lambda > \lambda^*$ where λ^* is defined as

$$\lambda^* = \frac{(1 - \frac{\alpha_g}{\alpha_b})\pi_b}{(\pi_g - \pi_b)},$$

which depends on the type distribution μ and the set of projects $\mathcal{N}$.

We can now state our commitment benchmark result including the form it takes, the conditions for each form to be optimal, and the upper bound they each generate for the principal.

Proposition 1. *The principal-optimal mechanism is either pooling or separating.*

a) The pooling mechanism is optimal when $\lambda \leqslant \lambda^*$ and the principal's maximum payoff is $v^p(\mu, \mathcal{N}) = \mu_g \pi_g + \mu_b \pi_b + \mu_m \pi_b$.

b) The separating mechanism is optimal when $\lambda > \lambda^*$ and the principal's maximum payoff is $v^s(\mu, \mathcal{N}) = \mu_g \pi_g + \mu_b \frac{\alpha_g}{\alpha_b} \pi_b + \mu_m \pi_g$.

This result tells us that the principal-optimal mechanism takes one of the two forms that are common in adverse selection problems. The result also suggests that we can see which form it takes by simply checking the relative likelihood ratio of types against a threshold.

In the context of our model's applications, we can think about this static commitment benchmark as the Board of Directors having the ability to sign court-enforced contracts with the CEO for each contingency. The result above implies that the Board would choose one of the two options for such a contract:

- A contract that grants full discretion to the CEO for taking the company in her preferred feasible direction.

- A contract that restricts her to choosing the Board's preferred action whenever it is feasible to do so. The contract allows the CEO's preferred action only when the alternative is infeasible, and even then with a small chance. Within this contract, the CEO can rarely carry out her empire-building plans: when it is the only feasible action and not with certainty.

Moreover, the Board can decide which contract to choose based on how likely the Board's preferred direction is to be feasible. Our result offers a simple solution to a seemingly complicated problem.

After establishing the principal's optimal commitment payoff, we turn our attention back to the sequential delegation game where the principal cannot commit to his responses. In particular, we are interested to explore how well the principal can do in equilibria of the game when he is constrained by sequential rationality and whether he can obtain his commitment payoff.

3.2 Implementing the Commitment Benchmark as Equilibria

In the sequential delegation game, the principal lacks proposal power and this translates into a reduced level of control over which project is implemented. This loss of control combined with the lack of commitment power and information may lead us to expect that the principal's sequential rationality results in forfeiting all the surplus.

Our main result, contradicting this initial expectation, establishes that there is always an equilibrium of the sequential delegation game where the principal attains his commitment payoff in the frequent-offer limit. In the absence of commitment power, the principal uses time as a screening device to achieve separation and in response, the agent uses the option of staying silent to signal her type through costly delay. We highlight the principal's lack of proposal power as the main force behind the possibility of separation. We can now state our main result where, focusing on the frequent-offer limit, we describe an equilibrium in which the principal's expected payoff is arbitrarily close to his commitment payoff for each form the optimal mechanism takes.

Theorem 1. *There is always an equilibrium of the sequential delegation game in which the principal's payoff approximates his commitment payoff in the frequent-offer limit, as $\delta \to 1$. On-path behavior of the equilibria that attain the commitment benchmark is as follows.*

a) (Pooling) When the pooling mechanism is optimal, $\lambda \leqslant \lambda^*$, the pooling equilibrium attains the principal's commitment payoff: Each type of the agent proposes her favorite available project at $t = 0$. The principal accepts any proposal at $t = 0$.

b) (Separating) When the separating mechanism is optimal, $\lambda > \lambda^*$, the separating equilibrium approximates the principal's commitment payoff:

- *The agent's good and mixed types propose the good project at $t = 0$ and the bad type stays silent until $t^* := \min\{t : \alpha_g \geqslant \delta^t \alpha_b\}$ at which point she proposes the bad project.*
- *The principal accepts the good project at $t = 0$ and the bad project at t^*.*

The key idea underlying our main result is that the lack of proposal power helps with the principal's commitment problem. Our separating equilibrium displays the art of waiting as it specifies that the agent and principal wait for several rounds before proposing and accepting the bad project, respectively. We show that this behavior is sequentially rational, even at histories where the principal attributes probability 1 to the agent's bad type. Most importantly, the principal's lack of proposal power makes separation through costly delay possible.

We focus on the more interesting case of the separating equilibrium that replicates the separating optimal mechanism and show how it is indeed an equilibrium. Recall that the separating optimal mechanism implements the good project from the good and mixed types, $q_{Gg} = q_{Mg} = 1$, and the bad project from the bad type with an interior probability, $q_{Bb} = \frac{\alpha_g}{\alpha_b}$.

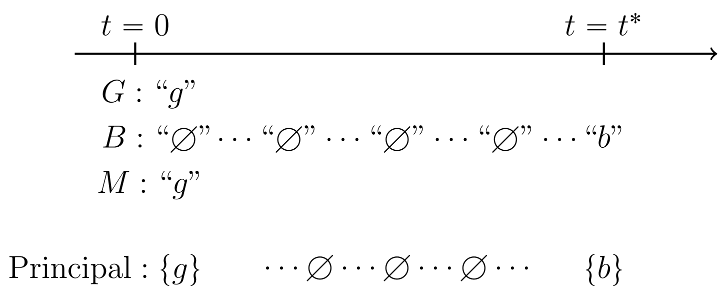

Figure 3: The timing of the proposals and accepted projects on the path of separating equilibrium of the sequential delegation game.

We can easily see that the on-path behavior of the separating equilibrium replicates the implementation probabilities of the commitment benchmark when it takes the separating form. We only need to make the observation that the discount for the bad project δ^{t^*} approaches $\frac{\alpha_g}{\alpha_b} = q_{Bb}$ as $\delta \to 1$. We now discuss the logic behind the separating equilibrium and argue that the strategies above indeed constitute an equilibrium with appropriate off-path beliefs.

> **Observation.** In order to separate the mixed type in the game, the bad project must be implemented at some period t^* such that $\delta^{t^*}\alpha_b \leqslant \alpha_g$.

It is important to note that separation is achieved in this equilibrium by delaying the implementation of the bad project as players use time as an instrument to signal and screen various types in the game. The delay in accepting the bad project should be long enough so that the mixed type proposes the good project at $t = 0$ rather than waiting until t^* to propose the bad project.

> **Observation.** In the strategy profile above, the agent's good and mixed types propose the good project at $t = 0$. How can we delay the implemen-

tation of the bad project until t^*? More precisely, how can we guarantee that

1. the bad project is not proposed before t^*;
2. and if proposed before t^*, the bad project is not accepted?

The principal's off-path beliefs are crucial to guarantee that the delay for the bad project can be sustained. In this equilibrium, if the bad project is proposed before t^*, the principal believes with probability 1 that the agent is of mixed type. The principal's certain belief dictates that he should accept only the good project starting at that period: he attributes probability one to the existence of the good project and anticipates the mixed type to propose it following his rejection of the bad project. Given the principal's off-path beliefs, no type of the agent proposes the bad project before t^*: The bad type would receive the status quo payoff of 0 if she did so, and the mixed type would get α_g at some $t' > 0$ which she could obtain at $t = 0$. Thus, "punishment through beliefs" is the answer to the question of how to guarantee the delay in equilibrium.

It is worth highlighting the power of the principal's beliefs in the separating equilibrium. Given the agent's strategy of proposing the good project immediately and delaying the proposal of the bad project, between the periods 0 and t^*, the principal believes that only the bad project can be available. In stark contrast, if the bad project is proposed in any of those time periods, the principal's updated belief now attributes probability one to both projects being available. Combined with the principal's inability to make proposals, this "punishment through beliefs" plays an important role in holding the separating equilibrium together.

If we revisit our organizational application, what are the implications of our main result for the case of a company's Board of Directors and CEO deciding on the direction it should move in? Given the concerns regarding the practicality of signing legally binding contracts, our main result about attaining a high payoff in the game has more relevance to organizations and other applications. We claim that for each type of contract the Board would sign if they could, Board can actually attain the same benefits in the corresponding equilibrium. If both directions were highly likely to be feasible, implementation of the CEO-preferred project would be delayed to ensure the CEO's attempts at empire-building are only allowed when the company cannot move in the Board's preferred direction.

We should emphasize the role of the principal's lack of proposal power in our game: It is a significant force behind the emergence of costly delay as a signaling device in equilibrium. The simple fact that a project must be proposed to be implemented helps the possibility of separation through delay. In combination with extremal off-path beliefs, this structure lends the principal an instrument akin to commitment power to conduct effective screening in equilibrium. We now make this point explicit by showing that separation would not be possible in an alternative game where the principal is the proposing party.

Consider an alternative dynamic delegation game where the principal makes proposals and the agent responds. It may initially seem that if the principal makes proposals, he has greater influence over what is implemented. However, it is precisely this control over proposals that leads to the principal's inability to achieve the commitment payoff in a game where he makes all the offers.

> **Observation.** The dynamic delegation game where the principal makes all the proposals has no equilibrium that attains the commitment payoff when it involves separating.

If the principal attempts to replicate the separating optimal mechanism by always permitting the good project and only permitting the bad project after period t^*, his sequential rationality would urge him to deviate. Given that the agent's good and mixed types accept the good project immediately, after observing silence in the first period, the principal would start accepting the bad project right away instead of waiting until t^* which leads to the mixed type also staying silent at $t = 0$. In the spirit of Coase conjecture, the principal allows both projects at the beginning of the game, rendering delay useless as a screening device. In fact, Li (2022) establishes that Coasian forces are prevalent across all equilibria and pooling is the unique equilibrium outcome of this game.

We conclude that in the absence of commitment power, the lack of proposal power helps the principal with overcoming the sequential rationality constraint to achieve his commitment payoff. This stark contrast in the outcomes suggests that if the Board of Directors decided to choose the direction of the company by issuing top-down commands to the CEO, obtaining a high payoff might not be possible as it is with the bottom-up process. Similarly, if an antitrust authority conducts merger reviews by imposing restrictions instead of seeking proposals from the firm, they might not be able to guarantee the consumer-optimal outcome.

4 A General Model with N projects

In this section, we consider the case when there are $N > 2$ possible projects. The set of all *possible* projects is $\mathcal{N} \equiv \{1, 2, \ldots, N\}$ where project i corresponds to (α_i, π_i). As in the two-project model, only the agent knows which projects are *available*; her *type* is $S \subseteq \mathcal{N}$ representing the set of available projects. The agent's type is drawn from $\mathcal{S} \equiv 2^{\mathcal{N}}$ according to the probability distribution $\mu : \mathcal{S} \to [0, 1]$.

In this N project setup, the interaction between the principal and agent might take several forms, which we call *delegation protocols*, depending on the institutional setting. For example, a possible delegation protocol is one where the agent proposes exactly one available project in each period, and the principal either accepts or rejects this proposal. Another possibility is the principal specifying the set of projects from which the agent is permitted to choose in each period. We first establish a general commitment result for a general class of delegation protocols and then tie this result back to our sequential delegation game.

4.1 A General Commitment Result

Until now, we only considered the two project case, where payoff from the static commitment benchmark can always be achieved in an equilibrium of the particular extensive form we consider. However, when we move beyond two projects, the problem becomes much harder and it is not clear if the commitment payoff can be achieved in an equilibrium of the sequential delegation game.

There are potentially two reasons why this could happen: (i) a wedge between the static stochastic mechanism and commitment in the dynamic game, or (ii) a sequential rationality consideration that impedes the principal from achieving his dynamic commitment payoff in an equilibrium of this game. To elucidate the source of this tension, we investigate here what the principal can achieve if he could commit to a strategy in the dynamic game.

We first define a general class of extensive forms, delegation protocols, that represent various forms of interaction between the principal and the agent. We then prove a Revelation Principle and show that anything that can be achieved in any extensive form can be achieved by a direct, static, stochastic mechanism with type-dependent message spaces. Then, we show that our sequential delegation game imposes no restrictions in the frequent-offer limit. More precisely, as $\delta \to 1$, for any static mechanism, there exists a commitment strategy of the principal in the sequential delegation game that attains the same payoff. Finally, we provide an example of a protocol where, even with the ability to commit to a strategy, the principal is not always able to achieve the optimal commitment payoff from a static mechanism. This highlights that some protocols may impose constraints beyond sequential rationality.

4.1.1 Protocols

We define a delegation protocol to be an extensive form game that specifies the proposer and what they are allowed to offer at any history. Formally, at any history h_t, the proposer is $P(h_t)$ and the set of permissible offers is $\mathcal{O}(h_t) \subseteq 2^{\mathcal{N}}$, so any offer $O(h_t)$ is a subset of $\mathcal{N}$.[3] When an offer is made by the proposer, the other responds by either accepting a project in the offer or rejecting the offer altogether. A history is a sequence of offers that have been rejected. [4]

The set of actions feasible for the agent at any history is type-dependent; if she is the proposer, she can only include available projects in her offer, and if the principal is the proposer, she can only accept an available project. So, if the agent of type S is the proposer at h_t, we must have $O(h_t) \subseteq S$. On the other hand, if the principal makes the offer $O(h_t)$, the agent can only accept a project in $O(h_t) \cap S$.

A delegation protocol is therefore simply a dynamic game with type-dependent action space for the agent at any history. If any project from an offer is accepted, the game ends, and players get their discounted payoffs; otherwise the game proceeds to the next period. Our equilibrium concept is Perfect Bayesian Equilibrium; both players play sequentially rationally and the principal's beliefs about the agent's type are updated according to Bayes' rule whenever possible.

[3] An offer can be $\varnothing$, this can be interpreted as the principal allowing nothing to be chosen by the agent at that history, or the agent choosing to stay silent.

[4] Tying this definition back to our sequential delegation game, it corresponds to the protocol where the proposer is the agent at each history, and the set of permissible offers is all singleton subsets of $\mathcal{N}$.

4.1.2 Revelation Principle

We consider the class of mechanisms defined in section 3.1, where the message space is type-dependent, so a type can report only subsets of her available projects. A mechanism q maps any report to a probability of implementation of each project in that report. These are mechanisms with evidence, as each type is only able to report a subset of the projects she has. The message space here satisfies the *normality* condition from Bull & Watson (2007)

We now prove a Revelation Principle for this setting; we show that any social choice function implementable in any protocol is also implementable by a mechanism in this class. We now define a social choice function and an *induced* social choice function to state our result.

Definition. *A social choice function (SCF) is a function f that maps a set $S \subseteq \mathcal{N}$, to a probability of implementation of each project in S, where $f_S(i)$ denotes the probability of implementing project $i \in S$ from type S.*

An *induced* social choice function is defined to be an SCF that is induced by a strategy of the principal in a protocol and a best response to that strategy. In order to better understand what an induced SCF is, fix a protocol. Consider any strategy of the principal in this protocol and a best response of the agent to this strategy. This pair of strategy and best response induce a probability distribution over outcomes for each type S, where an outcome (i, t) denotes project i being implemented at time period t. For any type S and project $i \in S$, we can condense the *discounted* probabilities of implementing i at different histories into a single probability, and this probability is denoted by $f^I_S(i)$.[5] The induced SCF is then the function f^I that maps any type S to a probability $f^I_S(i)$ of implementation of each $i \in S$.[6]

Proposition 2. *For any delegation protocol and any SCF f^I induced by a strategy of the principal and a best response of the agent, the static mechanism f^I is incentive compatible.*

Proof. Fix a delegation protocol and an induced SCF F^I in the protocol. Recall, from our description of protocols, that at any history, any action that is available to a type

[5]For example, if $i \in S$ is implemented with probability $\frac{1}{2}$ at $t = 0$, and with probability $\frac{1}{2}$ at $t = 1$, then $f^I_S(i) = \frac{1}{2} + \delta\frac{1}{2}$.

[6]The details of collapsing the probability of various outcomes involving i into a single probability can be found in the Appendix A.2.

S' is also available to S, where $S' \subseteq S$. This is because within the constraints of what the protocol permits, anything that S' can propose or accept, S can as well since S has all the projects S' has. Since the induced SCF comes from a best response of the agent, it must be that the payoff for S from f^I_S is weakly better than the payoff from $f^I_{S'}$, which is what S would get if she imitated the best response of S'. Thus, the incentive compatibility in the mechanism, which requires that no type should find it optimal to report a strict subset, is satisfied. □

Note that the standard Revelation Principle from Bull & Watson (2007) does not follow directly here as we do not start out with a fixed evidentiary structure under which we compare various static and dynamic mechanisms. Instead, we start out with type-dependent evidentiary actions, which can be taken at multiple nodes. Moreover, the proposal of the agent also limits what the principal can choose when the agent is the proposer, which is not a feature of standard mechanisms with evidence.

We now argue that as $\delta \to 1$, any SCF that is implementable in a static, stochastic mechanism is implementable in the sequential delegation game if the principal is able to commit to a strategy. It means that our sequential delegation game is an optimal protocol under commitment and sequential rationality is the only restriction it imposes on what is attainable in equilibrium.

Theorem 2. *Fix an SCF f. There exist a strategy of the principal and a best response of the agent in the sequential delegation game such that the induced SCF from this strategy and best response approaches to f as $\delta \to 1$.*

We provide the proof in the Appendix, but the idea is to fix an SCF f and construct a corresponding strategy in the sequential delegation game:

* According to this strategy, the first N time periods are reserved for *information elicitation* where the agent proposes the available projects in a particular order. Proposals in the first N periods are analogous to reports in the mechanism, and since the agent can only report projects she has, this captures the fact that a type can only report her subsets in the mechanism.

* In the time periods that follow, these reported projects are proposed again and accepted with probabilities such that the agent finds it optimal to report all her projects in the first N periods. As $\delta \to 1$, these probabilities approach the

implementation probabilities from the mechanism and the principal's payoff from this strategy approaches his payoff in the mechanism.

This result tells us that the sequential delegation game—the delegation protocol where the agent makes proposals and the principal responds—does as well as any delegation protocol under commitment, and always achieves the payoff from the optimal static stochastic mechanism. In contrast, there exist delegation protocols where the ability to commit to a strategy may not be enough to attain the commitment payoff from the optimal static mechanism.

Consider the example where there are three possible projects, $\mathcal{N} = \{1, 2, 3\}$, and three equally likely types in the support of μ with $\mathcal{S} = \{\{1, 2\}, \{2\}, \{2, 3\}\}$. The payoffs are:

$$\pi_1 = 8, \pi_2 = 3, \pi_3 = 1$$

$$\alpha_1 = 3, \alpha_2 = 8, \alpha_3 = 9$$

The optimal mechanism is as follows:

- From type $\{1, 2\}$, project 1 is implemented with probability one
- From type $\{2\}$, project 2 is implemented with probability $\frac{3}{8}$.
- From type $\{2, 3\}$, project 2 is implemented with probability one.

Consider the delegation protocol where the principal makes proposals by choosing a set of permissible projects in each period. We can show that there does not exist a commitment strategy for the principal in this protocol that would attain the payoff from the above mechanism.

The details are in the Appendix, but the intuition is as follows: The construction of the commitment strategy in the sequential delegation game (Theorem 2) includes eliciting information from the agent about her type through her proposals and conditioning future responses on these initial proposals. So, the implementation of project 2 in $\{2, 3\}$ can be made conditional on the proposal of 3. In this alternative protocol, the agent can only accept or reject, and not propose projects herself, and it limits the scope for information elicitation. Without the ability to condition future implementation of projects on the agent's own past proposals, the principal cannot separate type $\{2, 3\}$ and $\{2\}$.

4.2 Attaining the Commitment Payoff with N Projects

After establishing that our sequential delegation game is the principal-optimal protocol in a class of delegation protocols under commitment, we turn our attention back to attaining the commitment benchmark. It is natural to ask whether our main result holds beyond two projects and we find that it is not clear that this would always be the case.

In this section, we provide two sufficient conditions on the parameters under which there is always an equilibrium of the game that attains the commitment benchmark. We show through an example that the conditions are not necessary and also highlight another signaling opportunity for the agent by proposing redundant projects.

While the number of possible projects does not alter the game itself or the commitment benchmark, working in an unordered type space makes the problem significantly more complex. Even establishing the benchmark becomes a difficult problem itself as it now includes IC conditions for each subset of each type and the meaning of separating also becomes unclear. As a result, the problem loses its tractability.

In order to recover some of the lost tractability, we first turn our attention to a restricted class of parameters. More specifically, we consider the model under three assumptions about the payoffs of the projects and the types in the support of μ.

Assumption 1. *(Conflicting preferences) The set of projects $\mathcal{N}$ satisfies*

$$\pi_1 > \pi_2 > \ldots > \pi_{N-1} > \pi_N > 0;$$

$$\alpha_N > \alpha_{N-1} > \ldots > \alpha_2 > \alpha_1 > 0.$$

We start by assuming that the set of projects is such that the preferences of the players are diametrically opposed. When there are two projects, the only alternative to opposite preferences is identical preferences in which case the problem would not be interesting. Beyond two projects, however, there is an array of possible preferences. Under Assumption 1, the case where the principal and agent have exact opposite preferences simplifies the problem and allows us to obtain a clean extension beyond two projects.

Assumption 2. *(Linear payoffs) Any two projects* $i, j \neq 1$ *satisfy*

$$\frac{\pi_1 - \pi_i}{\alpha_i - \alpha_1} = \frac{\pi_1 - \pi_j}{\alpha_j - \alpha_1}.$$

We further simplify the complex incentives beyond two projects by Assumption 2 which requires all possible projects to lie on a line on $\mathbb{R}^2_{++}$. The linearity of the projects ensures that the magnitude of the payoffs does not become an additional consideration for the optimal mechanism.

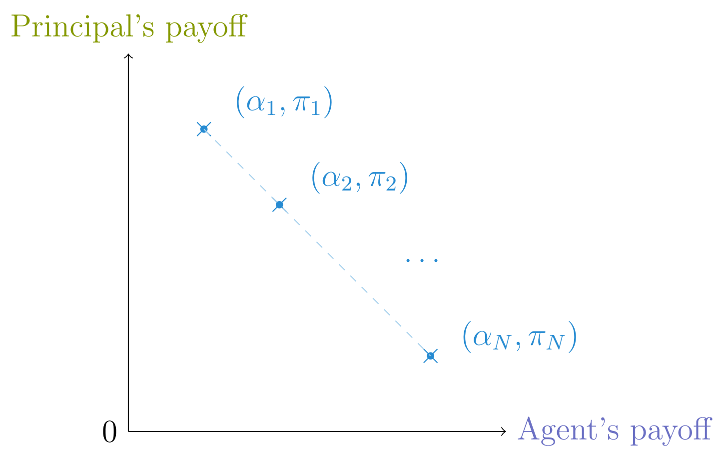

Figure 4: The project space with $N > 2$ projects under Assumptions 1 and 2. We see the conflicting preferences of the principal and agent, and the linear payoffs of the projects.

Assumption 3. *(Nested types) The probability distribution over the agent's types* μ *is such that for any* $S, S' \in \mathcal{S}$ *with* $\mu(S), \mu(S') > 0$*, either* $S \subseteq S'$ *or* $S' \subseteq S$*.*

Assumption 3 requires the set of possible types to be nested in a way that a type of the agent is either a subset or a superset of any other type. This assumption provides a structure to possible types and simplifies the incentives. Under Assumption 3, there can be at most one type with n projects for each $n \in \{1, 2, \ldots, N\}$.

When the parameters $\mathcal{N}$ and μ satisfy Assumptions 1, 2, and 3, we refer to this restricted type space with the restricted payoff structure as *nested linear type space*. The nested linear type space reduces the number of incentive compatibility constraints to at most $(N - 1)$, simplifying the problem significantly.

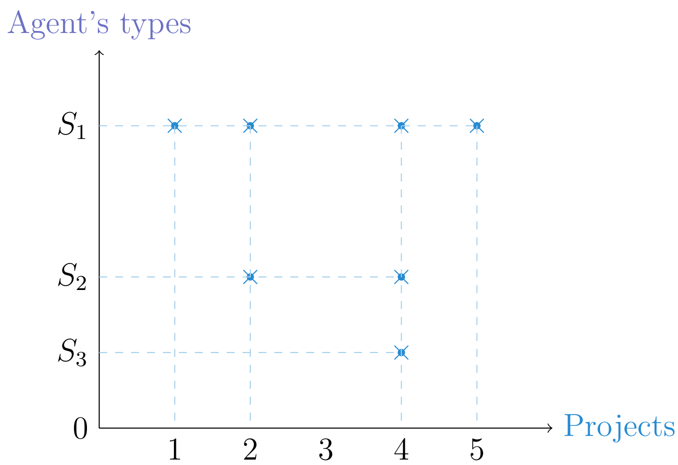

Figure 5: The type space $\mathcal{S}$ with $N > 2$ projects under Assumption 3 where project i refers to (α_i, π_i). The type space is nested such that if we take any two types, one would be a subset of the other.

Under these regularity conditions provided by Assumptions 1, 2, and 3, our main result extends to the general model, and the commitment payoff is always attainable in an equilibrium of the sequential delegation game.

Theorem 3. *In the nested linear type space, there exists an equilibrium of the sequential delegation game that attains the principal's commitment payoff as $\delta \to 1$.*

The main idea behind the proof is that we can divide solving for the principal-optimal mechanism into two parts. We first establish that in any optimal mechanism, each type's each possible report generates the same expected payoff v for the agent. Then, we can solve the optimization problem for a fixed value of v for each type. Combined with the fact that the differences between payoffs are linear, the optimal mechanism takes a very clean separating structure. The optimal mechanism can then be replicated in equilibrium with similar strategies as in the separating equilibrium in the two project case.

We can show with an example that our result for the nested linear type space is not tight: there are examples of type spaces outside this class where the commitment payoff can be achieved in equilibrium. Recall the example from Section 4.1 where there are three possible projects, $\mathcal{N} = \{1, 2, 3\}$, and three equally likely types in the support of μ with $\mathcal{S} = \{\{1, 2\}, \{2\}, \{2, 3\}\}$.

Note that we are outside the linear nested type space introduced in the previous section as the types are not nested and the payoffs are not linear. This type space can be thought of as augmenting our two project case with a type where the bad project is paired with an even worse project. Recall the optimal mechanism:

- From type $\{1, 2\}$, project 1 is implemented with probability one
- From type $\{2\}$, project 2 is implemented with probability $\frac{3}{8}$.
- From type $\{2, 3\}$, project 2 is implemented with probability one.

The structure of an equilibrium that attains the payoff from this mechanism is similar to the separating equilibrium but it exhibits a novel signaling opportunity. The principal always accepts project 1 and never accepts project 3. If project 3 is proposed at $t = 0$, then project 2 is accepted with certainty at $t = 1$. Otherwise, project 2 is only accepted with a delay at $t^* = \min\{t|\delta^t \leqslant \frac{3}{8}\}$. We should highlight that even though project 3 is never implemented, its proposal acts as a screening device and the agent has an opportunity to signal her type by proposing redundant projects.

5 Conclusion

In this paper, we study a dynamic principal-agent problem where the agent is privately informed about the feasibility of projects, and the interests of the parties are not aligned. Our main focus is on a dynamic delegation game where the informed agent makes proposals over time and the uninformed principal has the authority to approve without the power to commit to his future responses.

We ask whether the principal forfeits all surplus, given that the agent can easily hide principal-preferred projects by never proposing them. Our main result indicates that when there are two possible projects, the principal attains his commitment payoff in the frequent-offer limit. We highlight the principal's lack of proposal power as a major force behind capturing a high payoff despite the informational asymmetry and the lack of commitment power. Delaying the implementation of the principal's dispreferred project, the artful waiting, emerges as a signaling device for the agent and presents a screening opportunity for the principal to attain his commitment payoff.

Our setup has natural applications to organizational economics and merger reviews. We can consider the interaction between the Board of Directors of a firm and its CEO to decide on the direction of the firm. The CEO, who may be motivated by empire-building, is better informed about the feasible directions and the Board depends on the proposals coming from the CEO. Alternatively, we can think of an antitrust authority deciding which mergers to approve to maximize consumer surplus when the profit-maximizing firm is proposing feasible mergers.

Beyond two projects, we find that in a general class of delegation protocols, the sequential delegation game we consider does as well as any other protocol under commitment. We also extend our main result to N projects under regularity conditions. Combining the two results, we point to the advantage of implementing projects in a bottom-up manner instead of issuing top-down commands in the organizational setting or seeking merger proposals from firms instead of issuing restrictions on permissible mergers in the antitrust context.

References

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

Ali, S. N., Kartik, N., & Kleiner, A. (2022). Sequential veto bargaining with incomplete information. *arXiv preprint arXiv:2202.02462*.

Armstrong, M., & Vickers, J. (2010). Competitive non-linear pricing and bundling. *The Review of Economic Studies*, *77*(1), 30–60.

Bester, H., & Krähmer, D. (2017). The optimal allocation of decision and exit rights in organizations. *The RAND Journal of Economics*, *48*(2), 309–334.

Bull, J., & Watson, J. (2007). Hard evidence and mechanism design. *Games and Economic Behavior*, *58*(1), 75–93.

Che, Y.-K., Dessein, W., & Kartik, N. (2013). Pandering to persuade. *American Economic Review*, *103*(1), 47–79.

Deneckere, R., & Severinov, S. (2008). Mechanism design with partial state verifiability. *Games and Economic Behavior*, *64*(2), 487–513.

Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies*, *69*(4), 811–838.

Gerardi, D., Hörner, J., & Maestri, L. (2014). The role of commitment in bilateral trade. *Journal of Economic Theory*, *154*, 578–603.

Glazer, J., & Rubinstein, A. (2001). Debates and decisions: On a rationale of argumentation rules. *Games and Economic Behavior*, *36*(2), 158–173.

Glazer, J., & Rubinstein, A. (2004). On optimal rules of persuasion. *Econometrica*, *72*(6), 1715–1736.

Green, J. R., & Laffont, J.-J. (1986). Partially verifiable information and mechanism design. *The Review of Economic Studies*, *53*(3), 447–456.

Guo, Y., & Shmaya, E. (2022). Regret-minimizing project choice.

Li, W. (2022). Discretion and dynamics in delegated project choice. *Working Paper*.

Nocke, V., & Whinston, M. D. (2010). Dynamic merger review. *Journal of Political Economy*, *118*(6), 1200–1251.

Nocke, V., & Whinston, M. D. (2013). Merger policy with merger choice. *American Economic Review*, *103*(2), 1006–33.

Schneider, J. (2015). Persuasion, pandering, and sequential proposal. *Working Paper*.

Appendices

A Proofs

A.1 Proof of Theorem 1

Proof. 1. Pooling equilibrium and separating equilibrium are perfect Bayesian equilibria of the sequential delegation game.

a) Pooling equilibrium:

In this equilibrium, in any time period and following any history, each type of the agent proposes her favorite available project: the empty type stays silent; the good type proposes the good project; the bad and mixed types propose the bad project. The principal accepts any project at any time period and following any history.

Given the strategy of the agent, independent of the history, when the good project is proposed, the principal believes that it is the good type with probability 1 and expects that the good project will be proposed in the next period and in all subsequent periods if the current proposal is rejected, so he accepts the current proposal. When the bad project is proposed, independent of the history, the principal believes that it is the bad type with probability $\frac{\mu_b}{\mu_b+\mu_m}$ and the mixed type with probability $\frac{\mu_m}{\mu_b+\mu_m}$. He expects the bad project to be proposed in the next period and in all subsequent periods if the current proposal is rejected, so he accepts the current proposal.

Given the principal's beliefs and best responses, it is optimal for each type of the agent to propose her favorite project in any period, independent of the history.

b) Separating equilibrium:

In this equilibrium, strategies of the agent's each type are described as follows:

- The empty type stays silent in any period and following any history.
- The good type proposes the good project in any period and following any history.

- The bad type stays silent at any $t < t^*$ and proposes the bad project at any $t \geqslant t^*$ following any history where $t^* := \min\{t : \alpha_g \geqslant \delta^t \alpha_b\}$. Following any history where she proposes the bad project at some $t < t^*$, she proposes the bad project in any subsequent period.
- The mixed type proposes the good project at $t = 0$. If she stays silent at $t = 0$, she keeps staying silent at any $t < t^*$ and proposes the bad project at any $t \geqslant t^*$. Following any history where she proposes the bad project at some $t < t^*$, she proposes the good project in any subsequent period. If the proposal of the good project is rejected in any period, the mixed type proposes the good project in the next period.

the principal accepts the good project in any period following any history. He rejects the bad project at any $t < t^*$ following any history; he also rejects the bad project in any period following the history where it is proposed at some $t < t^*$; he accepts the bad project at any $t \geqslant t^*$ following the history where the agent stays silent at every $t < t^*$.

Whenever the good project is proposed, the principal forms appropriate beliefs and accepts the proposal following any history. When the bad project is proposed at some $t < t^*$, the principal believes that it is the mixed type with probability 1 and expects the good project to be proposed in the next period, so he rejects the current proposal. In fact, following any history where the bad project is proposed at some $t < t^*$, the principal only accepts the good project in any following period. Only when the bad project is proposed at some $t \geqslant t^*$ following any history where the agent stays silent at every single $t < t^*$, the principal believes that it is the bad type with probability 1 and accepts the proposal.

the agent's empty and good types are clearly best responding by staying silent and proposing the good project, respectively, in every period and following any history. Given the principal's beliefs and the threshold t^*, the bad type is best off by staying silent until the threshold and proposing the bad project only then as it is the only way to get it implemented. On the other hand, the threshold is defined precisely to deter the mixed type from waiting to get the bad project implemented, so she does not find mimicking the bad type profitable at the outset of the game. However, in the event that she stays silent in the first period, her best response is to follow the bad type's on-path strategy. In any history where she proposes the bad project before t^*, given that the principal will only accept the good project from that point on, she always proposes the good project in any period that follows.

2. Pooling equilibrium and separating equilibrium attain the principal's commitment payoff as $\delta \to 1$.

   a) Pooling equilibrium:

   We can express the outcome of the pooling equilibrium in terms of the implemented project and the time of implementation for each type.

   - The good type: the good project at $t = 0$.
   - The bad type: the bad project at $t = 0$.
   - The mixed type: the bad project at $t = 0$.

   Then, the expected payoff of the principal from this equilibrium is

   $$\mu_g \pi_g + \mu_b \pi_b + \mu_m \pi_b$$

   which is exactly equal to the expected payoff of the principal from the optimal mechanism when it takes the pooling form.

   b) Separating equilibrium:

   Again, we can express the outcome of the separating equilibrium in terms of the implemented project and the time of implementation for each type.

   - The good type: the good project at $t = 0$.
   - The bad type: the bad project at $t = t^*$.
   - The mixed type: the good project at $t = 0$.

   Then, the expected payoff of the principal from this equilibrium is

   $$\mu_g \pi_g + \mu_b \delta^{t^*} \pi_b + \mu_m \pi_g$$

   We can observe that, by the definition of t^*, we have $\delta^{t^*} \to \frac{\alpha_g}{\alpha_b}$ as $\delta \to 1$. Then, the expected payoff from the separating equilibrium gets arbitrarily close to the expected payoff of the principal from the optimal mechanism when it takes the separating form as $\delta \to 1$.

$\square$

A.2 Induced Social Choice Function

We provide the details of collapsing the probability of various outcomes involving i into a single probability here.

Fix a protocol, a strategy the principal has committed to, and a best response of the agent. Let (i, t) denote the outcome that project i is implemented (proposed and accepted) at t. No project ever being implemented is also a possible outcome. The proof proceeds in two steps. We first show that the strategy and the best response induce, for any type S, a probability distribution over outcomes. We then condense these probabilities to arrive at the *induced* SCF.

Let $x_S(h_t, i)$ be the probability of the agent proposing i at history h_t, according to the best response, and $y(h_t, i)$ be the probability of the principal accepting i at history h_t, where $i \in S \cup \{\varnothing\}$ and $y(h_t, \varnothing) = 0$ for any $t,\ h_t$. We define the probability of any history inductively. We denote period 0 history by ϕ, so at $t = 1$, for any $h_1 = (\phi, i)$, where $i \in S \cup \{\varnothing\}$, we define $\nu(h_1) = x(\phi, i)(1 - y(\phi, i))$, which is just the probability that i was proposed at $t = 0$ but not accepted, and thus the probability of history h_1 at $t = 1$. This is clearly a number in $[0, 1]$. Given that we have defined $\nu(h_t)\ \forall\ h_{t'},\ t' \leqslant t$, and any $h_{t+1} = (h_t, i)$ for some $i \in S \cup \{\varnothing\}$, we have that $\nu(h_{t+1}) = \nu(h_t)x(h_t, i)(1 - y(h_t, i))$.

We define the probability of outcome (i, t) as

$$p_S(i, t) := \sum_{h_t} \nu(h_t)x(h_t, i)y(h_t, i)$$

for all $i \in S$. It can be verified that the sum of probabilities for all outcomes in which a project *is* implemented,

$$\sum_{t=0}^{\infty} \sum_{i \in S} p_S(i, t) \leqslant 1,$$

where the probability of the outcome that *no* project is ever implemented is

$$1 - \sum_{t=0}^{\infty} \sum_{i \in S} p_S(i, t).$$

Thus, for any type S, the strategy and best response induce a probability distribution over outcomes.

We now construct the corresponding Induced SCF. For any S', probability of implementation of $i \in S'$ is

$$f^I_{S'}(i) = \sum_{t=0}^{\infty} \delta^t p_{S'}(i,t).$$

A.3 Proof of Theorem 2

Proof. We now fix a static IC mechanism. In this mechanism, any report $S = \{i_1, i_2, \ldots, i_m\}$, is mapped to implementation probability q_{Sk} for project i_k, and $\alpha_{i_1} < \alpha_{i_2} \ldots < \alpha_{i_m}$. We first define, for any S,

$$y_S(\delta) := \min\{\frac{1}{q_{S1} + \frac{q_{S2}}{\delta} + \ldots \frac{q_{Sm}}{\delta^{m-1}}}, 1\}$$

and let $y(\delta) = \min y_S(\delta)$.

We now construct the corresponding strategy in the sequential delegation game. According to this strategy:

- At any $t \in \{0, 1, \ldots, N-1\}$, the principal accepts *no* proposal irrespective of history.
- At every history where at any $t \in \{0, 1, \ldots, N-1\}$ the agent proposed anything other than project $t+1$ or $\varnothing$, the principal rejects any proposal.
- Fix a history where the set of projects proposed from $t = 0$ until $t = N-1$ is $S' = \{i_1, i_2, \ldots, i'_m\}$, and each project i was proposed at $t = i-1$. We call this history $h^{S'}_N$. Let $q_{S'1}$, $q_{S'2}$, $\ldots$ $q_{S'm'}$ be the probabilities of implementation of each project in S', when S' is reported in the mechanism we have fixed.
- At h^S_N, at $t = N$, if i_1 is proposed, the principal accepts with probability $y(\delta)q_{S1}$.
- If agent does *not* propose i_1 at h^S_N, the principal rejects any proposal at any $t > N$.
- At the history $(h^S_N, i_1, i_2 \ldots, i_{k-1})$ if the agent proposes i_k, it is accepted with probability $\frac{y(\delta)q_{Sk}}{\delta^{k-1}(q_{S1}+\frac{q_{S2}}{\delta}+\ldots+\frac{q_{S(k-1)}}{\delta^{k-2}})}$. If agent does *not* propose i_k at $(h^S_N, i_1, i_2, \ldots, i_{k-1})$, the principal rejects the current proposal and any proposal at any future period t.

- Period $N+(m'-1)$ onward (given that history until period N is $h_N^{S'}$), no project is accepted, irrespective of history.

Note that since the set of projects proposed in the first N periods is S', it is optimal for the agent to report projects in decreasing order of the principal's preference in the next m' periods. If the agent stays silent at any of the m' periods that follow, or recommends a project *out of turn*, the principal never accepts any project again. Let the expected payoff from reporting S' in the mechanism be $E_{S'}$. It is easy to see that if the agent proposes all projects she has in the first N periods, she gets an expected payoff of $\delta^N y(\delta) E_{S'}$. Thus, since the mechanism was IC, it is indeed a best response for the agent to propose *all* projects she has in the first N periods. The principal's payoff from this strategy and best response is therefore the product of $\delta^N y(\delta)$ and the expected payoff from the mechanism. It can be easily verified that $y(\delta) \to 1$ as $\delta \to 1$. So, the principal's payoff from this strategy approaches the payoff from the mechanism as $\delta \to 1$.

$\square$

We now provide an example that shows that with certain delegation protocols, the ability to commit to a strategy may not be enough to attain the commitment payoff from the optimal static mechanism. Recall the example in Section 4.2 where there are three possible projects, $\mathcal{N} = \{1, 2, 3\}$, and three equally likely types in the support of μ with $\mathcal{S} = \{\{1, 2\}, \{2\}, \{2, 3\}\}$. The payoffs are:

$$\pi_1 = 8, \pi_2 = 3, \pi_3 = 1$$

$$\alpha_1 = 3, \alpha_2 = 8, \alpha_3 = 9$$

The optimal mechanism is as follows:

- From type $\{1, 2\}$, project 1 is implemented with probability one
- From type $\{2\}$, project 2 is implemented with probability $\frac{3}{8}$.
- From type $\{2, 3\}$, project 2 is implemented with probability one.

Consider the delegation protocol where the principal makes proposals at each period. At each time period, he chooses a delegation set, from which the agent can choose a project. If a project is accepted, players get their discounted payoffs and the game ends; otherwise in the next period the principal sets another delegation set.

In this protocol, even if the principal could commit to a strategy, he cannot separate type $\{2, 3\}$, from type $\{2\}$ as he can do in a static mechanism. Since a strategy of the principal merely conditions a history of rejected delegation sets to a delegation set, and the ability to reject projects is not type-dependent, type $\{2\}$ can always imitate type $\{2, 3\}$ to get the allocation probability of 2 in $\{2, 3\}$. This is because for any strategy of the principal, histories along with 2 is implemented with positive probabilities, are precisely histories that involve all previous delegation sets being rejected by the agent. Not having 3 does not hinder $\{2\}$'s ability to replicate these histories. In absence of a mechanism to condition implementation of one project on another project that the agent has, the ability to replicate the optimal mechanism breaks down.

A.4 Proof of Theorem 3

Proof. The proof proceeds in three steps. Broadly, we first show that in solving for the optimal mechanism, the optimization problem can be divided into two parts. Then we show that if the solution to the second part corresponds to a value in a certain set, the payoff from the optimal mechanism can be replicated in equilibrium. Lastly, we show that the solution to the second part must always correspond to a value in this set.

Let the set of types be $\{S_1, S_2 \dots S_M\}$ where for any $i < i'$, we have that $S_{i'} \subset S_i$. For any type S_i, let $\mu(S_i) = \mu_i$ and let $q_{i,j}$ be the probability of implementation of $j \in S_i$ in the mechanism when the report is S_i. Let the expected value to the agent, corresponding to any report S_i, be denoted by E_i, where $E_i := \sum_{j \in S_i} q_{i,j} \alpha_j$. Observe that due to type-dependent message spaces and the support of $\mu(.)$, the IC constraints here boil down to $(M - 1)$ inequalities:

$$E_1 \geqslant E_2, \dots \geqslant E_M,$$

and we refer to the inequalities $E_{i+1} \dots \geqslant E_M$ as the IC constraints *below* i and the inequalities $E_1 \geqslant E_2 \dots \geqslant E_i$ as the IC constraints *above* i.

Lemma 1. *In any optimal mechanism, any report must generate the same expected payoff for the agent. Formally, for any two reports S_i and $S_{i'}$, it must be that $E_i = E_{i'}$.*

Proof. We prove this by contradiction. Suppose, in an optimal mechanism there are i, i' such that $E_i \neq E_{i'}$. Without loss, let $i < i'$. This implies, given the nature of the support of $\mu(.)$, that $S_{i'} \subset S_i$. Since the optimal mechanism is IC, it must be that $E_i \geqslant E_{i'}$, and since $E_i \neq E_{i'}$, we have that $E_i > E_{i'}$. This in turn implies that we can find *consecutive types* $k, k+1$ such that $i \leqslant k < k+1 \leqslant i'$ and $E_k > E_{k+1}$. So without loss, let $i' = i+1$. Let i^* be the lowest indexed project in S_i.

- **Case 1**: *Corresponding to report* S_i, $q_{i,j} > 0$, *for some* $j > i^*$.

  In this case, consider the following perturbation: Let $q'_{i,j} = q_{i,j} - \varepsilon$ and $q'_{i,i^*} = q_{i,i^*} + \varepsilon$. Now, $E'_i = E_i - \varepsilon(\alpha_j - \alpha_i) < E_i$. The ε in the perturbation is small enough that $E'_i > E_{i+1}$. Other than this change, all allocation probabilities corresponding to all other reports are unchanged, relative to the original mechanism. This new mechanism is IC, because since the original mechanism was IC, and we have reduced E_i, all IC constrains *above* i still hold. The inequality between E'_i and E_{i+1} is preserved, so this IC still holds. All IC constraints *below* i still hold, clearly. Thus we have constructed another IC mechanism in $(*)$ that gives a strictly higher expected payoff to the principal, as his payoff from type S_i increases by $\varepsilon(\pi_{i^*} - \pi_j)$. Thus, the mechanism we started out with cannot be optimal.

- **Case 2**: *Corresponding to report* S_i, $q_{i,j} = 0$, *for every* $j > i^*$.

  We can again construct an IC mechanism in $(*)$ that gives strictly higher expected payoff to the principal. Since $q_{i,j} = 0$, for every $j > i^*$, we have that $E_i = q_{i,i^*}\alpha_{i^*} \leqslant \alpha_{i^*}$, as only i^* might have positive allocation probability in S_i. Also, $E_i > E_{i+1}$, so it must be that $\sum_j q_{i+1,j} < 1$. This is because all projects in S_{i+1} have weakly higher payoff for the agent than α_{i^*}, so if their allocation probabilities sum up to one, we would have $E_{i+1} \geqslant \alpha_{i^*} \geqslant E_i$, which cannot be. So, since $\sum_j q_{i+1,j} < 1$, in particular, $q_{i+1,(i+1)^*} < 1$. Consider the following perturbation: let $q'_{i+1,(i+1)^*} = q_{i+1,(i+1)^*} + \varepsilon$ where ε is small enough that $E'_{i+1} = E_{i+1} + \varepsilon\alpha_{(i+1)^*} < E_i$, so the IC constraint between $i, i+1$ is preserved. Everything else is unchanged with respect to the original mechanism. Clearly, all IC constraints *above* i hold, and all *below* i hold as well as we increased E_{i+1}. This mechanism gives the principal a higher expected payoff since the payoff from type S_{i+1} has increased. So, the mechanism we started out with cannot be optimal.

This completes the proof of our claim that in any optimal mechanism, *any* report must generate the same expected payoff for the agent.

□

Now that we have shown this, the principal's optimization problem (finding the payoff-maximizing mechanism among all IC mechanisms in ($*$)) can be divided into two parts. First, for any expected value v, find the optimal mechanism corresponding to *this* v; the mechanism that maximizes the principal's payoff when each report generates an expected payoff of v for the agent. Then, maximize the principal's payoff over the possible values of v, i.e. find the values of v the optimal mechanism corresponding to which generates the highest expected payoff for the principal. Our aim is not to solve for the optimal mechanism, but rather show it is always the case that the principal's expected payoff from the optimal mechanism can be attained in equilibrium. We do this in the steps that follow.

Lemma 2. *In any optimal mechanism, it cannot be that $v > \alpha_{M*}$, where M^* is the lowest indexed project in S_M.*

Proof. Suppose in the optimal mechanism, $v > \alpha_{M*}$. Then it must be that there is a project $j \in S_M$ such that $j > M^*$, since the expected value that type S_M gets is $> \alpha_{M*}$. So, we can perturb this mechanism as follows:$q'_{M,j} = q_{M,j} - \varepsilon$, and $q'_{M,M*} = q_{M,M*} + \varepsilon$. Clearly, the new mechanism is IC as there is no IC below M. And it results in higher expected payoff for the principal. □

Lemma 3. *Let the set of all projects in types $\{S_1, S_2 \dots S_M\}$ be $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$, where $\alpha_1 < \alpha_2 \dots \alpha_N$. For any $v \in \{\alpha_1, \alpha_2, \dots, \alpha_N\}$, such that $v \leqslant \alpha_{M*}$, there exists an equilibrium of the sequential delegation game in which the principal attains the payoff from the optimal mechanism corresponding to v, as $\delta \to 1$.*

Proof. Recall that for any project i, $\frac{\pi_1 - \pi_i}{\alpha_i - \alpha_1} = K$ where K is a constant. Let $v = \alpha_k < \alpha_{M*}$. For each S_i, we solve the following problem:

$$\begin{aligned} \max_{\{q_{i,j} | j \in S_i\}} \quad & \sum_{j \in S_i} q_{i,j} \pi_j \\ \text{subject to} \quad & \sum_{j \in S_i} q_{i,j} \alpha_j = \alpha_k \end{aligned} \tag{1}$$

There are two possibilities: Either $i^* < k$ or $i^* \geqslant k$. If $i \geqslant k$, the solution is $q^*_{i,i^*} = \frac{\alpha_k}{\alpha_i}$, because we cannot do better than assigning positive probability to *only* the principal-favorite project in S_i, which is i^*, and since $i^* \geqslant k \implies \alpha_i \geqslant \alpha_k$, we can do so.

Now, let us consider the case where $i < k$. Here, $\alpha_{i^*} < \alpha_k$, so we can no longer assign positive probability only to i^* in S_i. In this case, any solution to the above optimization problem must satisfy $\sum_j \{q^*_{i,j} | j \in S_i\} = 1$. If $\sum\{q^*_{i,j} | j \in S_i\} < 1$, then in particular $q^*_{i,i^*} < 1$, and $q^*_{i,j} > 0$ for *some* $j' > i^*$, because we must have $\sum_{j \in S_i} q^*_{i,j}\alpha_j = \alpha_k$. Fix any such $j' > i^*$, such that $q^*_{i,j'} > 0$. We can now perturb the allocation probabilities as follows: Let $q^{**}_{i,j'} = q^*_{i,j'} - \varepsilon$, $q^{**}_{i,i^*} = q^*_{i,i^*} + \varepsilon\frac{\alpha_j}{\alpha^*_i}$, and $q^{**}_{i,j''} = q^*_{i,j''}\ \forall\ j'' \neq \{i^*, j\}$. It is straightforward to check that $\sum_{j \in S_i} q^{**}_{i,j}\alpha_j = v$, the principal's expected payoff is strictly higher, and for ε small enough, $\sum_{j \in S_i} q^{**}_{i,j}\alpha_j \leqslant 1$. So, if $i < k$, we must have $\sum\{q^*_{i,j} | j \in S_i\} = 1$. The constraint in the optimization problem can be thus be rewritten substituting $q_{i,i^{**}} = 1 - \sum_{\{j \in S_i | j < i^{**}\}} q_{i,j}$, where i^{**} is the highest indexed project in S_i.

$$\sum_{\{j \in S_i | j < i^{**}\}} q_{i,j}(\alpha_{i^{**}} - \alpha_j) = \alpha_{i^{**}} - \alpha_k \tag{2}$$

We can also, after the same substitutions, rewrite the objective function and get:

$$\pi_{i^{**}} + \sum_{\{j \in S_i | j < i^{**}\}} q_{i,j}(\pi_j - \pi_{i^{**}}),$$

which, after substituting (2), is just equal to

$$\pi_{i^{**}} + \sum_{\{j \in S_i | j < i^{**}\}} q_{i,j}K(\alpha_{i^{**}} - \alpha_j) = \pi_{i^{**}} + (\alpha_{i^{**}} - \alpha_k)K$$

Observe that the last expression is a constant independent of allocation probabilities. So, *any* allocation probabilities that satisfy $\sum\{q_{i,j} | j \in S_i\} = 1$ and $\sum_{j \in S_i} q_{i,j}\alpha_j = \alpha_k$, solves the optimization problem. In particular, $q_{i,k} = 1$ solves (1) when $i < k$.

To sum up, for any $v \in \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$, *an* optimal mechanism *corresponding to*

v is as follows:

$$q^*_{i,i*} = \frac{\alpha_k}{\alpha^*_i}, q^*_{i,j} = 0 \ \forall \ j > i^*, \ if \ i^* > k,$$

and

$$q^*_{i,k} = 1, q^*_{i,j} = 0 \ \forall \ j \neq k, \ if \ i^* \leqslant k$$

We now construct an equilibrium that replicates the payoff from the above mechanism. The construction is very similar to our *separating* equilibrium from the two-project case. Fix $v = \alpha_k$. Consider the equilibrium where on path, at $t = 0$, all types that have project k report it, and this proposal is accepted right away. For every S_i such that $i^* > k$, there exists a threshold $t^*_{i*}(\delta)$, such that type S_i, which does not have project k, proposes i at t^*_i, which is then accepted by the principal. We define this threshold inductively:

$$t^*_{k+1}(\delta) := \min\{t : \alpha_k \geqslant \delta^t \alpha_{k+1}\},$$

and, given that we have defined t^*_{k+j}, we define t^*_{k+j+1} as follows:

$$t^*_{k+j+1}(\delta) := \min\{t : \delta^{t^*_{k+j}(\delta)} \alpha_{k+j} \geqslant \delta^t \alpha_{k+j+1}\}$$

We omit the details of the strategies, as they are very similar to the *separating* equilibrium. But intuitively, this on path behavior can be supported in equilibrium as if the principal sees a proposal $i^* > k$ before t^*_{i*}, his off path belief is that it is type S_1 with probability one, and if this proposal is rejected, S_1 will propose 1 in the next period. The thresholds are such that any type that does not have k will find it optimal to propose the principal's favorite project that it has, at the appropriate threshold.

Note that in this proof we have implicitly assumed that all types where $i^* < k$ have project k. In case they do not, this construction does not work. However, the next Lemma will show that we do not have to worry about these cases; if such an $\alpha_k = v$ in the optimal mechanism, we can find another v' that attains same or strictly higher payoff for the principal, such that the optimal mechanism corresponding to this v' is implementable in equilibrium.

□

We have thus shown that every optimal mechanism must have *some* v which is the expected payoff to each type of the agent, and if the optimal mechanism has $v^* \in \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$, there always exists an equilibrium where the principal attains the payoff from the optimal mechanism as $\delta \to 1$. We now show, in the next lemma, that there is always a $v \in \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$ such that the optimal mechanism *corresponding to* v is indeed optimal. This would complete our proof that commitment payoff can be attained for this case of nested types.

Lemma 4. *For any v such that $v \in (\alpha_k, \alpha_{k+1})$ for some $k \in \{1, 2 \ldots N-1\}$, either v cannot be part of the optimal mechanism, or there exists $v' \in \{\alpha_1, \alpha_2 \ldots \alpha_N\}$ such that the principal's payoff from the optimal mechanism corresponding to v' is the same as his payoff from the optimal mechanism corresponding to v.*[7]

Proof. Let $v \in (\alpha_k, \alpha_{k+1})$ for some $k \in \{1, \ldots, N\}$. The principal's objective is to maximize his expected payoff by choosing implementation probabilities for each type:

$$\max \sum_{j \geqslant i*} q_{i,j}\pi_j \quad \text{subject to} \quad \sum_{j \geqslant i*} q_{i,j}\alpha_j = v, \quad \forall i \in \{1, \ldots, M\}$$

For all types S_i with $i^* \geqslant k+1$, the optimal mechanism assigns $q_{i,i*} = \frac{v}{\alpha_{i*}} < 1$ and $q_{i,j} = 0$ for all other $j > i^*$.

For the rest of the types S_i with $i \leqslant k+1$, as we argued in Lemma 3, we have $\sum_j q_{i,j} = 1$. Since given the constraint that the agent's expected payoff equals v, any randomization is optimal, we consider one particular randomization as part of the optimal mechanism.

An optimal mechanism that corresponds to the expected value $v \in (\alpha_k, \alpha_{k+1})$ is as follows:

- Consider $i = \min\{i' | i'^* < k+1\}$. In this case, $(i+1)^*$ is the lowest indexed project of the type S_{i+1}, so it must be that $(i+1)^* \geqslant k+1$. Also all types above $(i+1)$ will have both i^* and i^*+1.

- for all types above i, only projects i^* and i^*+1 are implemented with positive probability, with the appropriate mixture to provide the expected payoff of v, Let these probabilities be q_{i*} and $q_{(i+1)*}$.

- for all types below i, only the project with the lowest index is implemented with positive probability.

[7] The perturbations that we construct here will also work if $v = \alpha_k$ sor some k butt all types where $i^* < k$ do not have k.

We now argue that there exists some v' such that there is an IC mechanism in which each type of the agent gets v' and the principal gets a strictly higher payoff than the mechanism we describe above. Consider the following perturbation:

- for all types such that lowest indexed project is $< k+1$, implement $(i+1)^*$ with probability $q_{(i+1)^*} - \varepsilon$ and i^* with probability $q_{i^*} + \varepsilon$;
- for all types such that lowest indexed project is $\geqslant k+1$, implement this lowest indexed project i' with probability $q_{i,i'} - \frac{\varepsilon(\alpha_{(i+1)^*} - \alpha_{i^*})}{\alpha_{i'}}$.

Now the gain for the principal is

$$\sum_{i' \leqslant i} \mu_{i'} \varepsilon(\pi_{i^*} - \pi_{(i+1)^*})$$

and the loss is

$$\sum_{i' > i} \mu_i \frac{\varepsilon(\alpha_{(i+1)^*} - \alpha_{i^*})}{\alpha_{i'}} \pi_{i'}.$$

We can see that ϵ gets canceled out, and the comparison only depends on the parameters, probabilities of the types and the payoffs.

Either the gain is greater or the loss, or they are exactly equal. If the gain is greater than the loss, then the perturbed mechanism is an IC mechanism where the principal is strictly better off, and the original mechanism cannot be optimal. If the loss is greater than the gain, then we can reverse the signs of the perturbation and achieve an IC mechanism where the principal is strictly better off again, making the previous mechanism not optimal. Finally, if the gain and the loss are exactly the same, then any perturbation would result in the same expected payoff for the principal. In this case, we can perturb the mechanism such that the expected payoff for all types is α_{k+1} and this would be an optimal mechanism as well. In addition, this optimal mechanism can be implemented in an equilibrium of the game as established in Lemma (3).

□

□