The Art of Waiting*

Vasundhara Mallick† Ece Teoman‡

July 17, 2023

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 a proposed project. Importantly, the principal cannot commit to his responses, and cannot implement a project unless it is proposed. 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 can credibly commit to rejecting 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. In particular, these results suggest that to curb a manager's *empire building* plans, eliciting proposals from her "bottom-up" might be better 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 a CEO (principal) of a firm and a manager (agent). The manager may be better informed about what actions the firm can undertake, and unlike the CEO, is motivated by empire building. In such cases, the CEO cannot blithely assume that the manager selects actions purely for shareholder interests. Another example is that of an antitrust authority deciding which mergers to approve: It only wants to approve 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 no proposed project is *ever* accepted, so 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 *proposal power*: The principal is effectively giving up bargaining or proposal 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, along with restricting the principal's action to accepting or rejecting a proposed project, 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, and rejects the proposal. 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. It is also crucial that the principal cannot implement projects unless the agent proposes them. Otherwise, at a history where the principal believes that only his dispreferred projects are available, he would implement them himself.

We view this finding to be of more than just theoretical interest. It suggests a gain to organizations from allowing the agents—be 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). Our work also suggests that antitrust authorities, venture capital boards, and grant funding agencies may gain from allowing the agent to be the one to propose projects flexibly rather than constraining the agent to propose only certain projects.

We also explore what may happen when the principal and the agent interact in ways other than the agent always proposing and the principal accepting or rejecting these proposals. We study a general class of *delegation protocols* and compare these under commitment; i.e. if the principal could commit to a strategy within a protocol. We show that the delegation protocol we study 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 studies a delegation problem in a project selection setup. This problem is studied by Armstrong & Vickers (2010) in a static setup. In their work, the formal authority to choose the project lies with the agent but the principal can restrict the set of projects the agent can choose from. This is equivalent to the principal being able to commit to a deterministic mechanism. In contrast, our game is dynamic, with the agent making proposals, and we use the optimal *stochastic* mechanism as the commitment benchmark. We show that this commitment payoff can be attained in an equilibrium of our game.[1]

[1] Nocke & Whinston (2013) study a similar problem in a static setup, in the context of mergers. An antitrust authority can commit ex-ante to its merger-approval rule. However, this is not a direct static counterpart of our setup. There are multiple firms (agents) here, and given the set of permitted and feasible mergers, the implemented merger is the result of a bargaining process among firms.

In Aghion & Tirole (1997), the uncertainty is about the payoffs from projects, rather than their feasibility. Both *formal* authority to make decisions, and *real* authority, where the agent's proposals are accepted by the principal, can incentivize the agent to exert effort to learn the payoffs. Although our setup is quite different, the agent having *real* authority in their paper is similar to the agent having the control over proposals in our model. In our work, the principal can only implement what the agent proposes, so even though the he has the ultimate authority to make decisions, agent has significant *real* control over the decision.

Our setup is one that involves *hard evidence*, since the the agent's private information is about the *feasibility* of decisions, and she can only propose available projects. This is an important way in which our paper differs from some of the broader literature on delegation starting with Holmström (1984), and more recently, Alsonso & Matouschek (2008). They also study a joint decision problem where the principal has the formal authority to take decisions but the agent possesses private information relevant to decision-making. In these models, if the principal can commit to a decision rule as a function of the agent's report, then the formal allocation of authority is irrelevant. This is because any type can imitate the report of another. So, the optimal mechanism can be implemented by constrained delegation, where the principal delegates decision-making to the agent, but restricts her to choose from a *delegation set*.

In contrast to this, in our setup, even if the principal can commit, the optimal static mechanism wouldn't, in general, be equivalent to choosing a delegation set. This is because in our commitment benchmark, since the agent can only include available projects in her report, this helps convey her private information more effectively, rather than if she simply chooses an available project from a delegation set. So the principal making decisions as a function of the agent's report results in more effective communication of agent's private information.

An alternative approach in the project selection literature is to model the interaction as a cheap talk game where the principal lacks commitment power, 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. This is the paper that closest to ours. The key difference is that our work establishes a commitment benchmark and for the case of two projects (and for N projects, under some conditions), characterises the equilibria that achieves the commitment payoff.

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.

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 from implementing the project, 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 have conflicting preferences over the projects: $\pi_g > \pi_b > 0$ and $\alpha_b > \alpha_g > 0$.

The challenge is that not every *possible* project may be *available* to implement, and only the agent knows which projects available: her *type* represents the set of available projects, and is her private information. The agent has four possible types:

- $E = \varnothing$, the *empty* type with no available projects;

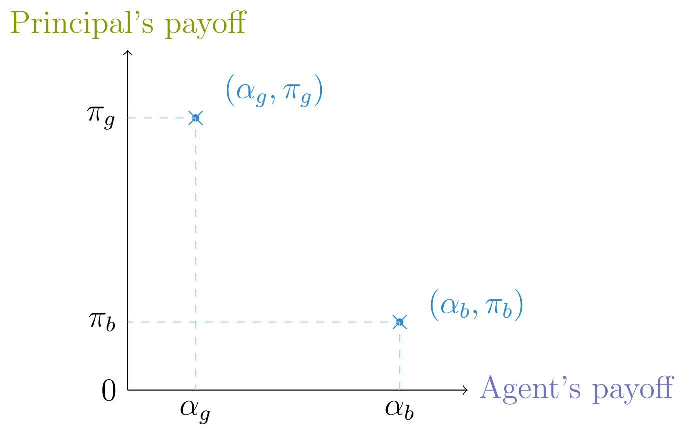

Figure 1: The project space with the principal-preferred good project (α_g, π_g) and the agent-preferred bad project (α_b, π_b).

- $G = \{g\}$, 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.

The set of all possible types is denoted by $\mathcal{S} \equiv 2^{\mathcal{N}} = \{E, G, B, M\}$.[2] An element of $\mathcal{S}$, or a possible type, is denoted by S. The agent's type is drawn from $\mathcal{S}$ and μ_S denotes the probability of type S.

We now describe how the principal and the agent solve the joint decision problem. We refer to their interaction as the *sequential delegation game*, which proceeds as follows. Time is discrete and the principal and the agent have a common discount factor $\delta \in (0, 1)$. In each period $t = 0, 1, 2, \ldots$, the agent makes a proposal and the principal responds to this proposal.

The set of actions, or proposals available to the agent, is type-dependent. If the agent is of type S, then at any time period t, the agent can make proposals from the set $A_S = \{\{i\} | i \in S\} \cup \varnothing$. This means that in any t, the agent of type S can either propose exactly one available project $i \in S$, or stay silent.[3] For the principal, if at any t, project $i \in \mathcal{N}$ is proposed by the agent, he can either *accept* i, or *reject* it. If the agent was silent and did not propose anything, then the only possible action for the principal is to *reject*.

[2] This is assuming that μ has full support on $2^{\mathcal{N}}$; this is not essential for our results.
[3] For type E, $A_E = \varnothing$, so staying silent is the only option.

If the principal rejects a proposed project i, or if the agent is silent, both players obtain 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 $\delta^t \alpha_i$. We focus on the case of $\delta \to 1$, which we interpret it as the frequent-offer limit of the game.

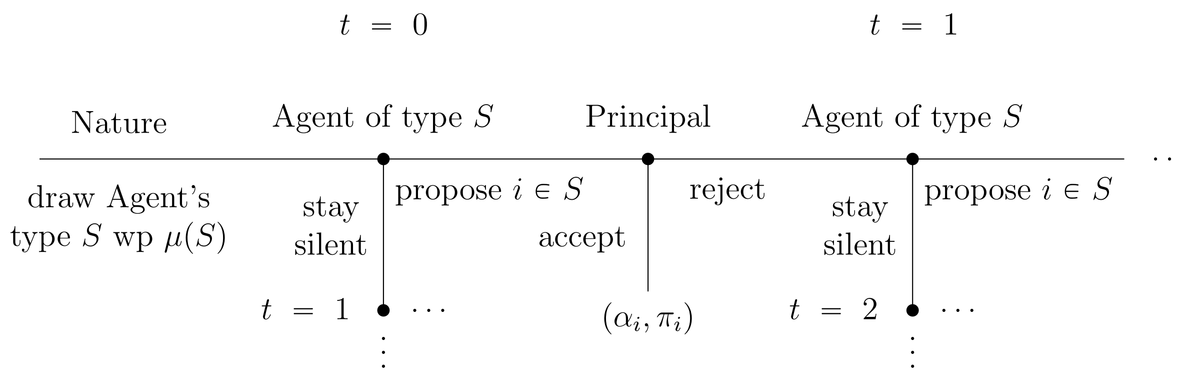

Figure 2: Timeline of the sequential delegation game.

Thus, in the sequential delegation game, for any time period t, the set of all possible histories at the beginning of period t is $\mathcal{H}_t = (\mathcal{N} \cup \varnothing)^t$. This captures the fact that if we are at t, for any $t' \leqslant t-1$, we can have two cases: (i) a project $i \in \mathcal{N}$ was proposed and rejected, or (ii) Agent was silent, and nothing was proposed, so there is nothing for the principal to accept. This is denoted by $\varnothing$. An element of $\mathcal{H}_t$ is denoted by h_t. If the agent is of type S, her strategy maps any history to a probability distribution over $\{\{i\}|i \in S\} \cup \varnothing$. A principal's strategy maps any history, and a proposal of project i at that history to a probability of accepting i. If agent is silent, then principal's only possible action is to reject. 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 The Benefits of Giving Up Control

This section presents our main result. We study how *well* the principal can do in an equilibrium of our sequential delegation game, despite being uninformed and lacking proposal power. We first establish a commitment benchmark that acts as an upper bound for what the principal could achieve if he could commit to a strategy in the sequential delegation game. Then, our main result shows that this commitment payoff is, in fact, attainable in an equilibrium of the sequential delegation game, even though the principal has no commitment power. Finally, we highlight the importance of *giving up control* over making proposals; in a game where the principal makes proposals, it may not be possible to attain the commitment payoff.

3.1 Commitment Benchmark

We first define a class of static, stochastic mechanisms with type-dependent message spaces, that we refer to 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 can commit to a strategy in the sequential delegation game. For now, we will take this fact as given.[4] In a mechanism in this class, the message space is type dependent, and 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 only subsets of her available projects as a message in the mechanism.[5] A mechanism is a tuple (M, q), where $M = \bigcup_{S \in \mathcal{S}} M(S)$ is the set of all possible messages, and $q : \mathcal{S} \to \Delta(S \cup \varnothing)$ is the outcome function. Therefore, when S is reported, only projects *in* S can be implemented, or no project at all, as captured by $\varnothing$. If no project is implemented, the players obtain the status quo payoff, which is zero for both the principal and the agent. For any type $S \in \mathcal{S}$ and any project $i \in S$, q_{Si} represents the probability of implementing project i when the type S is reported.[6] We define a mechanism to be *incentive compatible* (IC) if no type finds it optimal to report a strict subset.

Every mechanism determines an allocation, which is a vector $\{q_{Si}\}_{S \in \mathcal{S}, i \in S}$. A *feasible* allocation is one where for each type S, we have: $q_{Si} \in [0, 1]$ for each $i \in S$ and $\sum_{i \in S} q_{Si} \leqslant 1$. The principal-optimal mechanism maximizes the principal's payoff by choosing implementation probabilities for projects in each type, subject to feasibility and incentive compatibility constraints.

[4] It is in fact a *tight* upper bound in the sense that for any mechanism, there exists a strategy in our sequential delegation game such that commitment to this strategy gives the principal the same expected payoff as the mechanism as $\delta \to 1$. So, this upper bound is equivalent to commitment within the sequential delegation game. We show this in Section 4.1. There, we also show that this class of mechanisms actually acts as an upper bound for a very general class of interactions between the principal and the agent, not just the one where the agent makes all the offers.

[5] This includes the empty set.

[6] So, $1 - \sum_{i \in S} q_{Si}$ is the probability of not implementing any project when S is reported.

$$\max_{q_{Gg},q_{Bb},q_{Mg},q_{Mb}\in[0,1]} \quad \mu_G q_{Gg}\pi_g + \mu_B q_{Bb}\pi_b + \mu_M q_{Mg}\pi_g + \mu_M q_{Mb}\pi_b$$

$$\text{subject to} \qquad \textstyle\sum_{i\in S} q_{Si}\alpha_i \geqslant \sum_{i'\in S'} q_{S'i'}\alpha_{i'} \qquad (\mathrm{IC}_{SS'})$$

$$q_{Mg} + q_{Mb} \leqslant 1$$

where $\mathrm{IC}_{SS'}$ denotes the IC constraint for type S to not report type $S' \subseteq S$. The second constraint is just the feasibility constraint for the implementation probabilities when the mixed type is reported. Before we solve for the principal-optimal mechanism, we make a few simplifying observations about some properties that must be true for an optimal mechanism.

> **Observation.** No type finds it profitable to report the empty type, so IC_{GE}, IC_{BE}, and IC_{ME} are all redundant.

Recall that corresponding to any report, the mechanism only implements project *in* the report. So, the payoff from reporting the empty type is zero; it has no project, so no project is implemented when it is reported. 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, than she does by 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 an optimal mechanism, we must have $q^*_{Mg} + q^*_{Mb} = 1$.

When the mixed type is reported, the probabilities of implementing the good and the bad projects must sum up to 1 in an optimal mechanism. Suppose not, i.e. $q^*_{Mg} + q^*_{Mb} < 1$. Then, we can increase both q^*_{Mg} or q^*_{Mb} slightly, and have new implementation probabilities $(q^{**}_{Mg}, q^{**}_{Mb}) > (q^*_{Mg}, q^*_{Mb})$ and $q^{**}_{Mg} + q^{**}_{Mb} < 1$. It must be the case that the IC constraints involving the mixed type still hold, as

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

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

From these new implementation probabilities, the principal obtains a strictly higher payoff. So, $q^*_{Mg} + q^*_{Mb} < 1$ cannot be part of an 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, from reporting truthfully, will be at least α_g. On the other hand, her payoff from reporting the good type is at most α_g as $q_{Gg} \in [0, 1]$. Thus, the mixed type is always weakly better off by reporting truthfully than by pretending to be the good type, and IC_{MG}, is redundant.

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

Since IC_{MG} is redundant, and there is no type other than the nixed type that can report the good type, therefore when the good type is reported, an optimal mechanism must implement the good project with certainty.

Given the above *observations*, the problem of finding the optimal mechanism reduces to that of choosing q_{Mg} and q_{Bb} to maximise the principal's expected payoff, subject to IC_{MB}. This is because the other IC constraints are redundant, $q_{Gg} = 1$, and $q_{Mg} + q_{Mb} = 1$, so choosing q_{Mg} and q_{Bb} pins down the optimal mechanism.

The lone IC constraint, IC_{MB}, represents the trade off that the principal faces in implementing the good project with positive probability from the mixed type. If $q_{Mg} > 0$, the principal will have to set $q_{Bb} < 1$, so that the mixed type doesn't imitate the bad type. We now define two mechanisms, and it turns out that one of them is always *an* optimal mechanism.

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

It is easy to see that both these mechanism are IC. In the *pooling* mechanism, the outcome when the type is mixed, is same as the outcome when the type is bad. In both cases, the bad project is implemented with probability one. So, mixed type is *pooled* with the bad. On the other hand, in the *separating* mechanism, the outcome when the type is mixed is different from the outcome when the type is bad. In one case, the good project is implemented with probability one, and in the other, the bad project is implemented with probability $\frac{\alpha_g}{\alpha_b}$. So, this mechanism *separates* the mixed and bad types.

Before stating our result about the optimal mechanism, we establish some notation. Let $\lambda = \frac{\mu_M}{\mu_B}$ be the likelihood ratio of mixed type compared to bad type, and $\lambda^* = \frac{(1-\frac{\alpha_g}{\alpha_b})\pi_b}{(\pi_g-\pi_b)}$.

Proposition 1. *Either the pooling or the separating mechanism is always optimal.*

a) When $\lambda < \lambda^*$, the pooling mechanism is optimal.

b) When $\lambda > \lambda^*$, the separating mechanism is optimal

c) When $\lambda = \lambda^*$, any mechanism with $q_{Gg} = 1$, $q_{Mg}+q_{Mb} = 1$, and $q_{Mg}\alpha_g+q_{Mb}\alpha_b = q_{bB}\alpha_b$ is optimal. In particular, both the pooling and separating mechanisms are optimal.

The proof of the above result is in the appendix. To see the intuition behind the result, recall that we only need to worry about IC$_{MB}$. We can show that this reduces the problem to: (i) whether the principal wants to *separate* the mixed and bad types, as in the separating mechanism, or (ii) *pool* them, as in the pooling mechanism. Consider the case where $\lambda = \frac{\mu_M}{\mu_B} \to \infty$. Then it is *as if*, between the mixed and the bad types, only the mixed type exists. In this case, it would be optimal to set $q_{Mg} = 1$, and therefore, $q_{Bb} = 0$ bu IC$_{MB}$. On the other hand, if $\frac{\mu_M}{\mu_B} = 0$, i.e only the bad type exists, then it is optimal to set $q_{Bb} = 1$. So it makes sense that if λ is high enough, the separating mechanism would do better than the pooling mechanism. This is precisely λ^*, the threshold in our result.

Now that we have solved for the optimal mechanism, we turn our attention back to the sequential delegation game where the principal cannot commit to his responses. In particular, we are interested in exploring 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 Equilibrium

In the sequential delegation game, the agent's proposals are like *reports* of which projects are available, somewhat like the mechanism. However, unlike the mechanism, the principal cannot commit to his responses to the agent's proposals here. This therefore limits how effectively he can elicit the relevant information from the agent, and we might expect a gap between the commitment payoff and what the principal can achieve in an equilibrium of the game.

In the absence of commitment power, one might think that proposal power could help the principal; he could provide the necessary incentives by effectively restricting the choices of the agent. However, in our game, the principal lacks the ability to make proposals as well! This translates into a reduced level of control over which project is implemented, as he cannot implement something that the agent has *not* proposed. He can only accept or reject a project proposed by the agent. The principal seems to be at every possible disadvantage here.

Our main result however, contradicts the above intuition, and 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 particular, there is an equilibrium where delay emerges as a costly signalling device for the agent and allows for the separation of the mixed and the bad type. This equilibrium attains the optimal commitment payoff when the optimal mechanism is separating. We will also argue that the lack of proposal power in fact helps with this *signalling through delay.* We now state our main result where, for each form that the optimal mechanism takes, we describe an equilibrium that attains the payoff from this optimal mechanism in the frequent offer limit.

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 in 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 $t^*(\delta) := \min\{t : \alpha_g \geqslant \delta^t \alpha_b\}$, at which point she proposes the bad project.*
- *The principal accepts the good project at $t = 0$ and the bad project at $t^*(\delta)$.*

The details of the strategies and beliefs that constitute the pooling and the separating equilibria are in the Appendix. Here, we focus on the more interesting case; that of the separating equilibrium. We first argue that the on path behaviour we described indeed attains the principal's payoff from the separating mechanism. In this equilibrium, on the equilibrium path, the mixed type and the good type both propose the good project g at $t = 0$ and it is accepted. So, since g is implemented, i.e. proposed and accepted without delay, this replicates the implementation probabilities of $q_{Gg} = q_{Mg} = 1$. In the mechanism, the bad project b is implemented with an interior probability of $q_{Bb} = \frac{\alpha_g}{\alpha_b}$, and in the equilibrium is is implemented (proposed and accepted) with a delay, at $t^*(\delta)$. By definition of $t^*(\delta)$, we have that as $\delta \to 1$, $t^*(\delta) \to \frac{\alpha_g}{\alpha_b}$. Thus, as $\delta \to 1$, the principal's payoff is the separating mechanism is attained by the separating equilibrium.

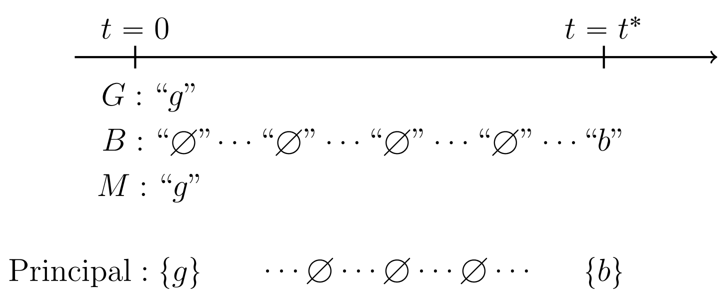

Figure 3: The timing of the proposals and accepted projects on the path of separating equilibrium of the sequential delegation game.

We now informally describe the separating equilibrium, and provide some intuition behind the key forces that hold this equilibrium together. The principal's strategy involves accepting the g whenever proposed, and rejecting b whenever it is proposed before the threshold $t^*(\delta)$. In fact, if b is proposed before $t^*(\delta)$, the principal's strategy is to reject not just this proposal, but *any* future proposal of b as well. More precisely, fix a history h_t at the beginning of period t. Suppose there is a $t^{'} < t$ such that along h_t, b was proposed at $t^{'}$, and $t^{'} < t^*(\delta)$. Then, if b is proposed at period t following h_t, it is rejected. If b is proposed at $t^*(\delta)$, and has never been proposed before this, then it is accepted.

This essentially means that a type that has b *cannot* get the principal to accept it before $t^*(\delta)$. Rather, if such a type proposes b before $t^*(\delta)$, then the principal will never accept b at any future time period. So, the mixed type faces a choice: it can get g accepted right away, at $t = 0$, or wait till $t^*(\delta)$ to get b implemented. By definition of $t^*(\delta)$, it (weakly) prefers to propose g at $t = 0$. The bad type has no option but to wait by staying silent till $t^*(\delta)$. At this point he proposes his only project, and it is accepted.

Thus, delay emerges as a costly signalling device in equilibrium; it is used by the bad type to signal that she indeed only has the bad project. But why does the principal reject b at *any* history that involves b being proposed before $t^*(\delta)$? This is because of his off-path beliefs. At any such history, he believes that it is the mixed type with probability one. Thus, the agent gets punished by the extremal off-path beliefs of the principal, if he ever proposes the bad project before $t^*(\delta)$.

However, ex-ante, it is not clear why this *punishment through beliefs* should be possible. Even if the principal attaches probability one to the mixed type, why does he find it optimal to always reject b given this belief? The agent still has control of the proposals after all, and the principal cannot implement something she doesn't propose. In this case, even if the principal *knows* the agent has both projects, its not obvious that he can make the agent propose g. The agent can just keep proposing b.

The intuition here is that in the complete information counterpart of our game, where it is common knowledge that the agent is of mixed type, there exists an equilibrium where, on path, the agent proposes g at $t = 0$. Consider the following strategies of the principal and the agent: the principal, irrespective of history, rejects b, and accepts g. The agent, irrespective of history, proposes g. In particular, at any history h_t, if the agent proposes g and it is rejected, then at this off-path history $h_{t+1} = (h_t, b)$, the agent's strategy is to propose g. It is easy to see that no party has a profitable one-shot deviation. For the principal, at any time period, if b is proposed, by rejecting it, he expects g to be proposed in the next period, which he would then accept. So, if he is sufficiently patient, it is optimal to reject b. For the agent, at any time period, if she proposes b, it would be rejected, and she would propose g in the next period, which would be accepted. So, her payoff from this deviation is $\delta\alpha_g$. If she doesn't deviate and proposes g in the current period, it is accepted and she gets α_g. Thus, the strategies constitute an SPE.[7]

It is *this* equilibrium that our analysis leverages. Consider a history h_t, where b was proposed and rejected at $t' < t^*(\delta)$ along this history. In the separating equilibrium, if the agent's type is mixed, her strategy is to propose g following such a history. So, if b is proposed at $t < t^*(\delta)$, the principal believes it is the mixed type with probability one, and therefore expects a proposal of g in the next period if he rejects b. This makes rejection of b sequentially rational for the principal at *any* such history, and holds this equilibrium together.

We have therefore argued that off-path beliefs can be used to exploit a complete information equilibrium, and separate the mixed type from the bad. We now argue that there is another force: *lack of proposal power*, that makes this leveraging this complete information equilibrium possible. Consider an alternative to our sequential delegation game, where, in each period, the principal makes an offer in form of a restriction set, which is a subset R of $\mathcal{N}$ that the agent is allowed to choose from. [8] If the agent's type is S, she can either implement a project $i \in R \cap S$, in which case the game ends, or not implement anything and reject R altogether. In this case the game moves to the next period, and the principal offers another restriction set.[9]

The complete information counterpart of this new game, where it is common knowledge that the agent's type is mixed, *also* has an equilibrium where, along the equilibrium path, the good project is implemented at $t = 0$. The strategy of the principal is to set $R = \{g\}$ at any history, and the agent's strategy is to always accept any project that's allowed. It is easy to verify that these strategies constitute an SPE.

However, it turns out that in the new principal-offer game, this complete information equilibrium cannot be exploited! When the separating mechanism is optimal, there is *no* equilibrium of this game that attains the optimal commitment payoff. This is surprising, since we might think that the principal having control of proposals means that he can exert greater influence over what is implemented. But there is a

[7]The *discreteness* of the offer space is important here. Consider the setting from Rubinstein (1982), but with one party making all the offers. Then, in the unique equilibrium, this party captures the entire surplus. However, as Van Damme, Selten, & Winter (1990) shows, any split can be supported if the offer space is discrete. A similar reasoning is at play here.

[8]R can be $\varnothing$, in which case, the agent has no choice and the game proceeds to the next period.

[9]We show in Section 4.1 that the commitment benchmark we established also serves as an upper bound for what the principal can achieve by committing to a strategy in this alternative game.

trade-off between control and sequential rationality here.

To understand the intuition behind this, suppose the principal attempts to replicate the separating mechanism by setting $R = \{g\}$ at all $t < t^*(\delta)$. At $t^*(\delta)$, he sets $R = \{g, b\}$. So the mixed type indeed finds it optimal to accept g at $t = 0$. But this strategy of the principal is not sequentially rational; at $t = 0$, if g is not accepted, the principal *knows* it is the bad type, and therefore would permit b in the very next period. The sequential rationality constraint actually prevents him from waiting till $t^*(\delta)$ before allowing the bad project. Also, since non-acceptance of g is on-path, there is no scope for the principal to form extremal off-path beliefs, and leverage the equilibrium of the complete information game. The limited action set of the agent therefore hurts the principal, and makes *punishment through beliefs* harder to achieve. In fact, Li (2022) establishes that this Coasian force is prevalent in any equilibrium of this alternative game, and the unique equilibrium outcome is the principal setting $R = \{g, b\}$ at $t = 0$.

So, there are various forces that work here and it is their combination that helps with the attainment of commitment payoff in an equilibrium of our game.

4 A General Model with N projects

When there are two possible projects, we showed that the payoff from the static commitment benchmark can always be achieved in an equilibrium of the sequential delegation game. However, when we move beyond two projects, it is not clear if the commitment payoff can be achieved in equilibrium.

There are potentially two reasons why this could happen: (i) there is a wedge between the static stochastic mechanism and what can be achieved by commitment to a strategy 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. We want to differentiate between these two forces. If commitment payoff in the sequential delegation game cannot be achieved, is it because of a limitation of the extensive form, as in (i), or is it because the principal cannot commit to his strategy, as in (ii).

Even with two projects, we might wonder if the principal can do better if the principal and the agent used an alternative extensive form to solve the joint decision problem of implementing a project. We therefore also want to explore the merits and demerits of the sequential delegation game, relative to other extensive forms. We want to understand if our sequential delegation game has any *limitations* that other extensive forms don't. But to answer this question, we need to define precisely what the general space of extensive forms is, and how we are comparing them.

In this section, we consider a general setting where there are $N > 2$ possible projects, and the interaction between the principal and the agent may take one of several possible forms. The set of all *possible* projects is now $\mathcal{N} \equiv \{1, 2, \ldots, N\}$ where project i corresponds to (α_i, π_i). As before, 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 $\mu : \mathcal{S} \to [0, 1]$. We refer to the different extensive forms we consider as *delegation protocols*.[10]

We first define a general class of delegation protocols. We then show that the class of mechanisms that we defined in Section 3.1 serves as a general commitment benchmark in this setting. For this, we prove a Revelation Principle and show that any outcome that can be achieved in any extensive form can be achieved by a direct, static, and stochastic mechanism with type-dependent message spaces. Furthermore, we show that for the sequential delegation game, this upper bound is tight. More precisely, for any static mechanism, there exists a commitment strategy of the principal in the sequential delegation game that attains the same payoff as $\delta \to 1$. So, the sequential delegation game imposes no constraints *beyond* sequential rationality. 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 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 the protocol specifies the proposer, $P(h_t)$ and the set of permissible offers,

[10] We can think of them as capturing the different institutional settings that the principal and the agent might interact in. The sequential delegation game is one possible delegation protocol. Another possible protocol is the principal specifying the set of projects from which the agent is permitted to choose in each period.

$\mathcal{O}(h_t) \subseteq 2^{\mathcal{N}}$, so any offer $O(h_t)$ is a subset of $\mathcal{N}$.[11] When an offer is made by the proposer, the other party 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. [12]

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. Formally, if, at h_t, the proposer is the agent, and the agent's type is S, then we must have $O(h_t) \subseteq S$. On the other hand, if the proposer is the principal, the offer $O(h_t)$ can be any subset of $\mathcal{N}$ that's in $\mathcal{O}(h_t)$, but 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.

4.1.1 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 first define a social choice function and an *induced* social choice function.

Definition. *A social choice function (SCF) is a function f that maps a set $S \subseteq \mathcal{N}$, to*

[11] An offer can be $\varnothing$, this can be interpreted as the party making the offer permitting no project to be chosen by the other, like the agent choosing to stay silent in the sequential delegation game. If the offer is $\varnothing$, then the party responding to the offer has no option but to reject it, as there is no project to accept.

[12] Tying this definition back to our sequential delegation game, it is the protocol where at any h_t, $P(h_t)$ is the agent, and the set of permissible offers is $\mathcal{O}(h_t) = \mathcal{O} = \{\{i\} | i \in \mathcal{N}\}$, i.e. the singleton subsets of $\mathcal{N}$.

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 any best response of the agent to this strategy. [13] 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 $f_S^I(i)$.[14] The induced SCF is then the function f^I that maps any type S to a probability $f_S^I(i)$ of implementation of each $i \in S$.[15]

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 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_S^I is weakly better than the payoff from $f_{S'}^I$, 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. $\square$

This result highlights an important point. In the two-project case, he principal can always attain the commitment payoff in an equilibrium of the sequential delegation game. The above Revelation Principle tells us that this commitment benchmark in fact serves as an upper bound for what the principle can achieve with commitment,

[13]This pair does not need to constitute an equilibrium, we can think of the principal as being able to commit to a strategy.

[14]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 $f_S^I(i) = \frac{1}{2} + \delta\frac{1}{2}$.

[15]The details of collapsing the probability of various outcomes involving i into a single probability can be found in the Appendix A.3.

and therefore in an equilibrium of a very general class of protocols. So, if the principal lacks commitment, and we have two projects, there is no equilibrium of any other protocol in this class that can do better than the principal-optimal equilibrium of the sequential delegation game.

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 a social choice function f. There exists 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 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, $t \in \{0, 1, \dots N-1\}$, 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 a *report* in the mechanism, and since the agent can only propose projects she has, this captures the fact that a type can only report her subsets in the mechanism.

* In time periods $t > N-1$, the projects that were proposed in the first $N-1$ periods 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 imposes no constraints as $\delta \to 1$ beyond sequential rationality. So, if the principal can commit to a strategy, but is constrained by the extensive form of the sequential delegation game, i.e. has to operate in a setting where the agent makes proposals, and he can merely respond to them, then this is not really a constraint. Even if he could choose the extensive form from a very general class of extensive forms, and *then* also commit to his strategy in that extensive form, he cannot do any better.

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, which is the common commitment benchmark for all the protocols we consider.

Consider the delegation protocol where, in each period, the principal makes proposals by choosing a restriction set, which is any subset O of $\mathcal{N}$ that the agent is allowed to choose from. So, at any history h_t , $P(h_t)$ is the principal, and $\mathcal{O}(h_t) = \mathcal{O} = 2^{\mathcal{N}}$. If the agent's type is S, she can either implement a project $i \in O \cap S$, in which case the game ends, or not implement anything and reject O altogether. In this case the game moves to the next period, and the principal offers another restriction set.

Now 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.

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) involves eliciting information from the agent about her type through her proposals and conditioning future responses on these initial proposals. Consider the following strategy of the principal: he rejects any proposal except 1 at $t = 0$. At $t = 1$, he accepts project 2 *only* if project 3 was proposed at $t = 0$, otherwise rejects 2 forever if it is proposed before a certain threshold t^*.[16]

In the alternative delegation 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 $\{2, 3\}$ from $\{2\}$.

4.2 Attaining the Commitment Payoff with N Projects

We now turn our attention back to the sequential delegation game and the possibility of attaining attaining the commitment payoff in equilibrium here. 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.

While the number of possible projects does not alter the game itself or the commitment benchmark, the problem significantly more complex. Even solving for the optimal mechanism is difficult as it now includes IC conditions for each subset of each type. As a result, the problem loses its tractability. In order to recover some of the lost tractability, focus attention on 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.

We show that these assumptions are sufficient conditions for existence of an equilibrium of the game that attains the commitment benchmark. We also show through an example that the conditions are not necessary and highlight another signaling opportunity for the agent by proposing redundant projects.

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

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

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

[16]The construction is similar to the separating equilibrium from the two-project case.

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 be trivial. Beyond two projects, however, there is an array of possibilities for conflicting preferences. Under Assumption 1, the principal and the agent have exactly opposing preferences.

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 $\mathbb{R}^2_{++}$.

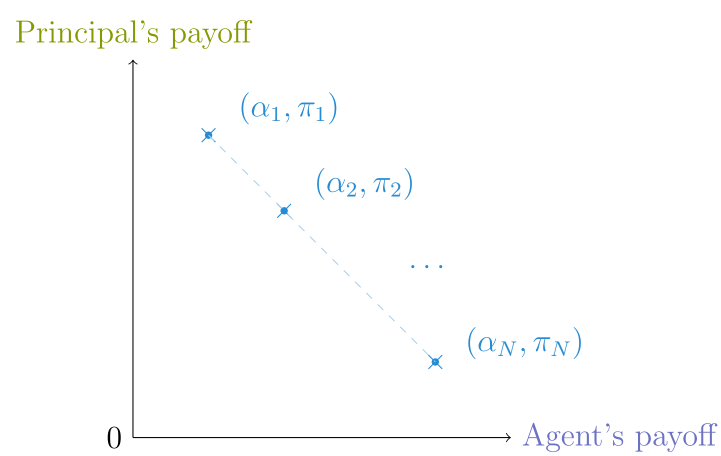

Figure 4: The project space with $N > 2$ projects under Assumptions 1 and 2. We see the conflicting preferences of the principal and agent, and the linear payoffs of the projects.

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 $n \in \{1, 2, \dots, N\}$.

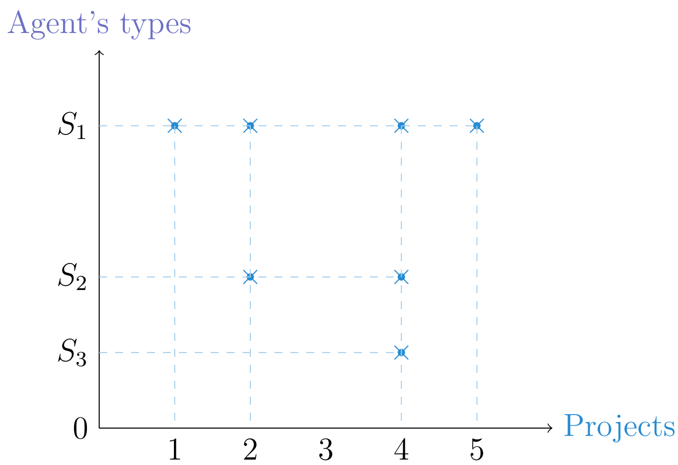

Figure 5: The type space $\mathcal{S}$ with $N > 2$ projects under Assumption 3 where project i refers to (α_i, π_i). The type space is nested such that if we take any two types, one would be a subset of the other.

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.

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 4.1.1 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 how much of a disadvantage the principal is at here; he lacks proposal power and the ability to commit to his responses to the agent's proposals. Since the principal can only implement the projects that are proposed, we might expect that that the agent can easily *hide* principal-preferred projects by never proposing them. Anticipating that his preferred projects may never be proposed, the principal would in turn capitulate and accept the agent-preferred projects when they are proposed. We show, however, that with two projects, there is always an equilibrium of the game that attains the optimal commitment payoff. We argue that it is in fact the inability to make proposals that enables the principal to *wait*, and for costly delay to emerge as a signalling device in equilibrium. For more than two projects, we identify sufficient conditions on parameters under which the commitment result still holds.

Our setup has natural applications to organizational economics, specifically empire-building by corporate managers. It is well known that managers may not always act in the best interests of the shareholders, but rather act to increase their own influence within the organization. Our analysis has implications for when the manager holds verifiable private information that's relevant to the optimal course of action for the firm, but is motivated by empire-building. We find that the by adopting a bottom-up approach, i.e. eliciting proposals from the manager, the CEO might be able to curb the manager's empire-building plans better than by issuing top-down commands to restrict what a manager can do.

We also define a general class of delegation protocols and show that if the principal is able to commit to his strategy, then the sequential delegation game we consider does as well as any other protocol. This comparison with other protocols highlights an important point. In the two project case, this commitment payoff is achieved in an equilibrium of the sequential delegation game. Therefore there is no equilibrium of any other protocol in our class that results is a strictly higher payoff for the principal. This comparison further reinforces our intuition about the bottom-up approach — even if the principal cannot commit, organizational structures that involve a bottom-up approach might do better than a whole range of other organizational structures.

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*.

Alsonso, R., & Matouschek, N. (2008). Optimal delegation. *The Review of Economic Studies*, *75*(1), 259–298.

Armstrong, M., & Vickers, J. (2010). A model of delegated project choice. *Econometrica*, *78*(1), 213–244.

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.

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. *Working Paper*.

Holmström, B. (1984). On the theory of delegation. *In Bayesian Models in Economic Theory,eds. M. Boyer and R. Kihlstrom*, 115–141.

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.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*, 97–109.

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

Van Damme, E., Selten, R., & Winter, E. (1990). Alternating bid bargaining with a smallest money unit. *Games and Economic Behavior*, *2*(2), 188–201.

Appendices

A Proofs

A.1 Proof of Proposition 1

Proof. Given that $q_{Gg} = 1$ and $q_{Mg} + q_{Mb} = 1$, our maximisation problem reduces to:

$$\max_{q_{Mg}, q_{Bb} \in [0,1]} \quad \mu_G \pi_g + \mu_B q_{Bb} \pi_b + \mu_M q_{Mg} \pi_g + \mu_M (1 - q_{Mg}) \pi_b$$

$$\text{subject to} \quad q_{Mg} \alpha_g + (1 - q_{Mg}) \alpha_b \geqslant q_{bB} \alpha_b$$

In an optimal mechanism, we must also have that $q_{Mg}\alpha_g + (1 - q_{Mg})\alpha_b = q_{bB}\alpha_b$. To see this, suppose the inequality is strict. Let ${q_{Mg}}' = q_{Mg} + \varepsilon$, and ${q_{Mb}}' = 1 - {q_{Mg}}' = q_{Mb} - \varepsilon$, where $\varepsilon > 0$. For ε small enough, IC_{MB} is still satisfied and the principal's expected payoff increases by $\varepsilon\mu_M(\pi_g - \pi_b)$. We can therefore substitute $q_{Mg}\alpha_g + (1 - q_{Mg})\alpha_b = q_{bB}\alpha_b$ into our objective function, and we get

$$\mu_G \pi_g + \mu_B \pi_b (1 - q_{Mg}(1 - \frac{\alpha_g}{\alpha_b})) + \mu_M q_{Mg} \pi_g + \mu_M (1 - q_{Mg}) \pi_b$$

$$= \mu_G \pi_g + \mu_M \pi_b + q_{Mg} \{\mu_M (\pi_g - \pi_b) - \mu_B \pi_b (1 - \frac{\alpha_g}{\alpha_b})\}$$

The only choice variable is q_{Mg} now, and whether the above expression is increasing or decreasing in q_{Mg} depends on the sign of it's coefficient, $\{\mu_M(\pi_g - \pi_b) - \mu_B\pi_b(1 - \frac{\alpha_g}{\alpha_b})\}$. If the coefficient is strictly positive, then the optimal mechanism has $q_{Mg} = 1$. Also, because IC_{MB} holds with equality, we have $q_{Bb} = \frac{\alpha_g}{\alpha_b}$ in the optimal mechanism. A bit of rearranging gives us that $\{\mu_M(\pi_g - \pi_b) - \mu_B\pi_b(1 - \frac{\alpha_g}{\alpha_b})\} > 0$ is equivalent to $\lambda > \lambda^*$. Similarly, if $\lambda < \lambda^*$, the optimal mechanism has $q_{Mg} = 0$, and therefore $q_{Bb} = 1$. If $\lambda = \lambda^*$, principal's expected payoff is constant in q_{Mg} and therefore any $q_M g \in [0, 1]$ is optimal, with q_{Bb} again being determined by the equality of IC_{MB}.

□

A.2 Proof of Theorem 1

We first establish some notation. Recall that for any time period t, the set of all possible period t histories is denoted by $\mathcal{H}_t$, where $\mathcal{H}_t = (\mathcal{N} \cup \varnothing)^t$. The representative period t history is denoted by $h_t \in \mathcal{H}_t$. The action (or proposal) space of an agent of type S at any history is given by $A_S(h_t) = A_S = \{S \cup \varnothing\}$, where $\varnothing$ represents remaining silent. An element of $A_S(h_t)$ is given by a_S^t. Given a history h_t, and a proposal in $(\mathcal{N} \cup \varnothing)$ by the agent, the principal can either *accept*, or *reject* this proposal.

A behaviour strategy maps histories and types into action spaces. For the agent of type S, $\sigma_S(a_S^t|h_t)$ denotes the probability of choosing a_S^t at history h_t. For the principal, $\sigma_P(h_t, i)$ denotes the probability of accepting proposal $i \in \mathcal{N}$ at history h_t. If at h_t, agent is silent, then $\sigma_P(h_t, \varnothing) = 0$, as there is no project to accept. At any history h_t, we denote the probability the principal attaches to type S by $\mu_S(h_t)$. If, following h_t, i is proposed, the updated beliefs are given by $\mu_S(h_t, i)$ for each S, and by $\mu_S(h_t, \varnothing)$ if the agent is silent at h_t. For any t and any $t' < t$, let $h_t(t') \in (\mathcal{N} \cup \varnothing)$ be the proposal at period t', along this history h_t. We denote by $h_{t-1}(h_t)$ the period $t-1$ history obtained by removing proposal $h_t(t-1)$ from h_t, so that $h_t = (h_{t-1}(h_t), h_t(t-1))$.

Our solution concept is Perfect Bayesian Equilibrium, as defined in Fudenberg and Tirole (1991). We want to highlight that here, something stronger than Bayes' Rule is used to update beliefs following any proposal at any history. To understand this, fix any history h_t. Even if this is a history that arises with probability zero along the equilibrium path, beliefs following a proposal i at this history are updated using Bayes' Rule if $\exists$ a type S such that $\mu_S(h_t) > 0$ and $\sigma_S(i|h_t) > 0$. Beliefs are allowed to be *completely arbitrary* only if given h_t, the proposal made by the agent had probability zero, according to the agent's strategy. However, even when following a proposal i at history h_t, beliefs are allowed to be completely arbitrary, they must still have support in $\{S \subseteq \mathcal{N} | i \in S\}$, i.e. the set of types that have i, because only these types can possibly propose i. This is in contrast to models without hard evidence.

A.2.1 Pooling equilibrium:

Strategies: The principal's strategy σ_P is: for any history h_t, and proposal i, $\sigma_P(h_t, i) = 1$, i.e. at each history, the principal accepts any proposal with probability one. If the agent is of the empty type, then for any h_t, $\sigma_E(\varnothing|h_t) = 1$. For an agent of type $S \neq E$, let $i^* = \min\{i|i \in S\}$, i.e. the most preferred project of the agent in S. At any history h_t, $\sigma_S(i^*|h_t) = 1$, so at any history, the agent always proposes her favorite available project with probability one.

Beliefs: Fix any h_t (if $t = 0$, this is the null history). (i) If the agent is silent at h_t, the beliefs are $\mu_E(h_t, \varnothing) = 1$ and $\mu_S(h_t, \varnothing) = 0$ for all $S \neq E$. (ii) If the agent proposes g at h_t, the beliefs are $\mu_G(h_t, g) = 1$ and $\mu_S(h_t, g) = 0$ for all $S \neq E$. (iii) If the agent proposes b at h_t, the beliefs are $\mu_B(h_t, b) = \frac{\mu_B}{\mu_B+\mu_M}$, and $\mu_M(h_t, b) = \frac{\mu_M}{\mu_B+\mu_M}$.

We now argue that the beliefs satisfy the requirements for PBE. At any h_t, if the agent takes a probability zero action, PBE imposes no restrictions on beliefs. So, we only need to worry about the case where the agent's proposal at h_t has positive probability given this history. In this case, PBE requires the beliefs to be determined by Bayes' Rule. At h_t, silence is a positive probability action only if $\mu_E(h_t) > 0$, since only type E has $\sigma_E(\varnothing|h_t) > 0$. In this case, the belief $\mu_E(h_t, \varnothing) = 1$ is precisely the one determined by Bayes' Rule, since $\mu_E(h_t) > 0$, $\sigma_E(\varnothing|h_t) = 1 > 0$, and $\sigma_S(\varnothing|h_t) = 0$ for all $S \neq E$. Similarly, g is a positive probability action only if $\mu_G(h_t) > 0$ and b is a positive probability action only if $\mu_B(h_t) > 0$. By identical reasoning as before, we can argue that in both these cases, the specified beliefs are the ones determined by Bayes' Rule.

Lemma 1. *The strategies and beliefs described above constitute a Perfect Bayesian Equilibrium.*

Proof. We have already argued that the beliefs that we described satisfy the requirements for being part of a PBE. Now, we only have to argue that in the continuation game starting at any history h_t, the strategies of the principal and agent constitute a Bayes Nash Equilibrium (BNE), given the principal's beliefs $\mu(h_t)$ at that history.

Fix any history h_t. If the agent is silent, there is no action for the principal to take. If the agent proposes g, then irrespective of $\mu(h_t)$, it is sequentially rational to accept g, since this is the highest payoff the principal can get. If the agent proposes b, again, the exact beliefs of the principal do not matter. Whatever the beliefs are, they have support in $\{S \subseteq \mathcal{N}|b \in S\} = \{B, M\}$. So, the principal must believe with probability one that it is a type that *has* b, and that therefore will propose b in the next period, if the principal rejects this proposal (in fact, in every future period). So, it is sequentially rational for the principal to accept this proposal too.

For the agent, if it is of type E, it can only stay silent. If it is of type G, it again has no profitable deviation to proposing g, as the principal will accept it if proposed. If it is of type B or M, again, the principal will accept b if proposed, no there is no profitable deviation to proposing b. □

Lemma 2. *The principal's payoff from the pooling equilibrium is the same as his payoff from the pooling mechanism.*

Proof. Along the equilibrium path, from type G, g is proposed and accepted at $t = 0$, and from types B and M, b is proposed and accepted at $t = 0$. This exactly replicates the implementation probabilities of $q_{Gg} = 1$, and $Q_{Bb} = q_{Mb} = 1$ from the pooling mechanism. □

A.2.2 Separating equilibrium:

Before describing the strategies and beliefs, we define two classes of histories.

Definition. *A* ***type 1*** *history h_t is one where there is* ***no*** *$t^{'} < t^*(\delta)$ such that $h_t(t^{'}) = b$. In other words, there is no $t^{'} < t^*(\delta)$ such that along h_t, b was proposed at $t^{'}$. The null history is a* ***type 1*** *history.*

Definition. *A* ***type 2*** *history h_t is one where $t \neq 0$ and there* ***is*** *a $t^{'} < t^*(\delta)$ such that $h_t(t^{'}) = b$. In other words, there is a $t^{'} < t^*(\delta)$ such that along h_t, b was proposed at $t^{'}$.*

Strategies:

- The principal's strategy σ_P is: At any history h_t, $\sigma_P(h_t, g) = 1$. So, g is accepted if proposed at any history. If h_t is such that $t < t^*(\delta)$, $\sigma_P(h_t, b) = 0$. If $t \geqslant t^*(\delta)$, and the history is of type 1, then $\sigma_P(h_t, b) = 1$. If $t \geqslant t^*(\delta)$, and the history is of type 2, then $\sigma_P(h_t, b) = 0$.

- If the agent's type is E, for any history h_t, $\sigma_E(\varnothing|h_t) = 1$.

- If the agent's type is G, for any history h_t, $\sigma_G(g|h_t) = 1$.

- If the agent's type is B, and h_t is such that $t < t^*(\delta)$, then $\sigma_B(\varnothing|h_t) = 1$. If $t \geqslant t^*(\delta)$ and h_t is of type 1, then $\sigma_B(b|h_t) = 1$. If $t \geqslant t^*(\delta)$ and h_t is of type 2, then $\sigma_B(\varnothing|h_t) = 1$.

- If the agent's type is M, then at h_0, $\sigma(g|h_0) = 1$. Now consider h_t with $t > 0$. If h_t is of type 2, then $\sigma_M(g|h_t) = 1$. If h_t is of type 1, and $t < t^*(\delta)$, then $\sigma_M(\varnothing|h_t) = 1$. If h_t is of type 1, and $t \geqslant t^*(\delta)$, then $\sigma_M(b|h_t) = 1$.

Beliefs:
At the null history h_0:

- If the agent is silent, the beliefs are $\mu_E(h_0, \varnothing) = \frac{\mu_E}{\mu_E+\mu_B}$ and $\mu_B(h_0, \varnothing) = \frac{\mu_B}{\mu_E+\mu_B}$. If the agent proposes g, the beliefs are $\mu_G(h_0, g) = \frac{\mu_G}{\mu_G+\mu_M}$, $\mu_M(h_0, g) = \frac{\mu_M}{\mu_G+\mu_M}$. If the agent proposes b, the beliefs are $\mu_M(h_0, b) = 1$.

- These beliefs all satisfy the conditions for PBE. The agent staying silent and proposing g both occur with positive probability at h_0, as $\sigma_E(\varnothing|h_0) = \sigma_B(\varnothing|h_0) = 1$ and $\sigma_G(g|h_0) = \sigma_M(g|h_0) = 1$, and these two cases, beliefs are the ones determined by Bayes' Rule. Proposing b is a probability zero action here, thus, beliefs can be arbitrary.

At a type 1 history h_t, where $t > 0$:

- If the agent is silent, beliefs are $\mu_B(h_t, \varnothing) = 1$. If the agent proposes g, beliefs are $\mu_G(h_t, g) = 1$. If the agent proposes b, and $t < t^*(\delta)$ beliefs are $\mu_M(h_t, b) = 1$. If $t \geqslant t^*(\delta)$, beliefs are $\mu_B(h_t, b) = 1$.

- The beliefs are consistent with PBE. Silence is a positive probability action at h_t only if $t < t^*(\delta)$ and $\mu_B(h_t) > 0$ or $\mu_M(h_t) > 0$. This is the case only if $h_t(t-1) = \varnothing$, and in this case beliefs are the same as the ones determined by Bayes' Rule. Proposing g is never a positive probability action at h_t so beliefs can be arbitrary in a PBE and we're done.

- Proposing b is a positive probability action only if $t \geqslant t^*(\delta)$. (1)If $t = t^*(\delta)$, this is the case only if $h_t(t-1) = \varnothing$. In this case, $\mu_B(h_t) = 1$, and $\sigma_B(b|h_t) = 1$, so beliefs are precisely the ones determined by Bayes' Rule. (2) If $t > t^*(\delta)$, then proposing b is a positive probability event only if $h_t(t-1) = \varnothing$ or $h_t(t-1) = b$. In this case, $\mu_B(h_t) = 1$ and $\sigma_B(b|h_t) = 1$, so again, the beliefs we specified are the ones determined by Bayes' Rule.

At a type 2 history:

- If the agent is silent, the beliefs are $\mu_E(h_t, \varnothing) = \frac{\mu_E}{\mu_E+\mu_B}$ and $\mu_B(h_t, \varnothing) = \frac{\mu_B}{\mu_E+\mu_B}$. If the agent proposes g, beliefs are $\mu_M(h_t, g) = 1$. If the agent proposes b, beliefs are $\mu_M(h_t, b) = 1$.

- The beliefs are consistent with PBE. If the agent is silent, there are two possibilities. Either, $h_t(t-1) = \varnothing$, in which case $\mu_E(h_t) = \frac{\mu_E}{\mu_E+\mu_B}$ and $\mu_E(h_t) = \frac{\mu_E}{\mu_E+\mu_B}$, so the beliefs that we mentioned are precisely the ones determined by Bayes' Rule. If $h_t(t-1) = b$, or $h_t(t-1) = g$, silence is a probability zero action at h_t, and PBE allows beliefs to be arbitrary.

- If the agent proposed b at h_t, irrespective of what was proposed at $t-1$, b is a probability zero action. This is because $\sigma_M(b|h_t) = \sigma_B(b|h_t) = 0$ at h_t. So again, PBE allows beliefs to be arbitrary. If agent proposed g at h_t, g is a positive probability action here only if $h_t(t-1) = b$, or $h_t(t-1) = g$. In this case since $\mu_M(h_t) = 1$ and $\sigma_M(g|h_t) = 1$, beliefs are precisely the ones determined by Bayes' Rule. If $h_t(t-1) = \varnothing$, beliefs can be arbitrary.

Lemma 3. *The strategies and beliefs described above constitute a Perfect Bayesian Equilibrium.*

Proof. We have already argued that the beliefs that we described satisfy the requirements for being part of a PBE. We must now argue that in the continuation game starting at any history h_t, the strategies of the principal and agent constitute a Bayes Nash Equilibrium (BNE), given the principal's beliefs $\mu(h_t)$ at that history.

For the principal, fix any history h_t. If the agent is silent, there is no action for the principal to take. If the agent proposes g, then irrespective of $\mu(h_t)$, it is sequentially rational to accept g, since this is the highest payoff the principal can get. If the agent proposes b, we need to consider three cases. (i) $t < t^*(\delta)$: The principal's strategy is to reject b at such a history. His beliefs following a proposal of b at h_t are $\mu_M(h_t, b) = 1$. If the principal accepts, he gets α_b, and if he rejects, he expects the agent to propose g in the next period. So, rejection is indeed sequentially rational if the principal is sufficiently patient. (ii) $t \geqslant t^*(\delta)$ and the history is of type 1. In this case, if b is proposed, the principal's beliefs are $\mu_B(h_t, b) = 1$ and if he rejects b, he expects b to be proposed again in the next period. So, accepting b is sequentially rational. (iii) $t \geqslant t^*(\delta)$ and the history is of type 2. In this case, the principal's strategy is to reject b. His beliefs following a proposal of b are $\mu_M(h_t, b) = 1$, so the same reasoning as case (i) follows, and rejection is sequentially rational.

For the agent, fix a history h_t. If her type is E, she can only be silent at any history. If her type is G, her strategy is to propose g, which is optimal, since the principal would accept it. If her type is B, and (i) $t < t^*(\delta)$, her strategy is to stay silent. Consider a one-shot deviation where she deviates by proposing b. This proposal would be rejected and since her strategy involves staying silent at any future period following this rejection, therefore getting a payoff of zero, she cannot be better off by this deviation. (ii) If $t \geqslant t^*(\delta)$ and the history is of type 1, proposing b is optimal since it would be accepted and α_b is the highest payoff the agent can get. (iii)If $t \geqslant t^*(\delta)$ and the history is of type 2, silence is optimal. Consider a one-shot deviation where the agent proposes b instead. It would be rejected, and the agent's strategy is to stay silent in each period that follows. So, this is not a profitable deviation.

If the agent's type is M, at h_0, her strategy is to propose g. Consider the one-shot deviation where she proposes b instead. It will be rejected and she will propose g at $t = 1$, which will get accepted. Clearly, this is not profitable, as she can propose g at $t = 0$ and it will get accepted. If the one-shot deviation involves silence at $t = 0$, her strategy then is to stay silent till $t^*(\delta)$, at which point she proposes b can it is accepted. So her payoff is $\delta^{t^*(\delta)}\alpha_b$ which is $\leqslant \alpha_g$. So, this deviation is not profitable either. We can similarly rule out deviations at other histories.

$\square$

A.3 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.

Fix a type S and a project $i \in S$. Recall that at any history h_t, the proposer is $P(h_t)$. Let $x(i|h_t)$ be the probability with which, according to $P(h_t)$'s strategy, the proposer $P(h_t)$ offers $O(h_t) \in \mathcal{O}(h_t)$ such that $i \in O(h_t)$ at h_t. Let $y(h_t, i)$ be the probability that the other party, who is not the proposer, *accepts* i in the offer $O(h_t)$. We define the probability of any history inductively. We denote period 0 history by h_0, so at $t = 1$, for any $h_1 = (h_0, i)$, where $i \in S \cup \{\varnothing\}$. We define $\nu(h_1) = x(i|h_0)(1 - y(h_0, i))$, 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 $\nu(h_{t+1}) = \nu(h_t)x(i|h_t)(1 - y(h_t, i))$.

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

$$p_S(i, t) := \sum_{h_t \in \mathcal{H}_t} \nu(h_t)x(i|h_t)y(h_t, i)$$

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.4 Proof of Theorem 2

Proof. 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 $h_N^{S'}$. 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_N^S, 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_N^S, the principal rejects any proposal at any $t > N$.

- At the history $(h_N^S, i_1, i_2 \ldots, i_{k-1})$ 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 $(h_N^S, i_1, i_2, \ldots, i_{k-1})$, 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 delegation protocol and 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 $\mathcal{S} = \{\{1, 2\}, \{2\}, \{2, 3\}\}$.

Proof. Suppose there is a commitment strategy of the principal in this alternative game, and a best response of the agent that attains the payoff from the optimal mechanism. This commitment strategy and best response give rise to an *induced* SCF. Recall from the construction of the induced SCF, that for any type S, and any $i \in S$, we have $f^I_{S'}(i) = \sum_{t=0}^{\infty} \delta^t p_{S'}(i, t)$. We therefore must have that $f^I_{2,3}(2) \to 1$ and $f^I_{1,2}(2) \to \frac{3}{8}$ as $\delta \to 1$. Recall that a history here is simply a sequence of restriction sets that have been rejected by the agent. Now, for type $\{2, 3\}$, consider any t, and history h_t at which 2 is implemented with positive probability. At this history, it must be that $2 \in O(h_t)$ and this is accepted by $\{2, 3\}$ with positive probability. But, even $\{1, 2\}$ can accept 2 at this history. Moreover, since the history involves a sequence of rejections, and the ability to reject is not type-dependent, *any* such history can also be reached when the type is $\{2, 3\}$. This contradicts the fact that $f^I_{1,2}(2) \to \frac{3}{8}$. $\square$

A.5 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:

$$E_1 \geqslant E_2, \ldots \geqslant E_M,$$

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

Lemma 4. *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 5. *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 6. *Let the set of all projects in types $\{S_1, S_2 \ldots S_M\}$ be $\{\alpha_1, \alpha_2, \ldots, \alpha_N\}$, where $\alpha_1 < \alpha_2 \ldots \alpha_N$. For any $v \in \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$, 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 7. *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.*[17]

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 6, 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:

[17] 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.

- 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.

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 (6).

$\square$

$\square$