Interacting mechanisms: a perspective on generalized principal-agent problems*

Françoise Forges † Frédéric Koessler ‡ Andrés Salamanca §

March 5, 2024

Abstract

Myerson (1982) formalizes general principal-agent problems, in which agents have private information and choose actions. His contribution is best known for a version of the revelation principle in the case of a single principal but he also introduces a model of interacting principals. We push the latter forward by studying the perfect Bayesian equilibrium outcomes of the corporations' game in which every principal proposes a mechanism to his agents. We show that several versions of the revelation principle hold in our framework and that, under certain conditions, every principals' equilibrium, as defined in Myerson (1982), is a perfect Bayesian equilibrium outcome of the corporations' game.

KEYWORDS: communication equilibrium; corporations; correlated equilibrium; mechanism design; multiple principals; perfect Bayesian equilibrium; revelation principle; robust equilibrium.

JEL CLASSIFICATION: C72; D82.

*Andrés Salamanca dicloses support of this research from the European Union-Next Generation EU through the Italian National Recovery and Resilience Plan (NRRP)—Mission 4, Component 2, Investment 1.1: Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN)—under the project "Political Persuasion," grant number 2022WS8AJY–CUP H53D23002600006. This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

†Université Paris-Dauphine, PSL University, LEDa, France

‡HEC Paris, GREGHEC-CNRS, France

§Ca' Foscari University of Venice, Italy

Contents

1 Introduction

Brief summary of Myerson (1982)

In the author's own words, his main goal is to formulate the "general principal-agent problem", in which an individual (such as the owner of a corporation) must make optimal decisions and also optimally coordinate agents (such as his managers and employees). Specific economic applications of this problem had already been studied in the years 1970's, e.g., in the insurance sector. The related literature identified two kinds of constraints potentially restricting the coordination systems that are feasible for the principal: the agents have private information that he cannot directly observe (which can generate "adverse selection") and they choose private actions that he cannot directly control (which can generate "moral hazard"). Myerson (1982) provides a general abstract framework to account for the principal's power given these constraints. A basic assumption is that the principal can design a costless, possibly multi-stage, communication process along which he can ask his agents to answer questions and send them instructions, as well as commit to decisions.

More precisely, by relying on Harsanyi (1967/8)'s insights, Myerson (1982) summarizes every agents' private information into a "type", which can take finitely many values. He assumes that finite sets of actions for the principal and the agents, as well as type dependent (von Neumann - Morgenstern) utility functions, are given exogenously. The problem of the principal is to coordinate his decision and those of his agents so as to maximize his own expected utility. To achieve this goal, he can use a variety of communication mechanisms, some of which may be so complex that they are hard to formalize. On the contrary, other ones are very simple; for instance, a *direct mechanism* just asks every agent to secretly report a type to the principal who in turn chooses an action for himself and privately recommends an action to every agent. Such a direct mechanism induces a simple (Bayesian) game among the agents. If the truthful and obedient strategies form a Nash equilibrium of this game, the mechanism is said to be *incentive-compatible*. Myerson (1982)'s main message is that it is not necessary to enter the details of complicated communication mechanisms: the principal's optimal incentive-compatible direct mechanism is also optimal in the (remarkably flexible) class of *all* communication mechanisms. This is a general version of the famous "revelation principle", which was already known in more specific settings (see, for instance, Holmström, 1977, Dasgupta, Hammond, and Maskin, 1979 and Myerson, 1979). The result had – and still has – a tremendous impact in mechanism design and information design.

Myerson (1982)'s Proposition 1 states that the problem of computing the principal's optimal incentive-compatible direct mechanism is a linear programming problem. The relevance of this result comes from the general revelation principle mentioned above, which is stated as Proposition 2. The proof of this result actually establishes that the set of all *equilibrium outcomes* that can be achieved with the help of an arbitrary communication mechanism coincides with the set of incentive-compatible direct mechanisms. Myerson (1982) notes the connection between his approach and Aumann (1974)'s correlated equilibrium, which accounts for the players' possible coordination over actions. However, he sticks to the formulation in which the principal, as opposed to a benevolent mediator, has a utility function which he seeks to maximize. As clear from Myerson (1991)'s Section 6.3, Myerson (1982)'s framework implicitly incorporates a solution concept for Bayesian games, to be known later as the "communication equilibrium".

About two thirds of Myerson (1982)'s article are devoted to the optimization problem described above, which involves a *single principal* and several agents. The other third extends the model by considering several principals, who *interact but do not compete*. By contrast, competing principals are at the core of a wide literature on multiple principals, a framework that is fully justified by a variety of economic applications (see Martimort, 2007 for a survey). Yet, in other relevant situations (see, e.g., Katz, 1991 and Martimort, 1996), "corporations"—consisting of a principal and his agents—are already constituted when decisions have to be made. The utility of every principal and every agent still depends on all types and actions, inside as well as outside his own corporation. But every principal can just propose a coordination mechanism to his own agents, who can just react to their principal's proposal, possibly by choosing an outside option. In other words, the agents do not have the opportunity to choose which principal to join.

Myerson (1982) formalizes his model of interacting principals as a *generalized game* (see Debreu, 1952), in which every principal chooses a direct mechanism that is incentive-compatible given the other principals' direct mechanisms. A "principals' equilibrium" is defined as a Nash equilibrium of the generalized game. Proposition 3 states that such an equilibrium does not always exist, which is established by a counterexample. The paper ends with the definition of a more permissive notion, the "quasi-equilibrium", for which an existence result applies (Proposition 4).

Our contribution

We push Myerson (1982)'s approach forward by studying the equilibria of an explicit "grand game" between corporations, consisting each of a principal and his agents. This corporations' game, which is formally defined in Section 3.2, starts with the simultaneous choice of a general mechanism by each principal. Knowing his type and the mechanism of his principal, every agent privately sends a report to his principal. Then, every agent gets a private message, which is selected by his principal's mechanism. Finally, at the last stage of the corporations' game, actions are taken. More precisely, for every principal, an ("enforceable") action is implemented according to his mechanism, while every agent makes his choice knowing the mechanism, his type, his report and the message from his principal's mechanism. The utility of every player (principal or agent) depends on all types and actions.

Games like the one described above, starting with the principals' choice of a mechanism, are studied in the literature on competing mechanisms (see, e.g., McAfee, 1993 and Martimort and Stole, 2002). An important difference is that our interacting principals only propose a contract to the agents of their own corporation. In a consistent way with this interpretation, we assume, for most of the paper, that every agent only observes the mechanism proposed by his own principal (and only communicates with him). Thanks to this assumption, the information of every agent, when he sends his report to his principal, only consists of his type, as chosen at the beginning of the corporation's game. We can thus hope to avoid the difficulties that lead to the failure of the revelation principle when principals are competing. Some care is nonetheless necessary. To start with, by contrast with Myerson (1982)'s analysis—which does not model a grand game—some refinement of Nash equilibrium is necessary, at least to prevent the agents to make non-credible threats against their principal if the latter proposes an unexpected mechanism. Even in the case of a single corporation, it has been observed that the revelation principle is not straightforward for refined equilibria (see, e.g., Dhillon and Mertens, 1996, Gerardi, 2004, and Gerardi and Myerson, 2007).

By adapting Fudenberg and Tirole (1991)'s insights, we propose a notion of perfect Bayesian equilibrium (PBE) for the corporations' game (Definition 5 in Section 3.2). To test whether given strategies (for the principals and the agents) form such an equilibrium, we have to account for sequential rationality in the "small" games that are induced by every principal's mechanism in his corporation, given the other corporations' strategies. These induced games are the topic of Section 2, which starts by defining incentive-compatible direct mechanisms for every corporation before formalizing general mechanisms. Not surprisingly, the revelation principle holds for Nash equilibria in the induced games. More precisely, for every corporation, keeping fixed the outcome in the other ones, the set of all Nash equilibrium outcomes that can be achieved by using some general mechanism coincides with the set of incentive-compatible direct mechanisms (Proposition 1).

If there is a single corporation (i.e., if in all corporations but one, all sets of types and actions are singletons), Proposition 1 reduces to a familiar form of the revelation principle, which, as already pointed out above, encompasses the canonical representation of communication equilibria in Bayesian games and correlated equilibria in games with complete information. These results are detailed in Section 2.6.

The analog of Proposition 1, with PBE instead of Nash equilibrium, does not necessarily hold. We illustrate this on Example 3, which is inspired by an example in Gerardi (2004) and Gerardi and Myerson (2007). The result can nevertheless be restored if, in the corporation under consideration, either the probability distribution over types has full support or no agent has a private action (i.e., agents have private information but only the principal has to choose an action). This is the content of Proposition 3.

Myerson (1982)'s solution concept, the "principals' equilibrium" mentioned above, can be defined in terms of incentive-compatible direct mechanisms. We refer to it as "M-equilibrium". We first show that M-equilibria coincide with *robust* equilibria, which are such that, in every corporation, keeping fixed the strategies of the other ones, no principal can unilaterally improve his expected utility by choosing another mechanism, *whatever his agents' reaction* (Proposition 2). We show that, under relatively mild assumptions, M-equilibria are—demanding—PBE outcomes of the corporations' game. More precisely, if, in every corporation, either the probability distribution over types has full support or no agent has a private action, then *every M-equilibrium is a PBE outcome of the corporations' game* (Proposition 4). Our corporations' game is thus an adequate substitute for the generalized game considered by Myerson (1982), a property that is not satisfied by conceivable variants of the game (see below). The converse of Proposition 4 does not hold. Indeed, even in the case of a single principal, the corporations' game (namely, the game starting with the principal's choice of a mechanism) may have a PBE in which the principal's expected utility is not as high as in any optimal incentive-compatible mechanism.

As recalled above, M-equilibria may fail to exist. We recall Myerson (1982)'s counterexample, in which each corporation consists of a principal and an agent with private information but no private action, as Example 2. Before that, we provide another counterexample (Example 1), in which each corporation involves only one agent, with private actions but no private information. Furthermore, in this counterexample, the principals have no enforceable action. It turns out that, in both examples, the corporations' game has a PBE, which leaves some hope for an existence result in this framework. In any case, the features of Example 1 are not incidental, in the sense that if principals have no enforceable action, the corporations' game indeed has a PBE (Proposition 5). A general existence result is nevertheless beyond the scope of our paper.

In the corporations' game, every principal can choose a general mechanism. Given the importance of the revelation principle, it seems natural to consider a variant of the game, the "direct corporations' game", in which the principals' choices are restricted to direct mechanisms. It is not difficult to show that, under suitable assumptions (the same as in Proposition 4 above), every PBE outcome of the corporations' game is a PBE outcome of the direct corporations' game. This result, which is stated as Proposition 6, already implies that there is "no loss of generality" to focus on direct mechanisms. Yet doing so may *enlarge* the set of PBE outcomes of the corporations' game. The next question is whether the converse of Proposition 6 holds, namely, whether every PBE outcome of the direct corporations' game is a PBE outcome of the corporations' game. At first sight, in the case of a single principal, the latter statement looks close to the trivial direction of the standard revelation principle, for which one considers all equilibrium outcomes of all games that are associated to some mechanism. However, even in this particular case, the converse of Proposition 6 looks like a strong result, because, our framework consists of a *single game*, in which the principal has a huge set of strategic choices—all the general mechanisms. The converse of Proposition 6 says that, if the principal cannot benefit from deviating to any direct mechanism (given the sequentially rational reaction of his agents), he cannot benefit either from deviating to any general mechanism. We establish the result for the corporations' game, with multiple principals, but under specific assumptions, namely, in every corporation, either there is a single agent or the principal has no enforceable action (Proposition 7).

Our corporations' game is consistent with Myerson (1982)'s framework in assuming that every agent, belonging to a single corporation, only observes the mechanism proposed by his own principal. But as noted by Attar, Campioni, and Piaser (2023), Myerson (1982)'s structure is compatible with a "public corporations' game", in which the principals' chosen mechanisms are announced to all agents. The definition of PBE in the latter game can be simplified by a representation in which nature chooses the agents' types after the principal's choices: agents' choices then become part of a proper *subgame*. Example 4 illustrates that M-equilibria may fail to be PBE outcomes of the public corporations' game, i.e., that Proposition 4 does not hold in the latter game. Example 5 illustrates that Proposition 6 does not survive either. In spite of these negative results, one might still argue, as Attar et al. (2023) do, that this alternate model of interacting principals is sensible in some applications and promising as far as existence of PBE is concerned.

2 Mechanism design for multiple corporations

2.1 Notations and basic definitions

We adopt the multiple corporations model presented by Myerson (1982, Section 4). For the sake of simplifying notations, we focus our attention on only two corporations. Let $k \in \{1, 2\}$ denote a typical *principal*, and let $l \neq k$ represent the other principal. Every principal k coordinates a set of *agents* numbered as $1, \ldots, n_k$, and the sets of agents for principals k and l are mutually exclusive. Principal k and his n_k agents collectively form what we refer to as *corporation* k.

Our notations are as follows: T_k^i is the set of *types* for agent i in corporation k, D_k^i is the set of *private actions* for agent i in corporation k and D_k^0 is the set of *enforceable actions* for principal k. These sets are assumed to be nonempty and finite. Our terminology is as follows: there is *no private action* (or no "moral hazard") in corporation k if $|D_k^i| = 1$ for every $i = 1, \ldots, n_k$; there is *no enforceable action* in corporation k if $|D_k^0| = 1$, i.e., if principal k has no direct control over actions; there is *no private information* (or no "adverse selection") in corporation k if $|T_k^i| = 1$ for every $i = 1, \ldots, n_k$.

For every corporation k and agent i in corporation k, let

$$D_k = \prod_{i=0}^{n_k} D_k^i, \quad D_k^{-i} = \prod_{\substack{j=0 \\ j \neq i}}^{n_k} D_k^j, \quad D = D_1 \times D_2,$$

$$T_k = \prod_{i=1}^{n_k} T_k^i, \quad T_k^{-i} = \prod_{\substack{j=1 \\ j \neq i}}^{n_k} T_k^j, \quad T = T_1 \times T_2.$$

The (common) *prior* distribution on the set of type profiles is denoted by $p \in \Delta(T)$. Without loss of generality, we assume that the marginal distribution of T_k^i has full support for every k and i. The utility function of principal k is $v_k : D \times T \to \mathbb{R}$. The utility function of agent i in corporation k is $u_k^i : D \times T \to \mathbb{R}$.

A *direct mechanism for corporation* k is a mapping $\pi_k : T_k \to \Delta(D_k)$. A direct mechanism can be interpreted as follows: Principal k asks the agents in corporation k to simultaneously and confidentially report their types. Then the principal chooses an action $d_k^0 \in D_k^0$ and makes a private recommendation $d_k^i \in D_k^i$ to each agent i in corporation k. We denote as $\pi_k(d_k^0, d_k^1, \ldots, d_k^{n_k} \mid t_k)$ the conditional probability that the principal chooses d_k^0 and recommends action d_k^i to each agent i, when the agents in corporation k have reported the type profile t_k.

2.2 Incentive-compatible mechanisms

For every corporation k, every agent i in corporation k, every $t_k^i, r_k^i \in T_k^i$, every function $\delta_k^i : D_k^i \to D_k^i$, and every pair of direct mechanisms (π_k, π_l), let

$$U_k^i(\pi_k, \pi_l, \delta_k^i, r_k^i \mid t_k^i) = \sum_{t_k^{-i} \in T_k^{-i}} \sum_{t_l \in T_l} p(t_k^{-i}, t_l \mid t_k^i) \sum_{d \in D} \pi_k(d_k \mid t_k^{-i}, r_k^i) \pi_l(d_l \mid t_l) u_k^i(d_k^{-i}, \delta_k^i(d_k^i), d_l, t). \quad (1)$$

That is, if principals use the direct mechanisms (π_k, π_l), then $U_k^i(\pi_k, \pi_l, \delta_k^i, r_k^i \mid t_k^i)$ is the conditionally expected utility of agent i in corporation k, given that his type is t_k^i, he reports r_k^i, and plans to play action $\delta_k^i(d_k^i)$ when d_k^i is recommended, while all other agents report their types truthfully and follow their recommended actions obediently.

Let Id_k^i be the identity function on D_k^i. For each corporation k, a direct mechanism for corporation k is (Bayesian) incentive compatible given the direct mechanism for corporation l if, for every agent in corporation k, it is optimal for him to report his type truthfully and follow his recommended action obediently when other agents also report their type truthfully and follow their recommended action obediently. Formally:

Definition 1. *A direct mechanism π_k for corporation k is* incentive-compatible (IC) for corporation k given π_l *iff*

$$U_k^i(\pi_k, \pi_l, \mathrm{Id}_k^i, t_k^i \mid t_k^i) \geq U_k^i(\pi_k, \pi_l, \delta_k^i, r_k^i \mid t_k^i),$$

for every $i \in \{1, \ldots, n_k\}$, $t_k^i, r_k^i \in T_k^i$, and $\delta_k^i : D_k^i \to D_k^i$.

The set of direct mechanisms for corporation k which are IC given π_l is non-empty, compact, and convex, denoted by $F_k(\pi_l) \subseteq \Delta(D_k)^{T_k}$. Let

$$V_k(\pi_k, \pi_l) = \sum_{t \in T} p(t) \sum_{d \in D} \pi_k(d_k \mid t_k) \pi_l(d_l \mid t_l) v_k(d, t),$$

be the expected utility of principal k given the direct mechanisms (π_k, π_l) when all agents are truthful and obedient.

2.3 Principals' M-equilibrium

The set of direct mechanisms (π_1, π_2) that are *jointly IC*, meaning that $\pi_1 \in F_1(\pi_2)$ and $\pi_2 \in F_2(\pi_1)$, is non-empty.[1] A pair of jointly IC mechanisms, for which no principal can improve his expected utility by choosing an alternative IC mechanism given the direct mechanism of the other principal, has been referred to as a principals' equilibrium by Myerson (1982). We refer to it as an M-equilibrium to distinguish it from other equilibrium concepts introduced later in the paper. Formally:

Definition 2 (Myerson, 1982). *A pair of direct mechanisms (π_1^*, π_2^*) is an* M-equilibrium *iff*

$$\pi_1^* \in \arg \max_{\pi_1 \in F_1(\pi_2^*)} V_1(\pi_1, \pi_2^*) \quad \textit{and} \quad \pi_2^* \in \arg \max_{\pi_2 \in F_2(\pi_1^*)} V_2(\pi_1^*, \pi_2).$$

We denote by $\mathsf{ME} \subseteq \Delta(D_1)^{T_1} \times \Delta(D_2)^{T_2}$ the set of M-equilibria.

[1] A direct way to prove this assertion is to observe that every correspondence $F_k : \Delta(D_k)^{T_k} \rightrightarrows \Delta(D_k)^{T_k}$ is upper hemi-continuous with non-empty, convex, and compact values, and then use the Kakutani fixed-point theorem.

2.4 Generalized mechanisms and the revelation principle

A (generalized) *mechanism for corporation* k is denoted by $\mathcal{M}_k = (R_k, M_k, \gamma_k)$, where $R_k = \prod_{i=1}^{n_k} R_k^i$, R_k^i is the nonempty and finite set of possible reports from agent i in corporation k to his principal, $M_k = \prod_{i=1}^{n_k} M_k^i$, M_k^i is the nonempty and finite set of possible messages from principal k to agent i in corporation k, and $\gamma_k : R_k \to \Delta(D_k^0 \times M_k)$. A mechanism for corporation k is direct if $R_k^i = T_k^i$ and $M_k^i = D_k^i$ and, with some slight abuse of notation, is simply denoted by $\pi_k : T_k \to \Delta(D_k)$ as in the previous subsections.

In this section, we fix a direct mechanism π_l for corporation l and assume that the agents in corporation l are truthful and obedient. Then, a mechanism $\mathcal{M}_k$ for corporation k induces an n_k-player multistage game $G_k(\mathcal{M}_k, \pi_l)$ played by the agents in corporation k, described as follows, where the choices of the agents in stage 2 (the reporting stage) and stage 4 (the action stage) are made simultaneously:

1. Nature selects the type profile $t_k \in T_k$ in corporation k according to the prior $p_k \in \Delta(T_k)$, where $p_k(t_k) = \sum_{t_l \in T_l} p(t_k, t_l)$. Every agent i in corporation k privately learns $t_k^i \in T_k^i$.
2. Every agent i in corporation k privately sends a report $r_k^i \in R_k^i$ to his principal.
3. Action $d_k^0 \in D_k^0$, and the profile of messages $m_k \in M_k$ are drawn with probability $\gamma_k(d_k^0, m_k \mid r_k)$. Every agent i in corporation k privately observes message $m_k^i \in M_k^i$ from his principal.
4. Every agent i in corporation k chooses an action $d_k^i \in D_k^i$.

In $G_k(\mathcal{M}_k, \pi_l)$, the payoff of each agent i in corporation k is $\tilde{u}_k^i(d_k, t_k; \pi_l)$, where for every t_k in the support of p_k:

$$\tilde{u}_k^i(d_k, t_k; \pi_l) := \sum_{t_l \in T_l} p(t_l \mid t_k) \sum_{d_l \in D_l} \pi_l(d_l \mid t_l) u_k^i((d_k, d_l), (t_k, t_l)).$$

If $\mathcal{M}_k = (T_k, D_k, \pi_k)$ is a direct mechanism, then the game $G_k(\mathcal{M}_k, \pi_l)$ is simply denoted by $G_k(\pi_k, \pi_l)$. For each agent i in corporation k, let

$$R_k^i \otimes M_k^i = \{(r_k^i, m_k^i) \in R_k^i \times M_k^i : \gamma_k(m_k^i \mid r_k^i, r_k^{-i}) > 0 \text{ for some } r_k^{-i} \in R_k^{-i}\}.$$

That is, given the mechanism $\mathcal{M}_k$, agent i can receive message m_k^i after reporting r_k^i for some strategies of the other agents in corporation k if and only if $(r_k^i, m_k^i) \in R_k^i \otimes M_k^i$. A (behavioral) *participation strategy* for agent i in corporation k in the game $G_k(\mathcal{M}_k, \pi_l)$ is given by a pair (ρ_k^i, σ_k^i), where

$$\rho_k^i : T_k^i \to \Delta(R_k^i),$$

is the reporting strategy of agent i, and

$$\sigma_k^i : T_k^i \times R_k^i \otimes M_k^i \to \Delta(D_k^i),$$

is the action strategy of agent i. In the direct game $G_k(\pi_k, \pi_l)$, the participation strategy of agent i is *truthful and obedient* if $\rho_k^i(t_k^i \mid t_k^i) = \sigma_k^i(d_k^i \mid t_k^i, t_k^i, d_k^i) = 1$ for every t_k^i and d_k^i.

A mechanism $\mathcal{M}_k$ and agents' participation strategies in corporation k induce an *outcome for corporation* k, denoted by $\phi_k : T_k \to \Delta(D_k)$, where $\phi_k(d_k \mid t_k)$ is the probability of the action profile $d_k \in D_k$ being played in corporation k, given the type profile $t_k \in T_k$ within corporation k. Formally, for every $t_k \in T_k$ and $d_k \in D_k$, we have

$$\phi_k(d_k \mid t_k) = \sum_{r_k \in R_k} \left(\prod_{i=1}^{n_k} \rho_k^i(r_k^i \mid t_k^i) \right) \sum_{m_k \in M_k} \gamma_k(d_k^0, m_k \mid r_k) \left(\prod_{i=1}^{n_k} \sigma_k^i(d_k^i \mid t_k^i, r_k^i, m_k^i) \right).$$

Observe that for a direct mechanism π_k, if agents in corporation k are truthful and obedient, then the outcome for corporation k is simply $\phi_k = \pi_k$.

The game $G_k(\mathcal{M}_k, \pi_l)$ is a finite game, so the set of Nash equilibrium outcomes of $G_k(\mathcal{M}_k, \pi_l)$, denoted by $\mathcal{E}_k(\mathcal{M}_k, \pi_l) \subseteq \Delta(D_k)^{T_k}$, is non-empty. By definition, π_k is IC given π_l iff π_k is the outcome of a Nash equilibrium of the game $G_k(\pi_k, \pi_l)$ in which agents' participation strategies are truthful and obedient. In particular, $\pi_k \in F_k(\pi_l)$ implies $\pi_k \in \mathcal{E}_k(\pi_k, \pi_l)$. The converse is also true, and more generally, Myerson (1982) proves that the set of all Nash equilibrium outcomes of $G_k(\mathcal{M}_k, \pi_l)$, for all possible mechanisms $\mathcal{M}_k$, has a canonical representation: it is simply the set of IC mechanisms for corporation k given π_l. This general and important result is referred to as the *revelation principle*.

Proposition 1 (Revelation principle: canonical representation)**.**

$$\bigcup_{\mathcal{M}_k} \mathcal{E}_k(\mathcal{M}_k, \pi_l) = F_k(\pi_l).$$

A direct consequence of this canonical representation, which is mostly used in applications, is that, given π_l in corporation l, to maximize the expected utility of principal k, it is without loss of generality for principal k to consider direct mechanisms as well as truthful and obedient strategies for the agents in corporation k (Myerson, 1982, Proposition 2).[2] Formally:

Corollary 1.

$$\max_{\substack{\mathcal{M}_k, \\ \phi_k \in \mathcal{E}_k(\mathcal{M}_k, \pi_l)}} V_k(\phi_k, \pi_l) = \max_{\pi_l \in F_k(\pi_l)} V_k(\pi_k, \pi_l).$$

Because the set of incentive-compatible mechanisms is characterized by finitely many linear inequalities, and the expected utility of principal k is linear in π_k, another consequence of the canonical representation is that the problem of computing the optimal mechanism of principal k given π_l is a linear programming problem (Myerson, 1982, Proposition 1).

2.5 M-equilibria as robust equilibria

Consider a pair of jointly IC mechanisms (π_1^*, π_2^*). We define this pair of direct mechanisms as a robust equilibrium if, for every k, given π_l^*, principal k cannot improve his expected utility by choosing any alternative (not necessarily direct) mechanism $\mathcal{M}_k$, regardless of the Nash equilibrium of $G_k(\mathcal{M}_k, \pi_l^*)$ played by agents in corporation k. In other words, the direct mechanism π_k^* is robustly optimal for principal k given π_l^* with respect to all possible mechanisms and all possible induced equilibria in corporation k. Formally:

Definition 3 (Robust equilibrium)**.** *A pair of direct mechanisms* (π_1^*, π_2^*) *is a* robust equilibrium *iff, for every corporation* k*, we have* $\pi_k^* \in F_k(\pi_l^*)$*, and for every mechanism* $\mathcal{M}_k$ *and every outcome* $\phi_k \in \mathcal{E}_k(\mathcal{M}_k, \pi_l^*)$*, we have*

$$V_k(\pi_k^*, \pi_l^*) \geq V_k(\phi_k, \pi_l^*). \tag{2}$$

The notion proposed here is akin to various concepts of *strongly robust equilibrium* that have been developed in the competing mechanisms literature (see, for example, Epstein and Peters, 1999, Peters, 2001, Han, 2007, Attar, Campioni, and Piaser, 2018). In this literature, strongly robust equilibria are defined in a game in which principals publicly post mechanisms. Proposition 4 in Section 3.3 will show that robust equilibria as defined above are also (perfect Bayesian) equilibria of a well-defined game,

[2] While only Corollary 1 is stated in the seminal paper, as Myerson (1982, Proposition 2), its proof actually establishes Proposition 1.

which differs from the previous one in that each principal proposes a mechanism to the agents of his own corporation only.

From the revelation principle (Corollary 1), we immediately obtain the following equivalence between M-equilibria and robust equilibria:

Proposition 2. *A pair of direct mechanisms* (π_1^*, π_2^*) *is an M-equilibrium if and only if it is a robust equilibrium.*

2.6 Particular cases with a single corporation

In this section, we examine a specific instance of the preceding framework, in which both the sets of types and actions of one of the two corporations are reduced to singletons. In other words, we assume that there is a single corporation. As pointed out in the Introduction, Myerson (1982) starts with this model and goes on with it for almost two thirds of the article. The main feature of this model is that privately informed agents take utility relevant actions that are not observed by the principal.

Throughout this section, we use similar notations as above but without mentioning any index k or l: the principal's set of enforceable actions is D^0, the set of types of agent i is T^i, his set of actions is D^i, $i = 1, \ldots, n$ and

$$D = \prod_{i=0}^{n} D^i, \quad D^{-i} = \prod_{\substack{j=0 \\ j \neq i}}^{n} D^j,$$

$$T = \prod_{i=1}^{n} T^i, \quad T^{-i} = \prod_{\substack{j=1 \\ j \neq i}}^{n} T^j.$$

The prior distribution on the set of type profiles is $p \in \Delta(T)$, the utility function of the principal is $v : D \times T \to \mathbb{R}$ and the utility function of agent i is $u^i : D \times T \to \mathbb{R}$, $i = 1, \ldots, n$.

A direct mechanism is a mapping $\pi : T \to \Delta(D)$. Definition 1 can be readily adapted to the current, single principal framework: a direct mechanism is incentive compatible if it is optimal for every agent to reveal his type and choose the recommended action when all other agents are truthful and obedient.

Let $F \subseteq \Delta(D)^T$ denote the set of IC direct mechanisms and let

$$V(\pi) = \sum_{t \in T} p(t) \sum_{d \in D} \pi(d \mid t) v(d, t)$$

be the expected utility of the principal given the direct mechanism π when all agents are truthful and obedient.

An M-equilibrium (see Definition 2) reduces to an *optimal direct mechanism*, namely, a direct mechanism π^* such that

$$\pi^* \in \arg\max_{\pi \in F} V(\pi).$$

Corollary 1 in turn reduces to a well-known form of the revelation principle, namely, any optimal direct mechanism π^* is also optimal in the class of all (generalized) mechanisms. A closer look at Proposition 1 is worthwhile because, in addition to the popular applications to mechanism design, this more general statement contains representation results for game theoretic solution concepts like correlated equilibrium or communication equilibrium.

A (generalized) mechanism $\mathcal{M} = (R, M, \gamma)$ can be defined as in subsection 2.4 by introducing a finite set R^i of possible reports for agent i, a finite set of possible messages M^i from the principal to agent i and a mapping $\gamma : R \to \Delta(D^0 \times M)$, where $R = \prod_{i=1}^n R^i$ and $M = \prod_{i=1}^n M^i$. A direct mechanism $\pi : T \to \Delta(D)$ corresponds to $R^i = T^i$ and $M^i = D^i$. Every mechanism $\mathcal{M}$ induces an n-player multistage game $G(\mathcal{M})$, with $\mathcal{E}(\mathcal{M})$ its set of Nash equilibrium outcomes. Recalling that F is the set of all IC direct mechanisms, Proposition 1 states that

$$\bigcup_{\mathcal{M}} \mathcal{E}(\mathcal{M}) = F.$$

Myerson (1982) explicitly identifies the particular case in which the principal has no enforceable action and the agents do not have private information (i.e., the sets D^0 and T^i, $i = 1, \dots, n$, are all singletons) and relates it to Aumann (1974)'s notion of correlated equilibrium. In this case, the set of actions D^i and the utility functions u^i, $i = 1, \dots, n$, define a *strategic form game*. Aumann (1974) extends such a game by means of a *correlation device*, which consists of finite sets of messages M^i, $i = 1, \dots, n$, together with a probability distribution γ over M. According to the current terminology, a correlation device is a particular mechanism $\mathcal{M} = (M, \gamma)$, in which the sets of reports are singletons. A *correlated equilibrium* can be defined as a Nash equilibrium of the extended game $G(\mathcal{M})$. A direct mechanism π is just a probability distribution over D, which can be used as a *canonical* correlation device to privately recommend actions to the players before they engage in the strategic form game. Definition 1 drastically simplifies in the particular case at hand: the direct mechanism π is IC if and only if it satisfies obedience conditions, expressing that the obedient strategies form a Nash equilibrium of $G(\pi)$. The *canonical representation* of correlated equilibria can be obtained as a corollary of Proposition 1: the set of all correlated equilibrium outcomes coincides with the set of probability distributions over D satisfying the obedience conditions.[3]

Between the particular case just described and the single principal framework of this section, there is a model that is implicitly part of Myerson (1982) but not yet developed there: the principal has no enforceable action (i.e., the set D^0 is a singleton) but the agents have private information. Then the prior p, the sets of types T^i, the sets of actions D^i and the utility functions u^i, $i = 1, \dots, n$, define a *Bayesian game* between the agents, which we denote as B. As pointed out in Forges (1986) and Myerson (1986) (see also Myerson, 1991), a possible generalization of the correlated equilibrium concept to Bayesian games is the *communication equilibrium*, which can be defined in terms of mechanisms. To be precise, a mechanism $\mathcal{M} = (R, M, \gamma)$ defines a *communication device* for the Bayesian game B and a communication equilibrium is a Nash equilibrium of the extended game $G(\mathcal{M})$, which is obtained by adding the communication device $\mathcal{M}$ to B. In this framework, a direct mechanism π can be called a *canonical* communication device. In the extended game $G(\pi)$ associated with the canonical device π, the truthful and obedient strategies are meaningful. They form a Nash equilibrium of $G(\pi)$ if and only π is IC (by Definition 1). We can deduce the *canonical representation* of communication equilibria from Proposition 1, namely, in a Bayesian game, the set of all communication equilibrium outcomes coincides with the set of *canonical* communication equilibrium outcomes, which are achieved with truthful and obedient strategies.

Note that the particular cases considered above, being inherited from a mechanism design problem, still specify a utility function v for the principal. To describe all correlated or communication equilibrium outcomes, the function v can be thought of as being constant. But as indicated in Myerson (1982), v can be interpreted in a strict sense, as the principal's objective function to be maximized over the set of correlated or communication equilibria, which is consistent with a focus on Corollary 1, rather than on Proposition 1.

The previous paragraphs illustrate the power of the revelation principle contained in Proposition 1

[3] This property appears in many textbooks but is stated neither in Aumann (1974) nor in Myerson (1982). As indicated above, the latter paper concentrates on Corollary 1. See Forges and Ray (this special issue of JME) for further comments on correlated equilibrium.

and Corollary 1 when there is a single principal, in the context of both mechanism design and communication in games. The result can also be formulated in the context of optimal *information design*. To see this, let us assume that the principal has no enforceable action (i.e., D^0 is a singleton) and that a special agent, say agent 1, represents nature: agent 1 is fully informed, has no private action (i.e., D^1 is a singleton) and his utility function is constant. Assume further that the other agents $i = 2, \ldots, n$ have no private information (i.e., T^i is a singleton), so that we can identify T with T^1. Consider a mechanism $\mathcal{M} = (R, M, \gamma)$ in which $R^1 = T$ and R^i is a singleton for $i = 2, \ldots, n$. Assuming that nature is truthful (which is without loss of generality), the mechanism reduces to a mapping $\gamma : T \to \Delta(M)$, which allows the principal to design the information of agents $i = 2, \ldots, n$ (and ultimately influence their choices of action) so as to maximize his own expected utility. Thanks to Corollary 1, the principal's choice of an optimal information structure amounts to choosing a direct mechanism $\pi : T \to \Delta(D)$ in such a way that agents $i = 2, \ldots, n$ are obedient (see, e.g., Bergemann and Morris, 2019 for more general versions of this result when the agents have preliminary private information).[4]

2.7 Existence of M-equilibrium

The best IC mechanism of a principal in a corporation given the outcome in the other corporation is always well-defined. Hence, if there is a single corporation, an M-equilibrium always exists: it is the optimal direct mechanism for the principal in that corporation (see Section 2.6). Unfortunately, because agents' payoffs in one corporation may depend on what happens in the other corporation, requiring principals to simultaneously obtain their best IC mechanism is not always feasible.

That is, an M-equilibrium may not always exist if there are multiple corporations, even if there is only one agent per corporation, and even if there is no private information or no private action. This existence problem is illustrated in the following two examples. The first example is a minimal example with no private information, one agent per corporation, and two actions for each agent. The second example, due to Myerson (1982), has one agent per corporation and no private action.

Example 1. Every corporation k has one agent, the principals do not have any enforceable action. Every agent has two possible private actions $D_k = D_k^1 = \{a_k, b_k\}$ and no private information ($|T_k| = 1$). The utility functions of principal k (first coordinate) and the agent in corporation k (second coordinate) are represented by the following table:

a_k	$1, z_k$
b_k	$0, 0$

where

$$z_1 = \begin{cases} -1 & \text{if } d_2^1 = a_2 \\ 0 & \text{if } d_2^1 = b_2 \end{cases} \quad \text{and} \quad z_2 = \begin{cases} 0 & \text{if } d_1^1 = a_1 \\ -1 & \text{if } d_1^1 = b_1. \end{cases}$$

In this example, principal 1 is able to get his best outcome (action a_1) if and only if principal 2 gets his worst outcome (action b_2), and principal 2 is able to get his best outcome (action a_2) if and only if principal 1 also gets his best outcome (action a_1). To show that there is no M-equilibrium, assume by way of contradiction that there exists a pair of IC mechanisms $(\pi_1, \pi_2) \in \Delta(D_1) \times \Delta(D_2)$ forming an M-equilibrium. If $\pi_2(a_2) = 0$, then the best IC mechanism for principal 1 is $\pi_1(a_1) = 1$, so the best IC mechanism for principal 2 is $\pi_2(a_2) = 1$, a contradiction. If $\pi_2(a_2) > 0$, then the unique IC mechanism for principal 1 is $\pi_1(a_1) = 0$, so the mechanism π_2 for principal 2 is not IC, a contradiction. $\diamond$

Example 2 (Myerson, 1982)**.** Every corporation k has one agent with two possible types, $T_k = \{\alpha_k, \beta_k\}$, and no private action ($|D_k^1| = 1$). The prior distribution of types is uniform. Every principal k has three

[4] General versions of the revelation principle have also been developed for *dynamic* mechanism design, *multistage* games (Forges, 1986, Myerson, 1986, Sugaya and Wolitzky, 2021) and *dynamic* information design (Makris and Renou, 2023).

possible enforceable actions, $D_k^0 = \{a_k, b_k, c_k\}$. The utility functions of principal k (first coordinate) and the agent in corporation k (second coordinate) are represented by the following table:

	α_k	β_k
a_k	$6, 1$	$0, z_k$
b_k	$0, z_k$	$6, 1$
c_k	$5, 0$	$5, 0$

where

$$z_1 = \begin{cases} 1 & \text{if } d_2^0 = c_2 \\ 2 & \text{otherwise} \end{cases} \quad \text{and} \quad z_2 = \begin{cases} 2 & \text{if } d_1^0 = c_1 \\ 1 & \text{otherwise.} \end{cases}$$

On the one hand, if principal 2 chooses the non-revealing mechanism $\pi_2(\alpha_2) = \pi_2(\beta_2) = c_2$, then $z_1 = 1$ with probability 1. Hence, principal 1 chooses his first-best mechanism $\pi_1(\alpha_1) = a_1$, $\pi_1(\beta_1) = b_1$, which is IC given π_2. Then $z_2 = 1$, so principal 2 deviates from the non-revealing mechanism to his first-best mechanism, a contradiction. On the other hand, if principal 2 does not choose the non-revealing mechanism $\pi_2(\alpha_2) = \pi_2(\beta_2) = c_2$, then $z_1 = 2$ with strictly positive probability. Hence, the best principal 1 can do is to choose the non-revealing mechanism $\pi_1(\alpha_1) = \pi_2(\beta_1) = c_1$. Then $z_2 = 2$, so the best principal 2 can do is to choose the non-revealing mechanism $\pi_2(\alpha_2) = \pi_2(\beta_2) = c_2$, a contradiction. We conclude that there is no M-equilibrium. $\diamond$

An M-equilibrium can be defined equivalently as an equilibrium of a generalized game (Debreu, 1952) in which a player's set of feasible strategies depends on the strategies chosen by the other player (this is actually the definition proposed by Myerson, 1982). Indeed, recalling Definition 2, an M-equilibrium is a Nash equilibrium of the 2-player generalized game $((V_1, V_2), (F_1(\cdot), F_2(\cdot)))$, where $F_k(\pi_l) \subseteq \Delta(D_k)^{T_k}$ is the set of feasible strategies of player k when the strategy of player l is π_l. Since the utility functions of the principals are linear in the mechanisms and the sets of IC mechanisms $F_k(\pi_l)$ are convex and non-empty, the existence of an equilibrium is guaranteed if the correspondences $F_k \rightrightarrows \Delta(D_k)^{T_k}$ are continuous (see, e.g., Debreu, 1952, Border, 1985, Tóbiás, 2022). In particular, if corporations are assumed to be "orthogonal," in the sense that the utility functions of the agents in corporation k do not depend on the actions in corporation l, then $F_k(\cdot)$ is constant and, therefore continuous. Actually, in that simple case, the existence of an M-equilibrium directly follows from Nash's existence theorem applied to the standard game $((V_1, V_2), (F_1(\cdot), F_2(\cdot)))$.

The correspondences of IC mechanisms are always upper hemi-continuous, but they may fail to be lower hemi-continuous. For instance, in Example 1, $F_1(\pi_2)$ is not lower hemi-continuous:

$$F_1(\pi_2) = \begin{cases} \Delta(D_1), & \text{if } \pi_2(a_2) = 0 \\ \{b_1\}, & \text{if } \pi_2(a_2) > 0. \end{cases}$$

To circumvent the possible non-existence of M-equilibria, Myerson (1982) introduces the weaker notion of quasi-equilibrium. A pair of jointly IC mechanisms (π_1, π_2) is a *quasi-equilibrium* if and only if there is a sequence of mechanisms $(\pi_1^t, \pi_2^t)_{t=1}^{\infty}$ converging to (π_1, π_2) such that π_k^t is IC given π_l^t for every k and t, and no principal k can improve upon π_k with a mechanism that is IC given π_l^t when t is sufficiently large. Clearly, an M-equilibrium is a quasi-equilibrium (by taking the constant sequence $(\pi_1^t, \pi_2^t) = (\pi_1, \pi_2)$ for every t), and the sets of M-equilibria and quasi-equilibria both reduce to the set of principal's optimal mechanisms when there is a single corporation. Proposition 4 in Myerson (1982) shows that a quasi-equilibrium always exists under our maintained assumptions (namely, the sets of actions and types are finite).

Intuitively, the idea of quasi-equilibrium is only to require that π_k is a "quasi best-response" mechanism of principal k given π_l, in the sense that π_k cannot be improved upon by another mechanism that remains IC for principal k for arbitrarily small perturbations of the mechanism of the other principal. For instance, in Example 1, the IC mechanism $(\pi_1, \pi_2) = (b_1, b_2)$ is not an M-equilibrium because principal 1 can profitably deviate to a mechanism $\tilde{\pi}_1$, such that $\tilde{\pi}_1(a_1) > 0$. In particular, the deviation to his first-best mechanism is feasible. However, if we perturb π_2 and consider the perturbed mechanism $\pi_2^t(a_2) = 1/t > 0$, then the probability that $z_1 = -1$ becomes strictly positive, which makes the deviation to $\tilde{\pi}_1$ infeasible because for every t, $\tilde{\pi}_1$ is not IC given π_2^t. More generally, in this example, a pair of mechanisms (π_1, π_2) is a quasi-equilibrium if and only if $\pi_2(a_2) = 0$.

3 Corporations' game and equilibria

In this section, we explicitly study a grand game between the principals and the agents, called the corporations' game. In this corporations' game, principals move first by simultaneously and secretly proposing a (not necessarily direct) mechanism for their corporation. We provide a definition of perfect Bayesian equilibrium (PBE) for this game and study the properties of PBE outcomes and their relationship with M-equilibria. In Section 3.1, we first define a version of PBE for the multistage game $G_k(\mathcal{M}_k, \pi_l)$ and provide conditions under which the revelation principle, as seen in Section 2.4 for Nash equilibria, can be extended to PBE.

3.1 Sequential rationality in the multistage game $G_k(\mathcal{M}_k, \pi_l)$

Whenever corporation k has more than one agent and there are both private actions and private information, the game $G_k(\mathcal{M}_k, \pi_l)$ is a sequential game in which some Nash equilibria may not necessarily satisfy sequential rationality conditions imposed by perfect Bayesian or sequential equilibrium (Kreps and Wilson, 1982; Fudenberg and Tirole, 1991). Specifically, it may be the case that some Nash equilibrium can only be supported by an action strategy that is not sequentially rational for some message off the equilibrium path. Such an action strategy would be necessary to deter another agent from deviating from his Nash equilibrium reporting strategy. This is illustrated in the following example adapted from Gerardi (2004) and Gerardi and Myerson (2007).

Example 3. Assume that corporation l is a "dummy" in the sense that the sets of actions and types in that corporation are singleton sets ($|D_l| = |T_l| = 1$). Hence, in this example, we remove the notations referring to corporation l and focus on the principal and agents in corporation k. Corporation k has two agents, whose type and action sets are given by

$$T_k^1 = \{\alpha^1, \beta^1\}, \quad T_k^2 = \{\alpha^2, \beta^2\},$$

$$D_k^1 = \{a, b, c\}, \quad |D_k^2| = 1.$$

The prior type distribution is given by $p_k(\alpha^1, \alpha^2) = p_k(\alpha^1, \beta^2) = p_k(\beta^1, \alpha^2) = \frac{1}{3}$, and $p_k(\beta^1, \beta^2) = 0$, and therefore it does not have full support. The utility functions of the agents are represented by the following table:

		α^2		β^2
α^1	a	$1, 0$	a	$0, 0$
	b	$-1, 0$	b	$-1, 0$
	c	$0, 2$	c	$1, 1$
β^1	a	$-1, -1$	a	$0, 0$
	b	$1, 1$	b	$1, 1$
	c	$-1, 1$	c	$0, 0$

The utility of the principal is the same as the utility of agent 1. The maximal expected utility of the principal is 1, which can only be obtained from a direct mechanism $\pi_k : T_k \to \Delta(D_k)$ satisfying:

$$\pi_k(\alpha^1, \alpha^2) = a, \quad \pi_k(\alpha^1, \beta^2) = c, \quad \pi_k(\beta^1, \alpha^2) = b.$$

($\pi_k(\beta^1, \beta^2)$ is irrelevant for the expected utility of the principal because $p_k(\beta^1, \beta^2) = 0$). Among these direct mechanisms, it is immediate to check that the only IC mechanism is:

$$\pi_k(\alpha^1, \alpha^2) = a, \quad \pi_k(\alpha^1, \beta^2) = c, \quad \pi_k(\beta^1, \alpha^2) = b, \quad \pi_k(\beta^1, \beta^2) = a, \tag{3}$$

which is represented by the grey cells in the utility table. In the game $G_k(\pi_k, \pi_l)$, the corresponding truthful and obedient Nash equilibrium relies on the "threat" of type β^1 of agent 1 to play action a when he is recommended to play a, which never happens on the equilibrium path. This threat allows preventing agent 2 from deviating in the reporting stage by reporting β^2 instead of α^2 when his type is α^2. However, a strategy that prescribes agent 1 to play action a when his type is β^1 cannot be made sequentially rational because action a is strictly dominated by action b for type β^1. Hence, the IC mechanism π_k, which is the unique IC mechanism that allows the principal to get an expected utility equal to 1, is not a PBE outcome of the game $G_k(\pi_k, \pi_l)$. $\diamond$

We show below that, under appropriate assumptions, situations like the one described in the previous example will not arise. Before that, we define formally a version of PBE in the game $G_k(\mathcal{M}_k, \pi_l)$.

Consider agent i at the action stage of $G_k(\mathcal{M}_k, \pi_l)$. Fix his information $(t_k^i, r_k^i, m_k^i) \in T_k^i \times R_k^i \otimes M_k^i$, i.e., such that $\gamma_k(m_k^i \mid r_k^i, r_k^{-i}) > 0$ for some r_k^{-i}. His belief over the types and reports of the other agents in corporation k can be described as a probability distribution

$$\mu_k^i(\cdot \mid t_k^i, r_k^i, m_k^i) \in \Delta(T_k^{-i} \times R_k^{-i}),$$

such that $\mu_k^i(r_k^{-i} \mid t_k^i, r_k^i, m_k^i) = 0$ if $\gamma_k(m_k^i \mid r_k^i, r_k^{-i}) = 0$.[5]

The belief $\mu_k^i(\cdot \mid t_k^i, r_k^i, m_k^i)$, together with the action strategies of the other agents in corporation k, σ_k^{-i}, induces a probability distribution over the other agents' types and actions, and the principal's action, in corporation k, which enables agent i to compute the expected utility $W_k^i(d_k^i; \mu_k^i, \sigma_k^{-i} \mid t_k^i, r_k^i, m_k^i)$ corresponding to a choice d_k^i. The precise expression of $W_k^i(d_k^i; \mu_k^i, \sigma_k^{-i} \mid t_k^i, r_k^i, m_k^i)$ is

$$\sum_{t_k^{-i}, r_k^{-i}} \mu_k^i(t_k^{-i}, r_k^{-i} \mid t_k^i, r_k^i, m_k^i) \sum_{d_k^0, m_k^{-i}} \gamma_k(d_k^0, m_k^{-i} \mid r_k, m_k^i) \left(\prod_{j \neq i} \sigma_k^j(d_k^j \mid t_k^j, r_k^j, m_k^j) \right) \tilde{u}_k^i(d_k, t_k; \pi_l),$$

where

$$\gamma_k(d_k^0, m_k^{-i} \mid r_k, m_k^i) = \frac{\gamma_k(d_k^0, m_k \mid r_k)}{\gamma_k(m_k^i \mid r_k)}, \text{ if } \gamma_k(m_k^i \mid r_k) > 0.$$

The action strategy σ_k^i of agent i is said to be *sequentially rational* at (t_k^i, r_k^i, m_k^i), given μ_k^i and σ_k^{-i}, if, for every $\tilde{d}_k^i$, and d_k^i in the support of $\sigma_k^i(t_k^i, r_k^i, m_k^i)$, we have:

$$W_k^i(d_k^i; \mu_k^i, \sigma_k^{-i} \mid t_k^i, r_k^i, m_k^i) \geq W_k^i(\tilde{d}_k^i; \mu_k^i, \sigma_k^{-i} \mid t_k^i, r_k^i, m_k^i).$$

The action strategy σ_k^i of agent i is *sequentially rational* given μ_k^i and σ_k^{-i} if, for every $(t_k^i$, r_k^i, $m_k^i) \in T_k^i \times R_k^i \otimes M_k^i$, it is sequentially rational at $(t_k^i$, r_k^i, $m_k^i)$, given μ_k^i and σ_k^{-i}.

[5]We could also add the restriction that $\mu_k^i(t_k^{-i}, r_k^{-i} \mid t_k^i, r_k^i, m_k^i) = 0$ if the type profile (t_k^i, t_k^{-i}) has zero prior probability (i.e., $p_k(t_k^i, t_k^{-i}) = 0$), so that type profiles with zero prior probability would never be assigned positive probability, even off the equilibrium path. Our results do not depend on whether this restriction is made or not.

Beliefs $(\mu_k^i)_i$ are said to be *consistent* with $(\rho_k^i)_i$ if every μ_k^i is obtained from p_k, γ_k, and $(\rho_k^i)_i$ by Bayes' rule whenever possible. That is, if we denote by $\mathbb{P}$ the probability distribution on $T \times R \times M$ induced by p_k, γ_k, and $(\rho_k^i)_i$, then

$$\mu_k^i(t_k^{-i}, r_k^{-i} \mid t_k^i, r_k^i, m_k^i) = \frac{\mathbb{P}(t_k^{-i}, r_k^{-i}, m_k^i \mid t_k^i, r_k^i)}{\mathbb{P}(m_k^i \mid t_k^i, r_k^i)}, \tag{4}$$

whenever $\mathbb{P}(m_k^i \mid t_k^i, r_k^i) > 0$, where

$$\mathbb{P}(m_k^i \mid t_k^i, r_k^i) = \sum_{t_k^{-i}} p_k(t_k^{-i} \mid t_k^i) \sum_{r_k^{-i}} \left(\prod_{j \neq i} \rho_k^j(r_k^j \mid t_k^j) \right) \gamma_k(m_k^i \mid r_k^i, r_k^{-i}). \tag{5}$$

As an illustration, consider again Example 3, with the direct IC mechanism $\gamma_k = \pi_k$ of Equation (3) that maximizes the utility of the principal, and consider the truthful and obedient strategy profile $(\rho_k^1, \rho_k^1, \sigma_k^2)$ in the direct game $G_k(\pi_k, \pi_l)$. Then, in the action stage, the information state of agent 1 at $t_k^1 = \beta^1$, $r_k^1 = \alpha^1$ or $r_k^1 = \beta^1$, and $m_k^1 = a$ is off the equilibrium path, and Bayes' rule cannot be applied to determine $\mu_k^i(t_k^2, r_k^2 \mid t_k^1 = \beta^1, r_k^1, m_k^1 = a)$ in Equation (4) because the denominator is 0:

$$\begin{aligned}\mathbb{P}(m_k^1 = a \mid t_k^1 = \beta^1, r_k^1) &= \sum_{t_k^2} p_k(t_k^2 \mid \beta^1) \sum_{r_k^2} \rho_k^2(r_k^2 \mid t_k^2) \gamma_k(a \mid r_k^1, r_k^2) \\ &= \underbrace{p_k(t^2 \mid \beta^1)}_{1} \underbrace{\gamma_k(a \mid r_k^1, \alpha^2)}_{0} + \underbrace{p_k(w^2 \mid \beta^1)}_{0} \underbrace{\gamma_k(a \mid r_k^1, \beta^2)}_{1} = 0.\end{aligned}$$

At this information state, regardless of the off-path belief $\mu_k^i(t_k^2, r_k^2 \mid t_k^1 = \beta^1, r_k^1, m_k^1 = a)$, action a is not sequentially rational for agent 1 because we have

$$W_k^1(a; \mu_k^1 \mid \beta^1, r_k^1, a) \leq 0 < W_k^1(b; \mu_k^1 \mid \beta^1, r_k^1, a) = 1.$$

Definition 4. *A profile of participation strategies $(\rho_k^i, \sigma_k^i)_i$ in corporation k is a PBE of $G_k(\mathcal{M}_k, \pi_l)$ iff it is a Nash equilibrium of $G_k(\mathcal{M}_k, \pi_l)$, and there exist beliefs $(\mu_k^i)_i$ that are consistent with $(\rho_k^i)_i$, such that the action strategy σ_k^i of every agent i is sequentially rational given μ_k^i and σ_k^{-i}.*

Let us go on with possible remedies to the difficulties raised in Example 3. Clearly, if agents in corporation k have no private action, or if there is a single agent in corporation k, then PBE and Nash equilibrium outcomes of $G_k(\mathcal{M}_k, \pi_l)$ coincide.

Consider next the case in which there are at least two agents as well as private actions in corporation k but assume now that p_k has full support. Let π_k be a direct mechanism that is IC for corporation k given π_l. Let us show that, in the associated direct game $G_k(\pi_k, \pi_l)$, a PBE with outcome π_k is easily constructed. First, at the reporting stage, every agent is truthful. Consider then agent i at the action stage. Fix his information (t_k^i, r_k^i, d_k^i), including a recommendation d_k^i that he can possibly get, i.e., such that

$$\pi_k(d_k^i \mid r_k^i, r_k^{-i}) > 0 \text{ for some } r_k^{-i}. \tag{6}$$

Proceeding as in (4) and (5), agent i's belief can be computed by Bayes' rule provided that

$$\mathbb{P}(d_k^i \mid t_k^i, r_k^i) > 0.$$

This condition turns out to be satisfied because according to (5), if agents $j \neq i$ are truthful at the reporting stage,

$$\mathbb{P}(d_k^i \mid t_k^i, r_k^i) = \sum_{t_k^{-i}} p_k(t_k^{-i} \mid t_k^i) \pi_k(d_k^i \mid r_k^i, t_k^{-i}),$$

which is positive by (6) and the full support assumption. In other words, given his information (t_k^i, r_k^i, d_k^i) at the action stage, agent i can use Bayes' rule, which amounts to believing with probability 1 that $t_k^{-i} = r_k^{-i}$. If he himself has been truthful (i.e., $r_k^i = t_k^i$), obedience (i.e., choosing action d_k^i) is sequentially rational since it is part of a Nash equilibrium of $G_k(\pi_k, \pi_l)$, π_k being IC given π_l. Otherwise, if $r_k^i \neq t_k^i$, agent i's strategy can just be completed so as to be sequentially rational.

Let us denote as $\mathcal{P}_k(\mathcal{M}_k, \pi_l)$ the set of PBE outcomes of $G_k(\mathcal{M}_k, \pi_l)$. By definition, $\mathcal{P}_k(\mathcal{M}_k, \pi_l) \subseteq \mathcal{E}_k(\mathcal{M}_k, \pi_l)$. Hence, using the arguments above and Proposition 1, we get the following result:

Proposition 3. *Assume that p_k has full support or that there is no private action in corporation k. If the direct mechanism π_k is IC for corporation k given π_l, then π_k is a PBE outcome of $G_k(\pi_k, \pi_l)$. That is,*

$$\pi_k \in F_k(\pi_l) \Rightarrow \pi_k \in \mathcal{P}_k(\pi_k, \pi_l),$$

and

$$\bigcup_{\mathcal{M}_k} \mathcal{P}_k(\mathcal{M}_k, \pi_l) = F_k(\pi_l).$$

Stronger notions of equilibria are studied in Gerardi and Myerson (2007) in the context of communication equilibrium (i.e., without actions for the principal) with a single principal by explicitly referring to sequences of perturbed strategies as in the concept of sequential equilibrium of Kreps and Wilson (1982). They study and compare two versions of sequential equilibrium, depending on whether or not we allow the possibility of "trembles" in the mechanism $\mathcal{M}_k$ (see also Sugaya and Wolitzky, 2021 who extend the analysis to dynamic games). The previous proposition still holds if we replace the concept of PBE with such versions of sequential equilibria, following the same argument: under the assumptions of the proposition, in a direct game $G_k(\pi_k, \pi_l)$, any unilateral deviation from a truthful reporting strategy profile is not observable by the other players. However, as shown by Gerardi and Myerson (2007), Proposition 3 does not extend to games with private actions when some combinations of types have zero probability. Relatedly, Dhillon and Mertens (1996) show that, even under complete information, the revelation principle fails when the solution concept is Selten (1975)'s trembling-hand perfect equilibrium (see also, e.g., Luo, Qiao, and Sun, 2022).

3.2 Corporations' game Γ

We now consider the corporations' game between the principals and the agents, denoted by Γ. Specifically, the corporations' game Γ is a multistage game described as follows, where the choices of the principals in stage 1, and the choices of the agents in stages 3 and 5 are made simultaneously:

1. Every principal k proposes a mechanism $\mathcal{M}_k = (R_k, M_k, \gamma_k)$.[6] Every agent of corporation k observes the mechanism $\mathcal{M}_k$ proposed by his principal but does not observe the mechanism $\mathcal{M}_l$ proposed by principal l.

2. Nature selects the type profile $t \in T$ according to the prior p. Every agent i in corporation k privately learns $t_k^i \in T_k^i$.

For every corporation k:

3. Every agent i in corporation k privately sends a report $r_k^i \in R_k^i$ to his principal.

[6]For the set of mechanisms to be well-defined, assume that every principal k chooses positive integers $\bar{R}_k^i \in \mathbb{N}^*$, $\bar{M}_k^i \in \mathbb{N}^*$, $i = 1, \ldots, n_k$, and a mapping $\gamma_k : R_k \to \Delta(D_k^0 \times M_k)$, where $R_k^i = \{1, \ldots, \bar{R}_k^i\}$ and $M_k^i = \{1, \ldots, \bar{M}_k^i\}$. Then, a direct mechanism is such that $\bar{R}_k^i = |T_k^i|$, $\bar{M}_k^i = |D_k^i|$, and we identify every type in T_k^i by an integer in R_k^i and every action in D_k^i by an integer in M_k^i.

4. Action $d_k^0 \in D_k^0$ and the profile of messages $m_k \in M_k$ are drawn with probability $\gamma_k(d_k^0, m_k \mid r_k)$. Every agent i in corporation k privately observes message $m_k^i \in M_k^i$ from his principal.

5. Every agent i in corporation k chooses an action $d_k^i \in D_k^i$.

For every k, the utility of principal k is $v_k(d,t)$, and the utility of every agent i in corporation k is $u_k^i(d,t)$.

To relate PBE outcomes of the corporations' game Γ to M-equilibria, we restrict attention to pure strategies for the principals. However, we allow mixed strategies for the agents. Hence, in the game Γ, a strategy for principal k is simply a – not necessarily direct – mechanism $\mathcal{M}_k$. A strategy for agent i in corporation k specifies a participation strategy (σ_k^i, ρ_k^i) (as defined in Section 2.4), one for each possible mechanism of principal k.

3.3 Equilibrium outcomes of Γ

In this section, we begin by defining a version of PBE for the corporations' game Γ. Subsection 3.1 shows that sequential rationality can be handled in a rather straightforward way in the continuation games played from stage 3 on. In the game Γ, the new issue is to take care of the reaction of the agents of corporation k when they observe that principal k has deviated from his equilibrium mechanism (denoted as $\mathcal{M}_k^*$), by choosing instead some other mechanism $\mathcal{M}_k$. If there is a single effective corporation k, as in Subsection 2.6, sub-game perfectness (Selten, 1965) is appropriate to avoid that agents punish their principal with *any* profile of participation strategies in the ensuing sub-game. However, there is no proper sub-game in Γ as soon as it involves two active corporations: having observed a deviation of their principal to $\mathcal{M}_k$, the agents of corporation k interact in $G_k(\mathcal{M}_k, \phi_l^*)$, which depends on the outcome ϕ_l^* in corporation l.

To address the previous issue, our definition of PBE relies on a version of the belief consistency principle referred to as "no-signaling-what-you-don't-know" (Fudenberg and Tirole, 1991). In our setting, this principle is formulated as follows: if principal k deviates to an off-path mechanism $\mathcal{M}_k$, all agents in corporation k believe that principal l did not deviate from his equilibrium mechanism. Additionally, they believe that all agents in corporation l have the same type, report, and action distributions as prescribed on the equilibrium path. The rationale is that, for each k, the mechanism $\mathcal{M}_k$ is secretly observed by agents in corporation k, and principal k has the same information as his agents about what happens in corporation l. This consistency condition implies that, given a candidate equilibrium strategy profile of Γ in which the pair of mechanisms $(\mathcal{M}_1^*, \mathcal{M}_2^*)$ is supposed to be chosen by the principals, and given the outcomes (ϕ_1^*, ϕ_2^*) induced by $(\mathcal{M}_1^*, \mathcal{M}_2^*)$ and agents' participation strategies given $(\mathcal{M}_1^*, \mathcal{M}_2^*)$, if principal k deviates to some mechanism $\mathcal{M}_k \neq \mathcal{M}_k^*$, then the participation strategies of agents in corporation k given $\mathcal{M}_k$ should constitute a PBE of the game $G_k(\mathcal{M}_k, \phi_l^*)$. Precisely:

Definition 5. *A pair of outcomes* (ϕ_1^*, ϕ_2^*) *is a* PBE outcome of the corporations' game Γ *if and only if, for every k, the following properties are satisfied:*

(i) There exists a mechanism $\mathcal{M}_k^*$ *such that* $\phi_k^* \in \mathcal{P}_k(\mathcal{M}_k^*, \phi_l^*)$*;*

(ii) For every mechanism $\mathcal{M}_k$*, there exists an outcome* $\phi_k \in \mathcal{P}_k(\mathcal{M}_k, \phi_l^*)$ *in corporation k such that*

$$V_k(\phi_k^*, \phi_l^*) \geq V_k(\phi_k, \phi_l^*). \tag{7}$$

We denote by $\mathsf{PBE} \subseteq \Delta(D_1)^{T_1} \times \Delta(D_2)^{T_2}$ the set of PBE outcomes of Γ. Under the assumption of Proposition 3 for each corporation k, the next proposition shows that every M-equilibrium (π_1^*, π_2^*) (as defined in Definition 2) is a PBE outcome of the corporations' game Γ (as defined in Definition 5).

Moreover, as demonstrated in the proof, the PBE outcome (π_1^*, π_2^*) can be induced by a strategy profile such that the principals choose the direct mechanisms (π_1^*, π_2^*), and all agents adopt truthful and obedient participation strategies on the equilibrium path.

Proposition 4. *Assume that for each corporation k, p_k has full support, or that there is no private action in corporation k. If (π_1^*, π_2^*) is an M-equilibrium, then (π_1^*, π_2^*) is a PBE outcome of the corporations' game Γ. That is,*

$$\mathsf{ME} \subseteq \mathsf{PBE}\,.$$

Proof. Let (π_1^*, π_2^*) be an M-equilibrium. We argue that there is a PBE of Γ in which for every k, principal k chooses the direct mechanism π_k^*, and the participation strategies of agents in corporation k (given π_l^*) are truthful and obedient, so that the outcome (π_1^*, π_2^*) is induced as a PBE outcome.

First, observe that condition (i) of Definition 5 is satisfied because, from Proposition 3 and the fact that $\pi_k^* \in F_k(\pi_l^*)$, we have $\pi_k^* \in \mathcal{P}_k(\pi_k^*, \pi_l^*)$, and the PBE outcome π_k^* of $G_k(\pi_k^*, \pi_l^*)$ can be induced by truthful and obedient participation strategies in corporation k.

The second condition (ii) follows from Proposition 2: because (π_1^*, π_2^*) is also a robust equilibrium according to Definition 3, for every $\mathcal{M}_k$ and $\phi_k \in \mathcal{E}_k(\mathcal{M}_k, \pi_l^*)$, we have $V_k(\pi_k^*, \pi_l^*) \geq V_k(\phi_k, \pi_l^*)$. Hence, since $\mathcal{P}_k(\mathcal{M}_k, \pi_l^*) \subseteq \mathcal{E}_k(\mathcal{M}_k, \pi_l^*)$, for every $\mathcal{M}_k$, there exists $\phi_k \in \mathcal{P}_k(\mathcal{M}_k, \pi_l^*)$ such that $V_k(\pi_k^*, \pi_l^*) \geq V_k(\phi_k, \pi_l^*)$. ■

If, for some corporation k, there are private actions ($|D_k^i| > 1$ for some agent i in corporation k) and p_k does not have full support, then an M-equilibrium (π_1^*, π_2^*) may not be a PBE outcome of Γ. The reason is that a direct IC mechanism for corporation k, π_k^* in $F_k(\pi_l^*)$, may not be a PBE outcome of $G_k(\mathcal{M}_k, \pi_l^*)$ regardless of the mechanism $\mathcal{M}_k$ used by principal k, as we previously saw in Example 3. That is, an M-equilibrium may fail condition (i) in Definition 5.

In general, the *converse* of Proposition 4 is not true, *even with a single active corporation* (see Subsection 2.6). As a trivial example, consider a single corporation, with a single agent, who is indifferent between his two actions, denoted as a and b, while the principal strictly prefers action a to action b. Then, action a is played with probability one at the unique M-equilibrium (i.e., optimal mechanism), but every outcome is a PBE outcome of the corresponding corporation's game.

As observed in Examples 1 and 2 in Section 2.7, an M-equilibrium does not always exist if there are multiple corporations. However, in these examples, the sets of PBE outcomes are non-empty. For instance, in Example 1, there is a PBE in which for every corporation k, the agent in that corporation chooses action b_k regardless of the mechanism proposed by his principal. Hence, the pair of mechanisms (π_1, π_2), with $\pi_1(b_1) = \pi_2(b_2) = 1$, is a PBE outcome. In the context of Example 2, Attar et al. (2023) also show that a PBE exists, although they consider a variant of our corporations' game (named a "competing hierarchies" game) in which the mechanisms are publicly observed by the agents in *both* corporations (see Section 3.5 for more discussion about such a variant of the corporations' game). It is readily observed that the non-revealing outcome identified by Attar et al. (2023), (π_1, π_2) with $\pi_k(c_k \mid \alpha_k) = \pi_k(c_k \mid \beta_k) = 1$ for $k = 1, 2$, is also a PBE outcome in our corporations' game Γ with privately observed mechanisms.

These examples suggest that, under suitable assumptions, a PBE could exist in the corporations' game, thus providing a solution to the possible non-existence of an M-equilibrium. However, we did not find any general existence result that would apply to the PBE of the corporations' games (in which, in particular, the principals are restricted to pure strategies). There is a simple special case in which at least one PBE always exists: when the principals have no enforceable action. In that case, there is a "no communication" PBE in which, regardless of the mechanisms proposed by the principals, the agents do not reveal any information to their principal (i.e., use constant reporting strategies) and ignore their principal's messages (i.e., choose their action as a function of their type only). The corresponding PBE outcome of Γ is then a Nash equilibrium outcome of the Bayesian game B between the agents, which always exists because B is a finite $(n_k + n_l)$-player Bayesian game. Therefore, we get the following proposition:

Proposition 5. *Assume that principals have no enforceable action. Then, the corporations' game* Γ *has at least one PBE.*

While this proposition only applies to the specific class of problems in which principals do not have direct control over actions, it is worth mentioning that, even in this class, an M-equilibrium does not necessarily exist (recall Example 1).

Relatedly, another open question remains: Is a quasi-equilibrium always a PBE outcome in the corporations' game? If this were the case, one could infer, using the existence result of quasi-equilibrium by Myerson (1982), that a PBE also exists. This is left for future research.

3.4 Direct Corporations' Game

Under suitable conditions (see Proposition 3), the revelation principle implies that in the corporations' game Γ, one principal's best response to the other's mechanism is a direct and incentive-compatible one. Given this, it seems reasonable to limit the principals' choices to direct mechanisms in the corporations' game Γ, without losing generality. While this holds true on the equilibrium path, deviating to a more general mechanism could potentially allow a principal to exert more influence over his agents. This concern is well-documented in the literature on competing mechanisms and motivates the concept of strongly robust equilibrium (see Sections 2.5 and 3.5).

Consider the *direct corporations' game* Γ_D, the multistage game that follows the same stages as game Γ, with the exception that in stage 1, each principal can only propose direct mechanisms. The definition of a PBE outcome (i.e., Definition 5) applies, *mutatis mutandis*, to Γ_D. We denote by $\mathsf{PBE_D} \subseteq \Delta(D_1)^{T_1} \times \Delta(D_2)^{T_2}$ the set of PBE outcomes of Γ_D.

Proposition 6. *Assume that, in every corporation* k, p_k *has full support or there is no private action. Then, every PBE outcome of* Γ *is a PBE outcome of* Γ_D. *That is,* $\mathsf{PBE} \subseteq \mathsf{PBE_D}$.

Proposition 6 directly follows from the revelation principle (Proposition 3) and the observation that if a principal cannot profitably deviate to a generalized mechanism, then *a fortiori* he cannot benefit from a deviation to a direct mechanism either. Bringing together the results in Propositions 4 and 6, under full support or no private action, we have that $\mathsf{ME} \subseteq \mathsf{PBE} \subseteq \mathsf{PBE_D}$.

Reversing the inclusion in Proposition 6 is much more delicate. To understand the difficulty, consider a situation in which a PBE outcome (π_k^*, π_l^*) of Γ_D is not an M-equilibrium, and suppose that there exists a generalized mechanism $\mathcal{M}_k$ such that the equilibrium outcome of $G_k(\mathcal{M}_k, \pi_l^*)$ is unique, given by π_k. By definition, because (π_k^*, π_l^*) is a PBE outcome of Γ_D, there exists an equilibrium outcome $\phi_k \in \mathcal{P}_k(\pi_k, \pi_l^*)$ such that $V_k(\phi_k, \pi_l^*) \leq V_k(\pi_k^*, \pi_l^*)$. Since ϕ_k is not necessarily the outcome of a truthful and obedient equilibrium of $G_k(\pi_k, \pi_l^*)$, it may differ from π_k. Hence, we might have $V_k(\pi_k, \pi_l^*) > V_k(\phi_k, \pi_l^*)$, and therefore, we cannot exclude that $V_k(\pi_k, \pi_l^*) > V_k(\pi_k^*, \pi_l^*)$, making the deviation to the generalized mechanism $\mathcal{M}_k$ profitable for principal k. It is worth noting that the above difficulty is not specific to the corporations model but also arises in the setting with a single principal.

Below, we identify two instances in which the inclusion of Proposition 6 can be reversed, and thus, the set of PBE outcomes of the direct corporations' game Γ_D coincides with the set of PBE outcomes of the corporations' game Γ: when in every corporation k, either there is a single agent, or p_k has full support and there is no enforceable action. The key step in establishing this result is the following lemma, which identifies, for every mechanism $\mathcal{M}_k$, a direct mechanism π_k that converts *every* equilibrium outcome of $G_k(\pi_k, \pi_l)$ into an equilibrium outcome of $G_k(\mathcal{M}_k, \pi_l)$.

Lemma 1. *Assume that in corporation k, there is a single agent or no enforceable action. Then, for every direct mechanism π_l of corporation l and for every mechanism $\mathcal{M}_k$ of corporation k, there exists a direct mechanism π_k such that $\mathcal{P}_k(\pi_k, \pi_l) \subseteq \mathcal{P}_k(\mathcal{M}_k, \pi_l)$.*

Proof. See the Appendix. ■

Using Lemma 1, we immediately obtain the following proposition:

Proposition 7. *Assume that, in every corporation, there is a single agent or no enforceable action. Then, every PBE outcome of Γ_D is a PBE outcome of Γ. That is, $\mathsf{PBE}_D \subseteq \mathsf{PBE}$.*

Under the assumption of Proposition 7, if a pair of direct mechanisms (together with the agents' continuation strategies) constitutes a PBE in the direct corporations' game Γ_D, then that same pair of direct mechanisms (alongside the same continuation strategies of the agents) also constitutes a PBE in the corporations' game Γ. Consequently, equilibrium direct mechanisms are robust to the availability of generalized mechanisms.

Although we do not have a counterexample, drawing from insights in the literature on full implementation and adversarial equilibrium selection in mechanism and information design (see, for example, Maskin, 1999, Mathevet, Perego, and Taneva, 2020, Halac, Lipnowski, and Rappoport, 2022, Morris, Oyama, and Takahashi, 2023, and references therein), we conjecture that the result in Lemma 1 (and therefore, Proposition 7) does not generalize if a corporation has multiple agents and enforceable actions, even if there is only one principal.

Combining Propositions 6 and 7, we finally obtain the following form of the "revelation principle" for corporations' games.

Corollary 2. *Assume that, in every corporation k, there is a single agent, or p_k has full support and there is no enforceable action. Then, the set of PBE outcomes of Γ coincides with the set of PBE outcomes of Γ_D. That is, $\mathsf{PBE} = \mathsf{PBE_D}$.*

3.5 Public mechanisms

In the corporations' game between the principals and the agents (Section 3.2), the agents in corporation k secretly observe the mechanism proposed by their principal. One may wonder whether Proposition 4, which connects M-equilibria to the PBE of this game, still holds if the mechanisms are publicly observed. In this section, we consider the *public corporations' game*, which differs from the game defined in Section 3.2 solely in that, at stage 1, every agent observes the mechanisms proposed by *all* principals.

In the following example, with one agent per corporation (thus with p_k having full support) and no private action, we construct an M-equilibrium that is not a PBE outcome of the public corporations' game. Hence, Proposition 4 does not hold if mechanisms are publicly observed.

Example 4. Every corporation k has one agent with no private action. The agent in corporation 1 is a dummy; he has no private information. The agent in corporation 2 has two equally likely types, $T_2 = \{\alpha_2, \beta_2\}$. Every principal k has two enforceable actions: $D_k = D_k^0 = \{a_k, b_k\}$. The utility functions of principal 2 and the agent in corporation 1 are constant. The utility functions of principal 1 (first coordinate) and the agent in corporation 2 (second coordinate) are represented by the following tables:

a_1

	α_2	β_2
a_2	$0, 1$	$1, 0$
b_2	$1, 0$	$0, 1$

b_1

	α_2	β_2
a_2	$0, 0$	$1, 1$
b_2	$1, 1$	$0, 0$

Consider the pair of direct mechanisms (π_1, π_2) defined by $\pi_1 = a_1$, $\pi_2(\alpha_2) = a_2$, and $\pi_2(\beta_2) = b_2$. These mechanisms are jointly IC because the agent in corporation 1 is indifferent between all outcomes, and the agent in corporation 2 gets his highest payoff. In addition, they form an M-equilibrium because, given π_2, principal 1 gets the same payoff (i.e., 0) regardless of his mechanism, and principal 2 is indifferent between all outcomes. However, it is not a PBE outcome of the public corporations' game: if principal 1 deviates to the mechanism $\pi_1 = b_1$, then the agent in corporation 2 is not truthful anymore after observing the deviation; he reports β_2 if his type is α_2, and α_2 if his type is β_2. This gives a strictly higher payoff to principal 1. $\diamond$

Attar et al. (2018), identify a model in which a result in the spirit of Proposition 4 holds for public mechanisms. They define a (direct) competing-mechanism game, in which principals publicly post (direct) mechanisms in order to attract the exclusive participation of agents with no private action (except for choosing a principal). In this model, restricting to direct mechanisms entails a loss of generality. However, strongly robust (subgame perfect) equilibrium outcomes of the direct game can be achieved as (subgame perfect) equilibrium outcomes of the competing-mechanism game.

The next example highlights that some PBE outcomes in the public corporations' game may not be PBE outcomes in the modified version of this game in which principals are restricted to direct mechanisms, even under conditions of full support and no private action. In other words, Proposition 6 does not hold if mechanisms are publicly observed.

Example 5. Every corporation k has one dummy agent with no private action and no private information. Every principal k has two enforceable actions: $D_k = D_k^0 = \{a_k, b_k\}$. The utility functions of principal 1 (first coordinate) and principal 2 (second coordinate) are represented by the following table:

	a_2	b_2
a_1	$2, 2$	$0, 3$
b_1	$3, 0$	$1, 1$

The game between the principals is a prisoner's dilemma, and an M-equilibrium is simply a Nash equilibrium. If the principals are restricted to direct mechanisms, the only PBE outcome of the direct corporations' game is the Nash equilibrium of the prisoner's dilemma.

Hence, the unique M-equilibrium, and the unique PBE outcome of the direct (public or private) corporations' game are given by $\pi_1 = b_1$, $\pi_2 = b_2$. Consider now the public corporations' game, and for every principal k, consider the (generalized) mechanism $\mathcal{M}_k = (R_k, M_k, \gamma_k)$, with $R_k = \{r^*, \tilde{r}\}$, $|M_k| = 1$, and

$$\gamma_k(r_k) = \begin{cases} a_k & \text{if } r_k = r^* \\ b_k & \text{if } r_k = \tilde{r} \end{cases}$$

For every corporation k, consider the following reporting strategy for the agent in that corporation: he reports r^* if the mechanism proposed by principal l is $\mathcal{M}_l$, and he reports $\tilde{r}$ otherwise. The induced outcome is $\pi_1 = a_1$, $\pi_2 = a_2$. Sequential rationality of the agents is satisfied because they are indifferent between all outcomes. In addition, if some principal k deviates by proposing another mechanism than $\mathcal{M}_k$, then his highest payoff is 1, which is lower than the payoff he gets by not deviating (i.e., 2). Hence, in this example, there is a PBE outcome of the public corporations' game that is different from (and Pareto dominates) the unique M-equilibrium and the unique PBE outcome of the direct (public or private) corporations' game. $\diamond$

In the previous example, an agent in corporation k, who observes the mechanism chosen by the principal of the other corporation l, can report a deviation of principal l to his principal k. Since principal k commits to an enforceable action as a function of his agent's report, a generalized mechanism allows him to commit to a given action based on the mechanism proposed by the other principal. In this example, a generalized mechanism enables each principal to indirectly commit to some action, conditionally on the action induced by the other principal, as in the contract games of, e.g., Myerson (1991, Section 6.1) and Kalai, Kalai, Lehrer, and Samet (2010), and the competing-mechanism games of, e.g. Yamashita (2010), Peters and Szentes (2012), and Peters and Troncoso-Valverde (2013).

References

Andrea Attar, Eloisa Campioni, and Gwenaël Piaser. On competing mechanisms under exclusive competition. *Games and Economic Behavior*, 111:1–15, 2018.

Andrea Attar, Eloisa Campioni, and Gwenaël Piaser. Equilibrium (non-) existence in games with competing principals. *Economics Letters*, 222:110932, 2023.

Robert J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

Kim C Border. *Fixed point theorems with applications to economics and game theory*. Cambridge university press, 1985.

Partha Dasgupta, Peter Hammond, and Eric Maskin. The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, 46(2):185–216, 1979.

Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893, 1952.

Amrita Dhillon and Jean François Mertens. Perfect correlated equilibria. *Journal of Economic Theory*, 68(2):279–302, 1996.

Larry G Epstein and Michael Peters. A revelation principle for competing mechanisms. *Journal of Economic Theory*, 88(1):119–160, 1999.

Françoise Forges. An approach to communication equilibria. *Econometrica*, 54(6):1375–1385, 1986.

Drew Fudenberg and Jean Tirole. Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53:236–260, 1991.

Dino Gerardi. Unmediated communication in games with complete and incomplete information. *Journal of Economic Theory*, 114(1):104–131, 2004.

Dino Gerardi and Roger B Myerson. Sequential equilibria in bayesian games with communication. *Games and Economic Behavior*, 60(1):104–134, 2007.

Marina Halac, Elliot Lipnowski, and Daniel Rappoport. Addressing strategic uncertainty with incentives and information. *AEA Papers and Proceedings*, 112:431–437, 2022.

Seungjin Han. Strongly robust equilibrium and competing-mechanism games. *Journal of Economic Theory*, 137(1):610–626, 2007.

John C Harsanyi. Games with incomplete information played by bayesian players. Parts i, ii, iii. *Management Science*, 14:159–182, 320–334, 486–502, 1967–1968.

Bengt Robert Holmström. On incentives and control in organizations. PhD dissertation, Stanford University, 1977.

Adam Tauman Kalai, Ehud Kalai, Ehud Lehrer, and Dov Samet. A commitment folk theorem. *Games and Economic Behavior*, 69(1):127–137, 2010.

Michael L Katz. Game-playing agents: Unobservable contracts as precommitments. *The RAND Journal of Economics*, 22(3):307–328, 1991.

David M Kreps and Robert Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.

Xiao Luo, Yongchuan Qiao, and Yang Sun. A revelation principle for correlated equilibrium under trembling-hand perfection. *Journal of Economic Theory*, 200:105396, 2022.

Miltiadis Makris and Ludovic Renou. Information design in multistage games. *Theoretical Economics*, 18(4):1475–1509, 2023.

David Martimort. The multiprincipal nature of the government. *European Economic Review*, 40: 673–685, 1996.

David Martimort. Multi-contracting mechanism design. In R Blundell, W. Newey, and T. Person, editors, *Advances in Economic Theory, Proceedings of the 2005 World Congress of the Econometric Society*, pages 56–101. Cambridge University Press, 2007.

David Martimort and Lars Stole. The revelation and delegation principles in common agency games. *Econometrica*, 70(4):1659–1673, 2002.

Eric Maskin. Nash equilibrium and welfare optimality. *The Review of Economic Studies*, 66(1):23–38, 1999.

Laurent Mathevet, Jacopo Perego, and Ina Taneva. On information design in games. *Journal of Political Economy*, 128(4):1370–1404, 2020.

R Preston McAfee. Mechanism design by competing sellers. *Econometrica*, 61(6):1281–1312, 1993.

Stephen Morris, Daisuke Oyama, and Satoru Takahashi. Implementation via information design in binary-action supermodular games. *Available at SSRN 3697335*, 2023.

Roger B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47(1):61–73, 1979. ISSN 0012-9682.

Roger B Myerson. Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, 10:67–81, 1982.

Roger B Myerson. Multistage games with communication. *Econometrica*, 54:323–358, 1986.

Roger B Myerson. *Game Theory, Analysis of Conflict.* Harvard University Press, 1991.

Michael Peters. Common agency and the revelation principle. *Econometrica*, 69(5):1349–1372, 2001.

Michael Peters and Balázs Szentes. Definable and contractible contracts. *Econometrica*, 80(1):363–411, 2012.

Michael Peters and Cristián Troncoso-Valverde. A folk theorem for competing mechanisms. *Journal of Economic Theory*, 148(3):953–973, 2013.

Reinhard Selten. Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit: Teil i: Bestimmung des dynamischen preisgleichgewichts. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics*, H. 2:301–324, 1965.

Reinhard Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):25–55, 1975.

Takuo Sugaya and Alexander Wolitzky. The revelation principle in multistage games. *The Review of Economic Studies*, 88(3):1503–1540, 2021.

Áron Tóbiás. Equilibrium non-existence in generalized games. *Games and Economic Behavior*, 135: 327–337, 2022.

Takuro Yamashita. Mechanism games with multiple principals and three or more agents. *Econometrica*, 78(2):791–801, 2010.

Appendix

Proof of Lemma 1

We fix k and π_l in the proof, so we remove below the notations referring to corporation k and mechanism π_l. First, consider the case in which there is a single agent in the corporation. For every mechanism $\mathcal{M}$, let $\Pi(\mathcal{M})$ be the set of all outcomes that could be induced by some pure strategy of the agent in $G(\mathcal{M})$.

Claim 1. *For every mechanism* $\mathcal{M}$ *and outcome* $\pi \in \Pi(\mathcal{M})$*, we have* $\Pi(\pi) \subseteq \Pi(\mathcal{M})$*.*

Proof. Because there is a single agent in the corporation, we denote respectively by D^1, T, R, and M the set of actions, types, reports, and messages of the agent. Consider a mechanism $\mathcal{M} = (R, M, \gamma)$, with $\gamma : R \to \Delta(D^0 \times M)$. Consider a pure strategy (ρ, σ) for the agent in $G(\mathcal{M})$, where $\rho : T \to R$ and $\sigma : T \times M \to D^1$. Let $\pi : T \to \Delta(D^0 \times D^1)$ be the outcome induced by (ρ, σ) in $G(\mathcal{M})$. That is, for every t and $d = (d^0, d^1)$:

$$\pi(d \mid t) = \sum_{m:\sigma(t,m)=d^1} \gamma(d^0, m \mid \rho(t)). \tag{8}$$

Consider the game $G(\pi)$ and let $(\tilde{\rho}, \tilde{\sigma})$ be an arbitrary pure strategy for the agent in $G(\pi)$, where $\tilde{\rho} : T \to T$ and $\tilde{\sigma} : T \times D^1 \to D^1$. Let $\tilde{\pi} : T \to \Delta(D^0 \times D^1)$ be the outcome induced by $(\tilde{\rho}, \tilde{\sigma})$ in $G(\pi)$. That is, for every t and $d = (d^0, d^1)$:

$$\tilde{\pi}(d \mid t) = \sum_{a^1:\tilde{\sigma}(t,a^1)=d^1} \pi(d^0, a^1 \mid \tilde{\rho}(t)).$$

From (8), we have:

$$\tilde{\pi}(d \mid t) = \sum_{a^1:\tilde{\sigma}(t,a^1)=d^1} \sum_{m:\sigma(\tilde{\rho}(t),m)=a^1} \gamma(d^0, m \mid \rho \circ \tilde{\rho}(t)) = \sum_{m:\tilde{\sigma}(t,\sigma(\tilde{\rho}(t),m))=d^1} \gamma(d^0, m \mid \rho \circ \tilde{\rho}(t)). \tag{9}$$

Define the pure strategy (ρ^*, σ^*) for the agent in $G(\mathcal{M})$, where $\rho^* : T \to R$ and $\sigma^* : T \times M \to D^1$, as follows: For every $t \in T$ and $m \in M$,

$$\rho^*(t) = \rho \circ \tilde{\rho}(t),$$

$$\sigma^*(t,m) = \tilde{\sigma}(t, \sigma(\tilde{\rho}(t), m)).$$

It is immediately checked that the outcome induced by (ρ^*, σ^*) in $G(\mathcal{M})$ coincides with $\tilde{\pi}$. ■

Consider an optimal pure strategy of the agent in $G(\mathcal{M})$, and let π be the induced outcome. That is,

$$\pi \in \arg\max_{\tilde{\pi} \in \Pi(\mathcal{M})} W(\tilde{\pi}),$$

where $W(\tilde{\pi})$ is the (ex-ante) expected utility of the agent given $\tilde{\pi}$. Since $\pi \in \Pi(\pi)$, Claim 1 implies that $\pi \in \arg\max_{\tilde{\pi} \in \Pi(\pi)} W(\tilde{\pi})$, which is consistent with the revelation principle. Even more, we have that

$$\arg\max_{\tilde{\pi} \in \Pi(\pi)} W(\tilde{\pi}) \subseteq \arg\max_{\tilde{\pi} \in \Pi(\mathcal{M})} W(\tilde{\pi}). \tag{10}$$

Because there is a single agent, observe that the set of PBE outcomes in $G(\mathcal{M})$ coincides with the set of Nash equilibrium outcomes in $G(\mathcal{M})$; it is simply the set of outcomes induced by an optimal mixed strategy of the agent in $G(\mathcal{M})$. In addition, the set of outcomes induced by an optimal mixed strategy of the agent in $G(\mathcal{M})$ is the convex hull of the set of outcomes induced by an optimal pure strategy of the agent in $G(\mathcal{M})$. Hence,

$$\mathcal{P}(\mathcal{M}) = \mathcal{E}(\mathcal{M}) = \text{co}\,\{\arg\max_{\tilde{\pi} \in \Pi(\mathcal{M})} W(\tilde{\pi})\}.$$

Using (10), we conclude that $\mathcal{P}(\pi) \subseteq \mathcal{P}(\mathcal{M})$. This completes the proof of the lemma when there is only one agent.

Second, consider the case in which the principal has no enforceable action. Consider the Bayesian game B played between the agents in the corporation when there is no principal, or, equivalently, when the principal uses a constant mechanism that always sends the same message regardless of the reports of the agents. Let Π^{NC} be the set of Bayes-Nash equilibrium outcomes of B. Consider a direct mechanism $\bar{\pi}$ which is constant: the same message is sent to the agents regardless of the reports. We have

$$\mathcal{P}(\bar{\pi}) = \Pi^{NC}.$$

Now, consider any mechanism $\mathcal{M}$ and the corresponding game $G(\mathcal{M})$. Consider any Bayes-Nash equilibrium strategy profile in B. For each agent, consider in $G(\mathcal{M})$ the participation strategy such that the same report is sent with probability one regardless of the agent's type, and plays according to the previous Bayes-Nash equilibrium strategy profile in B regardless of the message received from the principal. Clearly, the induced outcome is in Π^{NC} and is a "no communication" equilibrium outcome in $G(\mathcal{M})$. That is, we have $\Pi^{NC} \subseteq \mathcal{P}(\mathcal{M})$. We conclude that $\mathcal{P}(\bar{\pi}) \subseteq \mathcal{P}(\mathcal{M})$. This completes the proof of Lemma 1.