Dynamic Contracting Competition*

Sambuddha Ghosh[†] Seungjin Han[‡]

August 21, 2021

Abstract

We model dynamic contracting competition, where multiple principals simultaneously and repeatedly offer short-term contracts to multiple agents. Since each principal's contract can take other contracts as inputs, the complexity of contracts blows up - McAfee (1993) circumvents this "infinite regress problem" by allowing only direct mechanisms. As players become patient, we show that such ad-hoc restrictions on contracts are unnecessary. We propose a slight extension of direct mechanisms that generates all equilibrium payoffs that can be obtained by arbitrarily complex mechanisms.

1 Introduction

We study a model of dynamic contracting competition. Multiple principals compete over the infinite horizon by offering non-exclusive short-term

*We thank Madhav Chandrasekher, Ani Dasgupta, Yuk-fai Fong, Ying Lei, Maxim Ivanov, Rumen Kostadinov, Bart Lipman, Dilip Mookherjee, Andy Newman, Juan Ortner, Frank Page, Larry Samuelson, Edward Schlee, Caixia Shen, Andy Skrzypacz, and Lixin Ye for comments. Seminar audiences at Arizona State University, Boston University, Concordia University, Decentralization Conference, Indiana University, Sydney University, University of Guelph, University of New South Wales gave useful suggestions.

[†]School of Economics, Shanghai University of Finance and Economics, China Email: sghosh@mail.shufe.edu.cn

[‡]Department of Economics, McMaster University, Canada. Email: hansj@mcmaster.ca

contracts to agents with private information. Each principal's contract is observed by agents only; in any period it specifies that principal's current action as a function of the current private messages from the agents. Thus principals can write contracts and take actions, while agents have private information that they can choose to reveal through the use of (private) messages. All players are long-lived. For ease of exposition we assume that all contracts become public knowledge at the *end* of the period in which they were offered. We are interesting in identifying the class of payoffs that can be supported in equilibrium, and the strategies needed to do so.

Our model captures a range of situations. Manufacturers offer short-term contracts in every period t to non-exclusive retailers who are better informed about the local market conditions. The period-t contract specifies his action (say prices charged to retailers) in that period conditional on the period-t messages from retailers about the state of demand. Intermediate good suppliers (e.g., Intel, AMD, etc.) non-exclusively supplying processors to multiple computer manufactures; the production cost of intermediate goods depends on the situation of the underlying input (e.g., semiconductors). As a final example, consider various levels of governments making policy decisions, such as levying corporate taxes, after private communication from lobbyists and experts with better information about the state of the industry; different jurisdictions may have different objectives, introducing competition through taxation and other contracts seems natural.

The literature on dynamic mechanisms can be divided into two classes based on whether incentives are provided through explicit contracts or implicitly through repetition. Starting with Rubinstein (1979), the former considers a single principal who can offer contracts over the infinite horizon to agents. The latter relies on a shared understanding or an unwritten 'relational contract' (Levin (2003)).

A small literature pioneered by Pearce and Stacchetti (1998) and Baker, Gibbons, Murphy (1994) sits in the middle, studying the interaction of explicit and implicit contracts. Like these papers we consider short-run contracts. Since contracts are ultimately based on laws, which themselves change in response to other forces (political, sociological etc.) it seems impracticable or inefficient to write infinite-horizon contracts. Short-run contracts can be also be thought of as a simple way to permit renegotiation in a decentralized market.

Competition among principals hardly needs justification. It is unexplored compared to the single-principal version because of the complexity involved rather than the lack of relevance; for example McAfee (1993) says "... It is also quite difficult to define the large strategy space since mechanisms must map the set of mechanisms into outcomes,..." The crux is that, unlike the single-principal models, private information is partly endogenous in such models. This is because agents observe the contract or terms of trade offered by the various principals, in addition to their exogenous private information. While principals can of course write contracts that ask agents to report all their private information including the information about other principals, reporting about a complicated contract requires one that is yet more complicated. The complexity of the contracts starts to blow up quickly—this is what McAfee refers to as the 'infinite regress problem'.

We provide a simple way out of this problem. We show that all payoffs exceeding certain player-specific cutoff values can be supported in equilibrium when principals and agents are patient. Our work differs from the typical folk theorem in two ways. First, unlike in that literature and in single-principal models, the payoff cutoffs (or minmax values) in our world are not easily specified in closed form. We manage to do so by showing that the minmax value without exogenous restrictions on the complexity of mechanisms can be computed as the maxmin value of a much simpler game, where the only contracts are constant-action contracts, DMs and a slight extension of DMs free from the infinite regress problem. Second, while folk theorems are about payoffs, our results are about the complexity of mechanisms, which are not investigated by the literature of repeated games because their stage-game typically has a simple form.

1.1 Related literature

There has been a growing interest in designing dynamic mechanisms, with a flurry of research in the last decade (Bergemann and Välimäki (2010), Athey and Segal (2013), Pavan, Segal and Toikka (2014), Guo and Hörner (2018), Battaglini and Lamba (2019), Sugaya and Wolitzky (2020)). These models study a single principal (i.e., mechanism designer or mediator) offering an infinite horizon contract to one or more agents.

It is only recently that dynamic Bayesian games with multiple players have been tackled by the literature of repeated games, most relevant being Hörner, Takahashi, Vielle (2015). The object of their study is rather the limit payoff set of truthful public perfect equilibria given a fixed action space, abstracting contracting processes away from the model. We put contracting processes between multiple principals and multiple agents to the forefront of equilibrium analysis. Close to our paper is perhaps Bergemann and Välimäki (2003) in the sense that they study dynamic contracting problems. Bergemann and Välimäki (2003) considers a dynamic common agency where principals compete in reward schedules to control the single agent's action choice. Their interest differs from ours in that they focus on the uniqueness of truthful equilibrium payoffs under symmetric information.

2 An example

This sections presents the key ideas using a simple model of repeated duopoly with contracting. Some of the assumptions are made for simplicity only, as will be clear in the later parts of the paper.

At each $t \in \mathbb{N}$ two manufacturers, 1 and 2, produce differentiated nonstorable products at zero marginal cost; each period's produce is sold in wholesale markets to three or more long-lived retailers indexed by $i \in \mathcal{I}$. Manufacturers and retailers have zero reservation profits. Retailers sell the products in retail markets operating *a la* Bertrand. If the retail price of product j (i.e. produced by manufacturer j, irrespective of which retailer sells it) is $p_j \geq 0$ and the other product is priced at $p_\ell \geq 0$, then the consumer demand function for product j is

$$Q_j = D_j(p_\ell, p_j, \theta) = 1 + \theta p_\ell - (2 + \sqrt{p_\ell})p_j, \tag{1}$$

where $\theta \in \{1, 2\}$. Every player knows that the state θ, which we can interpret as ground information about demand shocks, takes the values 1 and 2 with equal probability independently across periods; however, its realization is observable to retailers only, and is never revealed to manufacturers. In the demand function for product j in (1), p_ℓ interacts with both θ and p_j. As shown below, this makes it necessary for manufacturer ℓ to know both θ and p_j (or prices charged to retailers) in order to effectively punish manufacturer ℓ.

Manufacturers and retailers maximize the discounted sum of their respective per-period profits using a common discount factor δ. We assume that there is a public correlation device available at the start of each period to permit (independent) randomizations over mechanisms.

We now describe sales in the wholesale market, from manufacturers to retailers. In each period, each manufacture j offers a one-period contract that commits him to a profile of prices $(p_{ij})_{i\in\mathcal{I}}$, one charged to each retailer, as a function of the messages from the retailers, with the understanding that the retailers can buy whatever amount they want to at the offered prices. Any such contract comprises a set M_{ij} of messages that retailer i can send to manufacturer j in his contract, with M_j defined as $\times_{i\in\mathcal{I}} M_{ij}$. Let $\gamma_j := \left\{(p^{\gamma}_{ij})_{i\in\mathcal{I}}\right\}$ denote manufacturer j's contract, where $p^{\gamma}_{ij} : M_j \to \mathbb{R}_+$ for all $i \in \mathcal{I}$. No restrictions are imposed on the complexity of the message spaces and what information it can contain. Let Γ_j be the set of all possible contracts for manufacturer j. Manufacturer j can reach zero reservation profit by offering a contract in Γ_j that specifies extremely high prices for all retailers conditional on every profile of their messages.

Contracts and actions (i.e., prices) become public information at the end of each period, but states are never revealed. Motivated by industry practices, we assume that retailers sell the product at the price that incorporates a standard industry *mark up* Δ, say $\Delta = 0.1$ (i.e., 10%), on the wholesale price.[1]

The worst punishment that can be inflicted on each retailer i is equal to zero because a retailer can be effectively excluded: Both manufacturers choose contracts in which the price charged to the punished retailer regardless of the messages sent is so high that she cannot sell any products.

How about the worst punishment that the other players can inflict on manufacturer j in a phase where he is punished for a past deviation? We need to know this to derive j's worst equilibrium payoff $\underline{w}_j$. The simplest contract that manufacturer j can offer to retailers is a fixed action $(p_{ij})_{i \in \mathcal{I}}$. A fixed action can be thought of as a constant contract, and manufacturer j cannot do worse with more complex contracts than he does with fixed actions. We shall show that manufacturer j's worst equilibrium payoff is equal to the maximum payoff he can reach with a fixed action.

To see this point, we introduce with some notation. Given a fixed action $\mathbf{p}_j := (p_{ij})_{i \in \mathcal{I}}$ of manufacturer j, the retailer (or retailers) who receives product j at the lowest price from manufacturer j supplies product j to the whole local market that operates *a la* Bertrand. Given $\mathbf{p}_j$, let $p_j^{\circ}(\mathbf{p}_j) := \min \{p_{ij} : i \in \mathcal{I}\}$.

Suppose that manufacturer ℓ uses an *extended direct mechanism* with a strict majority rule in a phase where manufacturer j is punished for his past deviation. In this class of extended direct mechanisms, retailers are asked to report j's action and the state of the local market, i.e. the value of θ. If more than a half of retailers report the same action $\mathbf{p}_j$ and state θ, manufacturer ℓ takes $\{\mathbf{p}_j, \theta\}$ as manufacturer j's true action and state.[2] For any action $\mathbf{p}_j$ that manufacturer j chooses and any state θ, the strict majority rule induces

[1]For simplicity, we take the standard industry mark up as given but this too can be incorporated as part of equilibrium outcomes.

[2]What matters in $\mathbf{p}_j = (p_{ij})_{i \in \mathcal{I}}$ is the lowest price charged to retailers. Therefore, instead of asking retailers to report $\mathbf{p}_j$, manufacturer ℓ's extended direct mechanism can ask retailers to report $p_j^{\circ}(\mathbf{p}_j) = \min \{p_{ij} : i \in \mathcal{I}\}$.

truth telling given three or more retailers.

Then, for every $\{\mathbf{p}_j, \theta\}$, the worst punishment manufacturer ℓ can inflict on j is to choose $p_{i\ell} = p_\ell^*(\mathbf{p}_j, \theta)$ for all $i \in \mathcal{I}$[3] to minimize the demand for product j, i.e.,

$$p_\ell^*(\mathbf{p}_j, \theta) \in \arg\min_{p_\ell} D_j \Big((1+\Delta)\, p_\ell, (1+\Delta)\, p_j^\circ(\mathbf{p}_j), \theta \Big).$$

The solution is unique and it is

$$p_\ell^*(\mathbf{p}_j, \theta) = \frac{1}{1+\Delta} \left(\frac{(1+\Delta)\, p_j^\circ(\mathbf{p}_j)}{2\theta} \right)^2 \text{ for all } \{\mathbf{p}_j, \theta\}.$$

Note that in his extended direct mechanism, manufacturer ℓ chooses $p_\ell^*(\mathbf{p}_j, \theta)$ when more than a half of retailers report $\{\mathbf{p}_j, \theta\}$.

This implies that if manufacturer j were restricted to offer only a fixed action, then the maximum payoff he could achieve when he was punished would be

$$\underline{w}_j := \max_{\mathbf{p}_j} \left\{ p_j^\circ(\mathbf{p}_j) \times \mathbb{E}_\theta \left[\min_{p_\ell} \; D_j((1+\Delta)\, p_\ell, (1+\Delta)\, p_j^\circ(\mathbf{p}_j), \theta) \right] \right\}$$

The definition of $p_j^\circ(\mathbf{p}_j)$ implies that

$$\begin{aligned} \underline{w}_j &= \max_{p_j} \left\{ p_j \times \mathbb{E}_\theta \left[\min_{p_\ell} \; D_j((1+\Delta)\, p_\ell, (1+\Delta)\, p_j, \theta) \right] \right\} && (2) \\ &= \max_{p_j} \left\{ p_j \times \left[1 - 2.2 p_j - \frac{3.63}{8} p_j^2 \right] \right\} && (3) \end{aligned}$$

The objective function in (3) is strictly concave in $p_j \geq 0$. The value of $\underline{w}_j$ is 0.1088 and it is reached when manufacturer j charges the price of 0.21321 to each retailer. Subsequently, the consumer price becomes 0.234531.

Since a fixed action can be thought of as a constant contract, manufac-

[3] Since every retailer can purchase product ℓ at the same price from manufacturer ℓ, they supply equal quantities to the market for product ℓ.

turer j's payoff cannot go below $\underline{w}_j$ in a phase where he is punished when he can use any arbitrary contract in Γ_j. How much higher can it be? This is a hard question to answer directly as we cannot simply list all possible contracts. Our Theorems 1 and 2 show that even when we do not impose any exogenous restrictions on the complexity of mechanisms, manufacturer j cannot do any better using non-constant mechanisms when he is punished. Thus the complexity of contracts is significantly pared down.

Using this example we explain some of the key ideas that permit such a simplification. Suppose that manufacturer j offers any arbitrary contract when he is being punished. Given such a contract $\gamma_j = \left\{ (p_{ij}^{\gamma})_{i \in \mathcal{I}} \right\}$, let

$$\gamma_j(M_j) := \left\{ (p_{ij}^{\gamma}(m_j))_{i \in \mathcal{I}} : m_j \in M_j \right\}$$

be the image set of γ_j, so it includes all actions (price vectors) that can be induced from γ_j using all permitted messages. Now suppose that retailers are instructed to always send the messages $m_j^{\times} = \left(m_{ij}^{\times} \right)_{i \in \mathcal{I}}$ such that it induces $(p_{ij}^{\gamma}\left(m_j^{\times} \right))_{i \in \mathcal{I}} = \mathbf{p}_j^{\times}(\gamma_j)$, where

$$\mathbf{p}_j^{\times}(\gamma_j) \in \underset{\mathbf{p}_j \in \gamma_j(M_j)}{\arg\min} \left\{ p_j^{\circ}(\mathbf{p}_j) \times \mathbb{E}_{\theta} \left[\min_{p_\ell} D_j((1+\Delta)\, p_\ell, (1+\Delta)\, p_j^{\circ}(\mathbf{p}_j), \theta) \right] \right\}.$$

Such communication behavior *always* induces $\mathbf{p}_j^{\times}(\gamma_j)$ from γ_j, so for manufacturer j, offering contract γ_j to retailers is equivalent to offering fixed action $\mathbf{p}_j^{\times}(\gamma_j)$. All retailers report $\mathbf{p}_j^{\times}(\gamma_j)$ and the true state θ to manufacturer ℓ whose extended direct mechanism then chooses $p_{i\ell} = p_\ell^*(\mathbf{p}_j^{\times}(\gamma_j), \theta)$ for all $i \in \mathcal{I}$. This implies that retailers *block any information transmission* to manufacturer j and that they completely neutralize the effectiveness of more complex contracts than fixed actions. In this way, manufacturer j's payoff can be lowered to $\underline{w}_j$ in a phase where he is punished even if he can use any complex contracts. Importantantly retailers do not need to describe γ_j directly but only j's action to manufacturer ℓ, so extened DMs are free from the infinite regress problem.

It only remains to answer whether such communication is enforceable. The answer is yes if retailers are sufficiently patient ($\delta \to 1$). Suppose that an unexpected action, i.e. one other than $\mathbf{p}_j^{\times}(\gamma_j)$, is observed at the end of the period given γ_j. Manufacturers then know that at least one retailer has deviated from such communication behavior even though they may not know the identity of the deviating retailer. With a continuous public correlation device, manufacturers can punish each retailer with equal probability: They choose contracts in which the price charged to the chosen retailer is so high that she cannot sell any products but reasonable prices are charged to the other retailers. This suffices to deter any retailer's deviation in a phase where a manufacturer is punished because a retailer faces a positive probability that she loses the opportunity of making positive profits based on mark-up $\Delta > 0$.

We described the worst punishment that can be inflicted on each player when he or she deviates. Profiles of incentive compatible direct mechanisms suffice in the phases where no player has previously deviates or an agent is punished.

As a last note, the incentive compatibility over the states in this example is straightforward because the state is fully observable by all agents. If each agent has only a partial signal, called her type, on the state, the incentive compatibility over agents' types in the direct mechanisms and extended direct mechanisms has to be carefully imposed. We shall show that as $\delta \to 1$, repetition allows us to use a weaker notion of incentive compatibility than one used in one-shot games because an agent's punishment can be deferred in some cases.

3 Preliminaries

We first describe the underlying game. The sets of principals and agents are, respectively, $\mathcal{J} := \{1, \dots, J\}$ with $J \geq 2$ and $\mathcal{I} := \{J+1, \dots, J+I\}$. For the most part, we assume that $I \geq 3$ to simplify agents' incentives; its role will be clarified later. Under a weak condition such incentives can be also be provided when there are only two agents.[4]

Each principal j takes an action a_j from a finite[5] set A_j. The nature of actions depends on a specific application we consider. In the example with manufacturers and retailers, principal (manufacturer) j's action a_j is $a_j = (p_{ij})_{i\in\mathcal{I}}$, a profile of prices charged to retailers (agents). In the problem with final good producers (principals) and intermediate good suppliers (agents) , final good producer j's action is $a_j = [(t_{ij})_{i\in\mathcal{I}}, (q_{ij})_{i\in\mathcal{I}}]$, a profile of payments and quantities, one for each supplier i. In the problem with policy makers (principals) and interest groups (agents), policy maker j's action a_j is his policy decision.

A random action of principal j is denoted by $\alpha_j \in \mathcal{A}_j := \triangle\left(A_j\right)$. A profile of random actions is $\alpha = (\alpha_1, \ldots, \alpha_J) \in \mathcal{A} := \times_{j\in\mathcal{J}}\mathcal{A}_j$ and $\mathcal{A}_{-j} := \times_{k\neq j}\mathcal{A}_k$.

In the example with manufacturers and retailers, the state of the local market is perfectly known to all retailers. This may be too strong. We allow that each agent i has a partial signal θ_i, called her type about the state. Each agent i is privately informed about her type which is drawn from a finite set Θ_i according to a distribution μ_i; the profile of types $\theta = \left(\theta_{J+1}, \ldots, \theta_{J+I}\right)$ is drawn from $\Theta := \times_{i\in\mathcal{I}}\Theta_i$ according to the joint distribution $\mu = \times_{i\in\mathcal{I}}\mu_i$.[6] In the example with manufacturers and retailers, θ_i is the private information that retailer i knows about the state about the local market. In the problem with final good producers and intermediate good suppliers, θ_i is the private information that supplier i knows about the state on the market for inputs for intermediate goods. In the problem with policy makers and interest groups, θ_i is the private information that interest group i knows about the state that relevant in making policy decisions.

The vN-M (von Neumann-Morgenstern) expected utility function for player ℓ (principal or agent) is $u_\ell : \mathcal{A} \times \Theta \to \mathbb{R}$; payoffs are uniformly

[4]See Section 5.1 for details.

[5]Finiteness of the type and action spaces is not critical for our results. With a modicum of technicalities we can deal with a compact set of actions and a countable type-space.

[6]If types are perfectly correlated (i.e., $\theta_i = \theta_{i'}$ for all $i, i' \in \mathcal{I}$), then it is the case of complete information among agents.

bounded by $\overline{u} < \infty$, i.e. $|u_\ell(\alpha, \theta)| < \overline{u}$ for all $\alpha \in \mathcal{A}$, all $\ell \in \mathcal{I} \cup \mathcal{J}$, and all $\theta \in \Theta$. All this information is encapsulated in the underlying game:

$$G := \left(\mathcal{J}; \mathcal{I}; (A_j)_{j \in \mathcal{J}}; (\Theta_i)_{i \in \mathcal{I}}; (\mu_i)_{i \in \mathcal{I}}; (u_\ell)_{\ell \in \mathcal{I} \cup \mathcal{J}}\right). \tag{4}$$

Stage game Fix an underlying game G as in (4), and for each $j \in \mathcal{J}$ fix a collection of sets $\{M_{ij} | i \in \mathcal{I}\}$ and a set Γ_j which comprises continuous mappings γ_j from $M_j := \times_{i \in \mathcal{I}} M_{ij}$ to $\mathcal{A}_j$, where M_{ij} is the set of messages that agent i can send to principal j. γ_j is principal j's mechanism. Γ_j is the set of mechanisms available for principal j. We assume that Γ_j and M_j are compact for all $j \in \mathcal{J}$. Note that random mechanisms are allowed. We impose no exogenous restrictions on the complexity our communication mechanisms permit: For example, mechanisms in Γ_j may allow agents to report not only their own types, but also mechanisms offered by the other principals, etc. The stage game (G, Γ) is the game with the following timing of moves:

1. Each principal j simultaneously offers a mechanism γ_j from Γ_j.

2. After observing the profile of mechanisms $\gamma = \left(\gamma_1, \ldots, \gamma_J\right)$ offered from $\Gamma := \times_{j \in \mathcal{J}} \Gamma_j$, each agent sends private messages, one to each principal, without observing others' messages; agent i's message to j is $m_{ij} \in M_{ij}$.

3. A principal's action, which may be random, is determined by his mechanism, given the messages he receives, so that principal j takes action $\gamma_j(m_j) \in \mathcal{A}_j$ when he receives the profile of messages $m_j := (m_{ij})_{i \in \mathcal{I}} \in M_j$.

4. Finally, each player $\ell \in \mathcal{I} \cup \mathcal{J}$ earns the stage payoff $u_\ell\left(\gamma_1(m_1), \ldots, \gamma_J(m_j), \theta\right)$.

The following assumption is maintained throughout: messages from agent i to principal j are private, i.e. it is not observable by other players, principals or agents.

Repeated game We now describe the infinitely repeated game $(G, \Gamma)^{\infty}(\delta)$. It involves playing the stage-game (G, Γ) at each time $t \in \mathbb{N}$, with a common discount factor $\delta \in (0,1)$ across periods. At the start of each period, each agent i's type is drawn from the full support distribution μ_i independently of all past types and the current types of the other principals (hence the joint distribution is the product of the marginals, i.e., $\mu = \times_{i \in \mathcal{I}} \mu_i$).

In each period t, a public correlated device (PCD) produces a continuous signal $\omega^t \in \Omega$ from a probability distribution P, independent across periods. All players observe it before principals offer their mechanisms. Players can condition their behavior on the realization of the PCD, which is part of the history.[7]

We adopt the following notational convention: If $\varkappa$ is a variable in the stage game, we denote its period t value by $\varkappa^t$, with the understanding that t is a superscript and not an exponent. Each principal j can offer a mechanism $\gamma_j^t \in \Gamma_j$ and his mechanism offer can be conditional on the realization of the PCD.

At the end of each period t, both agents and principals observe period-t mechanisms offered and period-t actions chosen. Even when only agents can observe mechanisms and actions, our results go through if our model allows agents to send cheap talk messages to principals at the end of each period. (See Section 5.2 for details). Therefore, the only substantive assumption is that agents learn the actions taken given a profile of mechanisms offered by principals.[8]

Starting with the null history h^0, agent i's period-t history h_i^t is constructed from her period-$(t-1)$ history h_i^t according to the formula $h_i^t = h_i^{t-1} \circ (\gamma^t, \alpha^t, m_i^t, \theta_i^t, \omega^t)$, where $\circ$ denotes concatenation and $m_i^t := (m_{ij}^t)_{j \in \mathcal{J}}$. Principal j's period-t private history h_j^t is constructed from his $(t-1)$-period

[7]While public correlation is normally assumed in repeated games, it is not trivial to show when it may be dispensed with (see Fudenberg and Maskin (1991) and Dasgupta and Ghosh (2016)).

[8]Agents learn the random actions taken. While this is non-standard in all but the earliest papers on repeated games, it seems less objectionable in the mechanism design literature as principals can commit to randomize in an observable way. Such observability of actions can play a critical role in designs, as in Ortner and Chassang (2018).

history according to the formula $h_j^t = h_j^{t-1} \circ (\gamma^t, \alpha^t, m_j^t, \omega^t)$ with $m_j^t := (m_{ij}^t)_{i \in \mathcal{I}}$. The (average) discounted payoff of player $\ell \in \mathcal{I} \cup \mathcal{J}$ from period τ onwards is $(1-\delta) \sum_{t \geq \tau} \delta^{t-\tau} u_\ell \left(\alpha^t, \theta^t\right)$.

Our solution concept is perfect Bayesian equilibrium (PBE) defined in Fudenberg and Tirole (1991); see also Watson (2017). It imposes sequential rationality and Bayes' rule wherever possible; in other words, if, under the equilibrium strategies, an information set h' is off path but becomes on path when we condition on a subgame, then beliefs at h' are determined by applying Bayes' rule conditional on the said subgame being reached.[9] Furthermore all players share a common belief about any other player, and no player signals what he does not know.

4 Equilibrium characterization

We shall characterize equilibria in terms of model primitives as $\delta \to 1$ in three steps. Henceforth, when it does not create confusion, we drop the superscript t for notational simplicity.

In the *first step*, we formulate incentive compatibility defined over Θ given a profile of all principals' direct mechanisms and show that we can adopt a weaker notion of incentive compatibility than in the one-shot game. A direct mechanism (DM) $\pi_j : \Theta \to \mathcal{A}_j$ for principal j is a special mechanism where Θ_i is the set of messages that agent i can send to principal j. Let Π_j be the set of all possible direct mechanisms for principal j and $\Pi := \times_{j \in \mathcal{J}} \Pi_j$. We assume that Γ_j is bigger than Π_j ($\Gamma_j \succcurlyeq \Pi_j$) for all $j \in \mathcal{J}$.[10]

In the *second step*, we express every principal's worst equilibrium payoff (i.e., the threshold value of each principal's equilibrium payoff) in terms of incentive compatible DMs and actions. Our approach for this is different from the Revelation Principle for models with a single principal. The Rev-

[9]This condition is what Mailath (2020) calls "almost perfect Bayesian equilibrium".

[10]Formally, $\Gamma_j \succcurlyeq \Pi_j$ if there exists an embedding $\eta_j : \Pi_j \to \Gamma_j$. It implies that there are more mechanisms in Γ_j than in Π_j. We will also assume later that Γ_j is bigger than the set of principal j's extended DMs that ask agents to report their types and the punished principal's actions.

elation Principle allows us to be agnostic about equilibrium strategies that agents play because of the following logic. Any complex mechanism, together with agents' equilibrium strategies, can be converted to an incentive-compatible DM with truthful type reporting as agents' equilibrium strategies. Therefore, a single principal can only focus on incentive compatible DMs where agents simply report their true private information without knowing how agents would play given an arbitrary mechanism. More importantly, the tractability of the Revelation Principle in applications stems from an agent's type being exogenously drawn from a simple space Θ_i for each agent i.

The above approach does not extend to a model with multiple principals because the infinite regress problem introduced by competition among principals makes it extremely complicated for agents to describe principal j's mechanism to the other principals in the phase where principal j is punished. We offer an alternative route to characterize the extreme profile of agents' equilibrium strategies that is the worst for each principal in a phase here he is punished, regardless of the complexity of mechanisms allowed in the game. The extreme profile of agents' equilibrium strategies turns out to be very simple and this is the key step in establishing an extended revelation principle free from the infinite regress problem.

The *final step* establishes an extended revelation principle and a folk theorem simultaneously for the characterization and implementation of equilibrium allocations in terms of actions, DMs, and tractable extended DMs.

4.1 Incentive compatibility over types

Repetition with patient players allows us to relax incentive compatibility because an agent's punishment can in some cases be deferred. To develop this idea further, we introduce some notation. Given a profile of DMs $\pi = (\pi_1, \ldots, \pi_J)$, the expected stage-game payoff of agent i of type θ_i who reports θ_{ij} to each principal j, when the other agents report truthfully, is

$$\mathbb{E}_{\mu_{-i}}\left[u_i\left(\pi_1\left(\theta_{i1},\theta_{-i}\right),\ldots,\pi_J\left(\theta_{iJ},\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right)\right],$$

where $\mathbb{E}_{\mu_{-i}}$ is the expectation operator with respect to the probability distribution μ_{-i} over Θ_{-i}. In the one-shot game with no repetition, when an agent's lie is detected it is too late to punish her; this is why unconstrained incentive compatibility is used in the one-shot game to prevent all possible lies: A profile of DMs $\pi = (\pi_1, \ldots, \pi_J)$ satisfies 'unconstrained incentive compatibility' (UIC) if for all $i \in \mathcal{I}$ and all $\theta = (\theta_i, \theta_{-i}) \in \Theta$, we have

$$\begin{aligned}&\mathbb{E}_{\mu_{-i}}\left[u_i\left(\pi\left(\theta\right),\theta\right)\right] \geq \\ &\quad \mathbb{E}_{\mu_{-i}}\left[u_i\left(\pi_1\left(\theta_{i1},\theta_{-i}\right),\ldots,\pi_J\left(\theta_{iJ},\theta_{-i}\right),\theta\right)\right], \forall\left(\theta_{i1},\ldots,\theta_{iJ}\right) \in \left(\Theta_i\right)^J, \end{aligned} \tag{5}$$

where $\pi\left(\theta\right) := \left(\pi_1\left(\theta\right),\ldots,\pi_J\left(\theta\right)\right)$. Let Π^U be the set of all profiles of DMs satisfying UIC.

Given a profile of DMs, agents' type messages induce a profile of actions. This profile carries some information about the messages sent by agents, and can sometimes reveal which agent lied. For any agent i in the repeated game, incentive compatibility thus does not need to be imposed over those type messages by i that with positive probability lead to a profile of actions that reveals her as the unique agent who lied.

This motivates us to formulate a weaker notion of incentive compatibility than UIC for the repeated game. Given a profile of DMs π and a subset of agents S, let $\mathcal{A}(\pi, S)$ denote the set of all action profiles induced when those outside S report their types truthfully while messages sent by those in S are unrestricted. Clearly the set of actions profiles when all agents tell the truth is $\mathcal{A}(\pi, \emptyset)$ and satisfies $\mathcal{A}(\pi, \emptyset) \subset \mathcal{A}(\pi, i)$ $\forall i$. Action profiles in the set

$$\mathcal{A}(\pi, i) \setminus (\cup_{k \neq i} \mathcal{A}(\pi, k)) =: \mathcal{A}_L^i(\pi) \tag{6}$$

arise only when i lies to at least one principal. The set $L_i\left(\pi\right)$ comprises message profiles that i can send, one to each principal, so as to bring about actions in the set above in (6). Formally,

$$L_i(\pi) := \left\{ (\theta_{ij})_j \in (\Theta_i)^J \middle| \exists \theta_{-i} \in \Theta_{-i} \text{ s.t. } \Big(\pi_1 \left(\theta_{i1}, \theta_{-i}\right), \ldots, \pi_J \left(\theta_{iJ}, \theta_{-i}\right) \Big) \in \mathcal{A}_L^i(\pi). \right\}$$

In words, $L_i(\pi)$ is the set of agent i's type messages that have a positive probability of generating an action profile that reveals agent i as the unique agent who deviated from truth telling when faced with π. Note that any consistent message profile $(\theta_{i1}, \ldots, \theta_{iJ})$ of agent i (i.e., $\theta_{i1} = \theta_{i2} \ldots = \theta_{iJ}$) is not in $L_i(\pi)$ even though it is a lie. If a message profile $(\theta_{i1}, \ldots, \theta_{iJ})$ is in $L_i(\pi)$, it is an inconsistent message profile (i.e., $\theta_{ij} \neq \theta_{ij'}$ for some $j, j' \in \mathcal{J}$ with $j \neq j'$).

As we shall in the folk theorem, messages in $L_i(\pi)$ can be deterred because a positive probability of detection is enough. The notion below anticipates this and does not impose incentive compatibility over type messages in $L_i(\pi)$.

Definition 1 *A profile of DMs π satisfies constrained incentive compatibility (CIC) under a type distribution μ if for all $i \in \mathcal{I}$ and all $\theta = (\theta_i, \theta_{-i}) \in \Theta$, we have*

$$\mathbb{E}_{\mu_{-i}}\left[u_i\left(\pi\left(\theta\right), \theta\right)\right] \geq \mathbb{E}_{\mu_{-i}}\left[u_i\left(\pi_1\left(\theta_{i1}, \theta_{-i}\right), \ldots, \pi_J\left(\theta_{iJ}, \theta_{-i}\right), \theta\right)\right] \\ \forall (\theta_{i1}, \ldots, \theta_{iJ}) \in (\Theta_i)^J \setminus L_i(\pi). \quad (7)$$

Let Π^C be the set of all profiles of DMs satisfying CIC. Note that $\Pi^U \subset \Pi^C$.

Given a profile of mechanisms $\gamma_j^t \in \Gamma$ in period t, agent i's (pure) communication strategy s_i^t specifies an array of messages she sends to principals given her private history up to period $t-1$ and the period-t values of the public randomization device, the mechanisms offered, and the types:

$$s_i^t\left(h_i^{t-1}, \omega^t, \gamma^t, \theta_i^t\right) = \left[s_{i1}^t\left(h_i^{t-1}, \omega^t, \gamma^t, \theta_i^t\right), \ldots, s_{iJ}^t\left(h_i^{t-1}, \omega^t, \gamma^t, \theta_i^t\right)\right] \in M_{i1} \times \cdots \times M_{iJ}.$$

Letting $m_{ij}^t\left(\theta_i^t\right) := s_{ij}^t\left(h_i^{t-1}, \omega^t, \gamma_j^t, \gamma_{-j}^t, \theta_i^t\right)$, principal j's t-period DM is given by

$$\pi_j^t := \gamma_j^t \circ (m_{ij}^t)_{i\in\mathcal{I}}. \tag{8}$$

Remark 1 *We study PBE where the profile of agents' communication strategies induces CIC DMs from the mechanisms offered by principals given any histories of agents.*

Section 4.3 explains why CIC is generally needed in the phase where a player is punished. However, we relax it even more in the phase where no player has previously deviated (See Section 5 for details).

4.2 Worst equilibrium payoffs

Let $\Pi(\gamma) \subset \Pi^C$ denote the set of all profiles of CIC DMs that can be induced by all profiles of agents' equilibrium communication strategies given mechanisms γ offered by principals in the phase where principal j is punished for his past deviation. Given γ, principal j's lowest possible stage payoff is

$$\underline{u}_j(\gamma) := \min_{\pi\in\Pi(\gamma)} \mathbb{E}_\mu \left[u_j(\pi(\theta),\theta)\right],$$

where $\mathbb{E}_\mu[\cdot]$ is the expectation operator over Θ given the probability distribution μ. Subsequently, principal j's worst PBE payoff (i.e., the threshold value of principal j's PBE payoff), denoted by $\underline{w}_j^C$, is

$$\underline{w}_j^C := \min_{\gamma_{-j}\in\Gamma_{-j}} \max_{\gamma_j\in\Gamma_j} \underline{u}_j\left(\gamma_{-j},\gamma_j\right).$$

One could provide a folk theorem based on $\underline{w}_j^C = \min_{\gamma_{-j}\in\Gamma_{-j}} \max_{\gamma_j\in\Gamma_j} \underline{u}_j\left(\gamma_{-j},\gamma_j\right)$ but it does not provide any insight into the characterization and implementation of equilibrium allocations in terms of model primitives. For that, we first identify the lower bound of $\underline{w}_j^C$ in terms of model primitives such as actions and DMs independent of complex mechanisms allowed in the game.

Let us introduce a few notations. An action α_j can be thought of as a simple mechanism that always assigns the same action α_j regardless of agents' messages. If principal j chooses an action α_j instead of a more complex mechanism, then one can consider a profile of CIC DMs for principals except for j. Let $\Pi^C_{-j}(\alpha_j)$ be the set of all profiles of DMs for principals except j that are CIC conditional on α_j:

$$\Pi^C_{-j}(\alpha_j) := \left\{ \pi_{-j} | \left(\pi_{-j}, \alpha_j \right) \in \Pi^C \right\}.$$

Lemma 1 *For every principal $j \in \mathcal{J}$,*

$$\max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi^C_{-j}(\alpha_j)} \mathbb{E}_\mu \left[u_j \left(\pi_{-j} \left(\theta \right), \alpha_j, \theta \right) \right] \leq \underline{w}^C_j \tag{9}$$

Proof. Suppose that principal j considers a choice of a mechanism off the path where he is punished for his past deviation. Because the simplest mechanism he can choose off the path is a constant mechanism, that is, a single action $\alpha_j \in \mathcal{A}_j$, we have that

$$\min_{\gamma_{-j} \in \Gamma_{-j}} \max_{\alpha_j \in \mathcal{A}_j} \underline{u}_j \left(\gamma_{-j}, \alpha_j, \theta \right) \leq \underline{w}^C_j. \tag{10}$$

Suppose that principals except for j are perfectly informed about an action α_j principal j chooses when j restricts himself to $\mathcal{A}_j$. Conditional on each action α_j that principal j may take, the other principals cannot lower principal j's payoff below

$$\min_{\pi_{-j} \in \Pi^C_{-j}(\alpha_j)} \mathbb{E}_\mu \left[u_j \left(\pi_{-j} \left(\theta \right), \alpha_j, \theta \right) \right].$$

It implies that

$$\max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi^C_{-j}(\alpha_j)} \mathbb{E}_\mu \left[u_j \left(\pi_{-j} \left(\theta \right), \alpha_j, \theta \right) \right] \leq \min_{\gamma_{-j} \in \Gamma_{-j}} \max_{\alpha_j \in \mathcal{A}_j} \underline{u}_j \left(\gamma_{-j}, \alpha_j, \theta \right). \tag{11}$$

(10) and (11) lead to (9). ■

The value $\underline{w}^C_j$ is based on the profile of agents' equilibrium communication strategies that is the worst for principal j in the phase where he is punished for his past deviation. We construct a profile of equilibrium communication strategies that makes principal j no better off with any mechanism in Γ_j if he is being punished, than he does with a single action in $\mathcal{A}_j$ and that also makes the other principals perfectly informed about principal j's action. This makes $\underline{w}_j^C$ equal to the left hand side of (9). To this end we introduce class of mechanisms that is only slightly more general than DMs.

Definition 2 (EDM) *A mechanism $\zeta_\ell : T_\ell \to \mathcal{A}_\ell$ offered by principal ℓ is said to be an 'extended direct mechanism' (EDM) if for some $j \neq \ell$ we have $T_\ell = \times_{i\in\mathcal{I}} T_{i\ell}$ and $T_{i\ell} = \mathcal{A}_j \times \Theta_i$.*

Theorem 1 *Suppose that agents can be enforced to always induce an action in $\gamma_j\left(M_j\right)$ only conditional on principal j's mechanism γ_j in the phase he is punished for his past deviation. Then, for each principal j,*

$$\max_{\alpha_j\in\mathcal{A}_j} \min_{\pi_{-j}\in\Pi^C_{-j}(\alpha_j)} \mathbb{E}_\mu\left[u_j\left(\pi_{-j}(\theta), \alpha_j, \theta\right)\right] = \underline{w}_j^C \tag{12}$$

Furthermore,

1. *$\underline{w}_j^C$ is achieved if each principal ℓ ($\neq j$) offers an incentive compatible extended DM (EDM) where agents are asked to report their own types and principal j's action;*

2. *in the phase where principal j is being punished, he cannot do any better by offering a mechanism in Γ_j than by offering an action in $\mathcal{A}_j$.*

Proof. Suppose that principal j offers a mechanism $\gamma_j \in \Gamma_j$ in the phase where he is punished for his past deviation. Suppose that agents send messages to j so as to induce an action in $\gamma_j\left(M_j\right)$ only conditional on γ_j, say $g_j(\gamma_j) \in \gamma_j\left(M_j\right)$ irrespective of their own types. This means that offering γ_j is equivalent to offering a single action $g_j(\gamma_j)$.

Then, we can restrict principal j to choose an action in $\mathcal{A}_j$ in the phase where he is punished. For any given action α_j that principal j takes, for the other principals, punishing principal j is equivalent to choosing a profile of their CIC DMs conditional on α_j. The reason is that agents' equilibrium communication with them, given the mechanisms that they choose, induces CIC DMs. Therefore, principal j's lowest possible payoff conditional on α_j can be realized if other principals can implement $\varphi^j_{-j}(\alpha_j) = \left(\varphi^j_\ell(\alpha_j)\right)_{\ell \neq j} \in \Pi^C_{-j}(\alpha_j)$, where $\varphi^j_\ell(\alpha_j)$ is principal ℓ's DM and $\varphi^j_{-j}(\alpha_j)$ is defined as

$$\varphi^j_{-j}(\alpha_j) \in \underset{\pi_{-j} \in \Pi^C_{-j}(\alpha_j)}{\arg\min}\ \mathbb{E}_\mu\left[u_j\left(\pi_{-j}(\theta), \alpha_j, \theta\right)\right].$$

Each non-deviating principal ℓ can implement $\varphi^j_\ell(\alpha_j)$ by offering an EDM. A non-deviating principal's EDM asks agents to report the action that deviating principal j takes, along with their type messages. If a majority of agents report α_j, then principal ℓ implements the DM $\varphi^j_\ell(\alpha_j)$, which then assigns ℓ's action according to agents' type messages. Formally we can define an EDM as follows. Let ξ^j_ℓ be principal ℓ's EDM that is used in punishing principal j after his deviation. In this EDM, each agent i's message space is $T_{i\ell} := \mathcal{A}_j \times \Theta_i$. Let $T_\ell := \times_{i \in \mathcal{I}} T_{i\ell}$. Then, ξ^j_ℓ is a mapping from T_ℓ into $\mathcal{A}_\ell$. Let $(\tilde{\alpha}^j_{i\ell}, \widetilde{\theta}_{i\ell})$ denote a pair of messages that agent i sends to principal ℓ, where $\tilde{\alpha}^j_{i\ell}$ is principal j's action that agent i reports and $\widetilde{\theta}_{i\ell}$ is the type that she reports as her type. Let $\tilde{\alpha}^j_\ell := (\tilde{\alpha}^j_{i\ell})_{i \in \mathcal{I}}$ be a profile of all agents' messages on principal j's action and $\widetilde{\theta}_\ell := (\widetilde{\theta}_{i\ell})_{i \in \mathcal{I}}$ a profile of all agents' messages on their types. Then, $\xi^j_\ell(\tilde{\alpha}^j_\ell, \widetilde{\theta}_\ell)$ denotes principal ℓ's action when $(\tilde{\alpha}^j_\ell, \widetilde{\theta}_\ell)$ is a profile of agents' messages. For the implementation of the DM, $\varphi^j_\ell(\alpha_j)$, the majority rule of the EDM is specified as: For all $(\tilde{\alpha}^j_\ell, \widetilde{\theta}_\ell) \in T_\ell$,

$$\xi^j_\ell(\tilde{\alpha}^j_\ell, \widetilde{\theta}_\ell) := \begin{cases} \varphi^j_\ell(\alpha_j)\left(\widetilde{\theta}_\ell\right) & \text{if } \exists \alpha_j \text{ s.t. } \#\left\{i : \tilde{\alpha}^j_{i\ell} = \alpha_j\right\} > I/2, \\ \overline{\alpha}_\ell & \text{otherwise} \end{cases}, \tag{13}$$

where $\overline{\alpha}_\ell$ is some arbitrary action in $\mathcal{A}_\ell$. Recall that we assume that there are three or more agents. If all agents report α_j, any single agent's deviation from α_j cannot change principal ℓ's DM away from $\varphi^j_\ell(\alpha_j)$. Therefore, truthful action reporting can be sustained.

For the last step of the proof, we choose the mapping $g_j : \Gamma_j \to \mathcal{A}_j$ as follows. If principal j offers γ_j off the path following his deviation, let agents send messages to principal j that induce $g_j\left(\gamma_j\right)$ such that

$$g_j\left(\gamma_j\right) \in \min_{\alpha_j \in \gamma_j\left(M_j\right)} \mathbb{E}_\mu\left[u_j\left(\varphi_{-j}^j\left(\alpha_j\right)(\theta), \alpha_j, \theta\right)\right] \tag{14}$$

Agents report $\alpha_j = g_j\left(\gamma_j\right)$, together with their types, to non-deviating principals who offer EDMs. Theorem 2 below shows that this is part of an equilibrium in the phase where principal j is punished for his past deviation. This means that principal j cannot do any better by offering a mechanism in Γ_j than he does by offering an action in $\mathcal{A}_j$. Therefore, even when principal j can choose any mechanism in Γ_j, the maximum payoff principal j can achieve is

$$\max_{\alpha_j \in \mathcal{A}_j} \mathbb{E}_\mu\left[u_j\left(\varphi_{-j}^j\left(\alpha_j\right)(\theta), \alpha_j, \theta\right)\right] = \max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C\left(\alpha_j\right)} \mathbb{E}_\mu\left[u_j\left(\pi_{-j}(\theta), \alpha_j, \theta\right)\right], \tag{15}$$

which is the left-hand side of (9). ■

The crux in Theorem 1 is whether principals can deter any agent's unilateral deviation to induce an unexpected action other than $g_j(\gamma_j)$. If an unexpected action (i.e., some action other than $g_j(\gamma_j)$) is induced from γ_j, principals know that at least one agent deviated. They may not know the deviating agent's identity. Nonetheless, our folk theorem (Theorem 2) shows that an agent's deviation to induce such an unexpected action can be deterred if principals punish every agent with equal probability, $1/I$ even when the identity of the deviating agent is not known.[11].

Theorem 1 show how to express $\underline{w}_j^C$ in terms of incentive compatible DMs and actions and how to implement it. In the phase where principal j

[11]We partition the space $(0, 1]$ into $(i/I, (i+1)/I]$ for $i = 0, 1, ..., I-1$. A PRD in the ith partition at the start of period immediately following a deviation leads to all principals punishing agent i.

is punished, $\underline{w}_j^C$ is reached without loss of generality when each principal except for j offers an extended DM defined in (13) and agents block any information transmission to principal j. Consequently, in the phase where principal j is punished, he cannot do any better with any mechanism than he does with a single action.

Agent's worst equilibrium payoff Now, let us consider agent i's worst PBE payoff. In the phase where agent i is punished, CIC cannot be enforced to agent i because she will report her type to increase her current stage payoff. Therefore, UIC is the notion of IC that needs to be imposed for agent i. On the other hand, principals can enforce CIC to the other agents because their deviation that is detectable with a positive probability can be deterred. Let $\Pi^C(i)$ be the set of all profiles of DMs that are UIC for agent i but CIC for the other agents, that is a profile of DMs in $\Pi^C(i)$ imposes IC to agent ℓ ($\neq i$) only over the set of type messages in $(\Theta_\ell)^J \setminus L^\ell(\pi)$ while imposing IC to agent i over the set of type messages in $(\Theta_i)^J$.

Remark 2 *Agent i's worst PBE payoff is*

$$\underline{w}_i^C := \min_{\pi \in \Pi^C(i)} \mathbb{E}_\mu \left[u_i \left(\pi(\theta), \theta \right) \right]. \tag{16}$$

4.3 Extended revelation principle and a folk theorem

A stage allocation or social choice function (SCF) is a mapping $f : \Theta \to \triangle(A_1 \times \cdots \times A_J)$ from type profiles to (possibly correlated) probability distributions over actions. The set of (correlated) stage-SCFs is denoted by $\mathcal{F}$. A subclass of SCFs is obtained by the principals mixing independently of one another:

$$\mathcal{F}_0 := \{ f \in \mathcal{F} | f(\Theta) \subset \mathcal{A} \} \subset \mathcal{F}. \tag{17}$$

$\mathcal{F}_0$ is indeed equal to the set of all DM profiles: $\mathcal{F}_0 = \Pi$.

What stage-SCFs and payoff profiles can we support in a PBE of $(G, \Gamma)^\infty(\delta)$? We say $f \in \mathcal{F}$ is strictly individually rational (SIR) (w.r.t. $\mu \in \triangle\Theta$) if each player ℓ gets an expected payoff above $\underline{w}_\ell^C$:

$$\mathbb{E}_\mu \left[u_\ell \left(f\left(\theta\right), \theta \right) \right] > \underline{w}_\ell^C \text{ for all } \ell \in \mathcal{I} \cup \mathcal{J}. \tag{18}$$

$\mathcal{F}_0$ is the set of SCFs that are obtained by independently mixing principals' actions. The smaller class of SCFs that also satisfy CIC is written as $\mathcal{F}_0^C$; these may not satisfy SIR. We now define the class of SCFs that are SIR and result from picking a CIC SCF by observing the realization of the PCD:

$$\mathcal{F}^C\left(\mu\right) := \left\{ f^* \in \triangle \mathcal{F}_0^C | f^* \text{ is SIR w.r.t } \mu \right\}.$$

The theorem below shows that any (correlated) SCF $f^* \in \mathcal{F}^C\left(\mu\right)$ is supportable in a PBE of $\left(G, \Gamma\right)^\infty \left(\delta\right)$, provided players are sufficiently patient. Note that any $f^* \in \mathcal{F}^C\left(\mu\right)$ is a probability distribution over SCFs that are induced by profiles of CIC DMs. Each SCF (i.e., each profile of DMs) in the support of $f^* \in \mathcal{F}^C\left(\mu\right)$ does not need to be SIR. As in any repeated game, what matters is that the expected payoff is above the threshold even if the current period's payoff (after observing the PCD) isn't.

We make the standard full dimensionality assumption (FD) that the set of expected payoffs has the same dimension as the number of players, i.e. $\dim\left[u\left(\mathcal{F}^C\left(\mu\right)\right)\right] = J + I$. As is well known from the theory of repeated games (Fudenberg and Maskin 1986), this allows us to design player-specific punishments. We present the folk theorem and extended revelation principle in model primitives.

Theorem 2 *Consider i.i.d. types with distribution $\mu \in \triangle\Theta$. Under the standard full dimensionality assumption on the expected payoffs,*

1. *any (correlated) SCF $f^* \in \mathcal{F}^C\left(\mu\right)$ is the outcome of a PBE of $\left(G, \Gamma\right)^\infty \left(\delta\right)$ for high δ;*

2. *Incentive-compatible DMs suffice in phases where no one has previously deviated, or an agent is punished, or players stay after punishing a player, while in the phase where principal j is punished, all other principals employ*

incentive-compatible EDMs whereas j offers a single action, i.e. a constant mechanism.

Proof. See the appendix. ∎

Suppose that $f = (f_1, \ldots, f_J)$ in the support of $f^* \in \mathcal{F}^C(\mu)$ is a SCF that needs to be supported. Each principal j offers the DM $\pi_j = f_j$. As long as no principal deviates, agents truthfully report their types. If a principal deviates, agents play a one-shot equilibrium, expecting that the deviating principal will be punished from the next period. In phases where an agent is punished or players stay after punishing a player, principals can also simply offer incentive compatible DMs.

The extended revelation principle for the off-path mechanisms used in punishing a principal for his past deviation is based on the profile of agents' equilibrium communication strategies that is worst for the deviating principal. Only single actions are required for the punished principal off the path following his deviation because agents do not reveal any information to the punished principal regardless of his mechanism. This makes the infinite regress problem of mechanisms play no role and agents only need to report the action of the punished principal and their types to the other principals who offer incentive compatible EDMs. Further, the incentive compatibility in the EDMs is imposed only over agents' type messages conditional on their messages on the punished principal's action but not the messages on the punished principal's action because truthful action reporting can be always enforced.

We now provide some intuition for the folk theorem. The familiar stick-and-carrot schemes by Fudenberg and Maskin (1986) or Abreu, Dutta, and Smith (1991) are used to ensure sequential rationality based on player-specific punishments, except for allowing for random player-specific punishments when the identity of a deviating agent is not known in a phase where a principal is punished.

Notation: In discussions related to folk theorems, generic players are denoted by i and j unless explicitly noted otherwise.

Fix any (correlated) SCF $f^* \in \mathcal{F}^C(\mu)$ that yields v_j as player j's expected payoff for $j \in \mathcal{I} \cup \mathcal{J}$. This is the target equilibrium payoff for player j. Following Fudenberg and Maskin (1986), we choose a vector of payoffs $\left(v'_1, \ldots, v'_{I+J}\right)$, with $\underline{w}_j^C < v'_j < v_j$ for all $j \in \mathcal{I} \cup \mathcal{J}$. The full dimensionality assumption ensures that for each j, there exists $\epsilon > 0$ and a SCF in $\mathcal{F}^C(\mu)$ that yields expected payoffs

$$\beta^j := \left(\beta_i^j\right)_{i=1}^{i=I+J} = \left(v'_1 + \epsilon, \ldots, v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, \ldots, v'_{J+I} + \epsilon\right)$$

such that ϵ is small enough to satisfy $\beta_i^j < v_i$ for $i \neq j$. A principal's deviation can be easily deterred with the standard player-specific punishment. Punishing an agent's deviation is more subtle. For example, given the notion of CIC, a deviation from truthful type reporting can be detected only with a positive probability. Let $p > 0$ be the minimum probability with which an agent's deviation from truthful type reporting is detected across all agents and all possible profiles of CIC DMs that can be offered on and off the path. Following Abreu, Dutta and Smith (1994), after player j deviates, players stay in phase II_j where the deviator is punished with probability q and move to the phase where each player i $(\neq j)$ is rewarded with β_i^j with probability $1 - q$.

Recall that $\overline{u} = \max_{\mathcal{I} \cup \mathcal{J}, \mathcal{A}, \Theta} |u_i(\alpha, \theta)|$ denotes the maximum stage-game payoff. Find a parameter $q \in (0, 1)$ such that

$$\overline{u}(1-q) < \beta_j^j (1 + p - q) - p\underline{w}_j^C \text{ for all } j \in \mathcal{I} \cup \mathcal{J}. \tag{19}$$

Such a q exists because at $q = 1$ this inequality becomes $0 < p(\beta_j^j - \underline{w}_j^C)$, which is satisfied because $\beta_j^j = v'_j > \underline{w}_j^C$. We show that an agent has no incentive to report false types, given q that satisfies (19).[12]

When principal j offers any arbitrary mechanism γ_j in phase II_j where

[12] If every player's deviation is detected with probability one (i.e., $p = 1$) and players' lower bounds are normalized to zero, (19) is reduced to the condition that is used in setting up the value of q in Abreu, Dutta, and Smith (1994).

he is punished, agents are supposed to induce an action $g_j\left(\gamma_j\right) \in \gamma_j\left(M_j\right)$. If an agent i deviates to induce another action, such a deviation is detected even though the identity of the deviating agent may not be observed. In this case, each agent is punished with equal probability. If agent i conforms to induce $g_j\left(\gamma_j\right)$, she will be rewarded after punishing principal j is done. If agent i deviates but the other agents are punished instead, she will get the same reward after punishing the other agents are done. If she deviates, she is punished with probability $1/I$ and in this case she will lose the reward. Because this loss happens with probability $1/I$, agent i will not deviate if she is sufficiently patient.

A natural question arises as to whether we can further relax CIC. For example, it is tempting to remove IC over profile of messages that lead with positive probability to action profiles from which principals can infer that an agent deviated even though the identity of the deviating agent may not be known. We cannot dispose of IC over such message profiles off the path. Suppose that players went through phase II_i where agent i was punished and that the game reached the final phase III_i where all other players, except for agent i, are rewarded after punishing her in phase II_i. If agent i lies in this final phase and principals only know that at least one agent deviates but not the identity of the deviating agent, they cannot simply punish every agent with equal probability because that would, with positive probability, let agent i participate in punishing other agents and reap the rewards forever after punishment is done. This may trigger agent i's deviation in this final phase if she is sufficiently patient. Therefore, truth telling may not lead to a continuation equilibrium in the final phase and it is not clear what kind of a non-truthful equilibrium will then arise in the final phase. Depending on the non-truthful equilibrium that prevails in the final phase, the agent may want to deviate on the path in the first place.

However, CIC can be further relaxed on the equilibrium path where no one has previously deviated, over message profiles that lead with positive probability to action profiles from which principals can infer an agent's deviation but not the identity of the deviating agent. This is discussed in Section 5.3.

5 Discussions

Our paper studies dynamic contracting competition among principals, putting contracting processes between multiple principals and multiple agents to the forefront of equilibrium analysis. For the characterization and implementation of equilibrium allocations, the extended revelation principle only needs to be based on the extreme profile of agents' equilibrium strategies that is worst for a principal in a phase where he is punished for his past deviation. We construct them regardless of the complexity of contracts. The key insight into our results is that contrary to what the term "extended revelation principle" might suggest, agents block any information transmission to a principal in a phase where he is punished. This leads to the extended revelation principle and the characterization and implementation of equilibrium allocations in terms of model primitives and they are free from the infinite regress problem.

Our results do not change with extensions in several fronts.

5.1 Two agents

The assumption of three or more agents makes it easy to force agents to truthfully report principal j's action to other principals when the latter offer EDMs in the phase where principal j is punished for his past deviation. Given the majority rule employed in EDMs, a single agent's deviation from true action reporting has no effect when the remaining agents all report the same true action.

Suppose that there are only two agents. Consider a phase where each principal ℓ ($\neq j$) offers the EDM ξ^j_ℓ defined in (13) to punish principal j. If two agents' reports on j's action to principal ℓ are not consistent, principal ℓ knows that at least one agent has deviated. However, the other principals do not know that because agents' action reports to principal ℓ are not observable by them. In order to punish an agent together with the other principals, principal ℓ needs to let them know that at least one agent has sent to him a false message.

Note that given the two agents, principal ℓ's EDM ξ_ℓ^j assigns $\overline{\alpha}_\ell$ when the two agents send inconsistent messages on j's action. Can $\overline{\alpha}_\ell$ be a perfect signal to the other principals on agents' inconsistent messages on j's action to principal ℓ?

Principal ℓ's EDM assigns an action $\varphi_\ell^j(\alpha_j)(\theta)$ when θ is a profile of type messages that agents send to principal ℓ and more than a half of agents (both agents in the model with two agents) send α_j as j's action to principal ℓ. Recall that $\varphi_{-j}^j(\alpha_j)$ is the profile of CIC DMs for the principals except for j that minimizes principal j's payoff conditional on α_j. Then, $\{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\}$ is the set of principal ℓ's actions that can be induced by all profiles of agents' consistent messages on principal j's action and all profile of their type messages. If it is the strict subset of $\mathcal{A}_\ell$ (i.e, $\{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\} \subsetneq \mathcal{A}_\ell$), then $\mathcal{A}_\ell \backslash \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\}$ is non-empty. For $\overline{\alpha}_\ell$, principal ℓ can choose an arbitrary action in $\mathcal{A}_\ell \backslash \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\}$. Upon observing $\overline{\alpha}_\ell$ at the end of the period, the other principals know that two agents sent inconsistent messages on principal j's action to principal ℓ. After observing $\overline{\alpha}_\ell$, each agent is punished with equal probability.

Therefore, a sufficient condition for providing both agents with incentives to truthfully report principal j's action to each principal $\ell (\neq j)$ is $\{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\} \subsetneq \mathcal{A}_\ell$, equivalently,

$$\mathcal{A}_\ell \backslash \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\} \neq \emptyset. \tag{20}$$

We believe that this sufficient condition is very weak because it is satisfied as long as principal ℓ does not need to use all of his actions in punishing principal j. Let us go back to the example of manufacturers and retailers in Section 2. Manufacturer ℓ charges the price of 0.21321 to every retailer in a phase where manufacturer ℓ is punished. Therefore, (20) is satisfied. Manufactuere ℓ can charge any other price to signal if retailers send inconsistent messages on manufacturer j's action. Such a price becomes the perfect signal on retailers' inconsistent messages on manufacturer j's action. From the next period, both manufacturers exclude each retailer from the market with equal probability as punishment.

5.2 Observability

We assume that, at the end of each period, mechanisms and actions are observable to both principals and agents. Suppose that only agents, but not principals, observe actions and mechanisms at the end of each period. A principal could infer the other principals' actions to a certain extent from the payoff he receives at the end of the period. However, in general, he may not be able to pin down the true actions chosen by the other principals from his realized payoff. In this case, principals need to allow agents to report the identity of a player who deviates.

First, consider the case with three or more agents when mechanisms and actions are observable by agents only. For characterization and implementation of equilibrium allocations in terms of model primitives, we modify the model of dynamic contracting competition so that agents are allowed to send *privately observed cheap talk messages* to principals at the end of every period.

Consider the phases where no player has previously deviated, or an agent is punished, or players stay after punishing a player. In these phases, suppose that principals are supposed to offer a profile of CIC DMs $\pi = (\pi_1, \ldots, \pi_J)$ at the beginning of a period. Each principal j offers π_j and agents send type messages at the beginning of the period, and agents to send a cheap talk message from $\{1, \ldots, J\}$ at the end of the period. $\{1, \ldots, J\}$ is the set of cheap talk messages available for reporting a deviating principal. If more than a half of agents are silent at the end of the period, a principal believes no one has deviated in the current period. If more than a half of agents report j to each principal ℓ then those principals believe that j has deviated and it is what principal j expects when he deviates. Therefore, if principal j deviates, agents report j to every principal ℓ ($\neq j$) at the end of the period. After j's deviation, punishment starts from the next period.

Consider the phase where a principal is punished. Given the profile of EDMs $\xi^j_{-j} = (\xi^j_\ell)_{\ell \neq j}$ that principals except for j are supposed to use in the phase where j is punished, each principal ℓ ($\neq j$) offers ξ^j_ℓ at the beginning of a period and let agents send a cheap-talk message from $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$ at the end of a period. ($\mathcal{P}(\mathcal{I} \cup \mathcal{J})$ is the power set of $\mathcal{I} \cup \mathcal{J}$). Principal j would choose whichever mechanism he wants and let agents send a cheap-talk message from $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$. If more than a half of agents are silent at the end of the period, a principal believes that no one has deviated in the current period. If more than a half of agents send a cheap talk message $\{k\}$ for $k \in \mathcal{J}$ to each principal ℓ ($\neq k$), each principal ℓ believes that principal k has deviated in the curent period. If more than a half of agents send non-empty subsets of $\mathcal{I}$ to every principal, all principals believe that there is at least one agent has deviated in the current period: this is crucial to enforce agents to send messages that always induce $g_j(\gamma_j)$ from principal j's mechanism γ_j when j is punished. Given principals' beliefs, agents send the truthful cheap talk messages at the end of the period when a player deviates.

Given the availability of privately observed cheap talk messages at the end of each period, $\mathcal{F}^C(\mu) = \{f^* \in \triangle\mathcal{F}^C_0 | f^* \text{ is SIR w.r.t } \mu\}$ is still the set of equilibrium allocations in the case with three or more agents.

In the case with two agents, we need (20) for EDMs to enforce agents to report the expected action $g_j(\gamma_j)$ when principal j is punished and his mechanism is γ_j. In addition to it, when actions and mechanisms are observable by agents only, the other principals need to check whether the expected action $g_j(\gamma_j)$ is actually taken by the punished principal or there is an agent who deviates to induce an action other than $g_j(\gamma_j)$ at the end of the period. It is possible if a cheap talk message from $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$ that each agent sends to each principal is *publicly observable* (or each agent announces a publicly observable single cheap talk message). Only when neither agent sends a subset of $\mathcal{I}$ as her cheap talk message to a principal, all principals believe that no agents has deviated to induce an unexpected action from principal j's mechanism. Therefore, if at least one agent sends a subset of $\mathcal{I}$ to a principal, then all principals believe that at least one agent has deviated and it triggers the punishment of an agent by all principals with equal probability. Given principals' beliefs, agents send truthful cheap talk messages at the end of the period.

Given (20) and the availability of publicly observed cheap talk messages at the end of each period, $\mathcal{F}^C(\mu) = \{f^* \in \triangle\mathcal{F}_0^C | f^* \text{ is SIR w.r.t } \mu\}$ is also the the set of equilibrium allocations in the case with two agents.

5.3 Relaxing incentive compatibility

Our folk theorem and the extended revelation principle are based on our notion of incentive compatibility (CIC). Above we argued that off the equilibrium path incentive compatibility is imposed over message profiles that lead only to action profiles revealing that an agent deviated, but not the identity of the deviating agent. When the current mechanism is π, these are precisely those misreports that are not in $L^i(\pi)$ for any i. This is used to deter an agent from further deviation from her own punishment after her initial deviation. Therefore, CIC is the proper notion in constructing the worst PBE payoffs.

Contrary to the incentive compatibility imposed off the path, we can dispense with incentive compatibility on the equilibrium path, over messages profiles that lead with positive probability to an action profile that only reveals that some agent deviated, but not the identity of the deviating agent.

Recall that $\mathcal{A}(\pi, \varnothing)$ is the set of actions that can be induced from a profile of DMs π by agents' truthful type messages. Define $B_i(\pi) \subset (\Theta_i)^J$ as the set of all profiles of type messages of agent i, one message to each principal, that lead to an action profile in $\mathcal{A}(\pi, \varnothing)$ irrespective of the types of the other agents as long as the others report truthfully:

$$B_i(\pi) :=$$
$$\left\{(\theta_{i1},\ldots,\theta_{iJ}) \in (\Theta_i)^J \mid \left[\pi_1(\theta_{i1},\theta_{-i}),\ldots,\pi_J(\theta_{iJ},\theta_{-i})\right] \in \mathcal{A}(\pi,\varnothing)\ \forall\theta_{-i} \in \Theta_{-i}\right\}$$

Suppose that given the realization of the PCD, principals offer a profile of DMs π that impose incentive compatibility only over $B_i(\pi)$ for all i on the equilibrium path. If agent i sends a message profile outside $B_i(\pi)$, then she may induce an 'unexpected' action profile, which could not have been induced under truthful type reporting. If only agent i could have induced this unexpected action profile by sending a profile of messages outside her $B_i(\pi)$, then i is singled out for punishment.

However, even if principals do not know the identity of the deviating agent (because two or more agents could have caused this by some unilateral deviation), they can pick out any agent i with equal probability and punish her before moving on to Phase III. Following the notation in the appendix, let L_i^k denote agent i's expected average life-time payoff in the phase where player k is punished. If agent i deviates to lie on the equilibrium path, her expected average life-time payoff cannot be greater than $(1-\delta)\overline{u}+\delta\left[(1-p^*)v_i+\frac{p^*}{I}\sum_{k=J+1}^{J+I}L_i^k\right]$, where p^* is the maximum probability that an agent's lie is detected and $\overline{u}$ is the maximum utility. As $\delta \to 1$, this payoff approaches $(1-p^*)v_i+p^*v_i'+p^*\frac{I-1}{I}\epsilon$. This is less than v_i given that $v_i'+\epsilon < v_i$. Since v_i is the equilibrium payoff for agent i, a sufficiently patient agent i will not deviate to a lie on the path.

Jackson and Sonnenschein (2007) have shown how a single mechanism designer can 'link' several decisions together by giving each agent a 'budget' of reports that encompasses a large number of identical decisions, and thereby impose some 'budgets' on messages, instead of directly providing incentives on a period-by-period basis. Jackson and Sonnenschein (2007) show that an ex-ante Pareto efficient social choice function can be supported through such budgeted messages. Can we relax incentive compatibility even further on the equilibrium path using this approach? At this point we should like to point out how both our model and our result don't fit into their scheme. First, we have with multiple mechanism designers receiving private messages. Second, our target payoffs need not be ex-ante Pareto-efficient.

References

[1] Abreu, D., D. Dutta and L. Smith (1994), "The Folk Theorem for Repeated Games: A NEU Condition," *Econometrica*, 62(4), 939-948.

[2] Athey, S. and I. Segal (2013), "An Efficient Dynamic Mechanism," *Econometrica*, 81(6), 2463-2485.

[3] Battaglini, M. and R. Lamba (2019), "Optimal Dynamic Contracting: The First-Order Approach and Beyond," *Theoretical Economics*, 14, 1435-1482.

[4] Bergemann, D. and J. Välimäki (2003), "Dynamic Common Agency," *Journal of Economic Theory*, 111, 23 - 48.

[5] ——————————— (2010), "The Dynamic Pivot Mechanism," *Econometrica*, 78(2), 771-789.

[6] Brandenburger, A. and E. Dekel (1993), "Hierarchies of beliefs and common knowledge," *Journal of Economic Theory*, 59, 189 198.

[7] Dasgupta A. and S. Ghosh (2016), "Repeated Games with Public Randomisation: A Self-accessible approach," Boston University working paper.

[8] Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting Or with Incomplete Information," *Econometrica*, 54(3), 533-554.

[9] ——————————— (1991), "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory* 53(2), 428-438.

[10] Guo, Y. and J. Hörner (2018), "Dynamic Allocation without Money," Working paper, Northwestern University.

[11] Hörner, J., Takahashi, S., and N. Vielle (2015), "Truthful Equilibria in Dynamic Bayesian Games," *Econometrica*, 83, 1795-1848.

[12] Jackson, M. O. and H. F. Sonnenschein (2007), "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75, 241-257

[13] Levin, J. (2003), "Relational Incentive Contracts," *American Economic Review*, 93, 835-857.

[14] Mailath, G. (2020), *Modeling Strategic Behavior: A Graduate Introduction to Game Theory and Mechanism Design*. https://cpb-us-w2.wpmucdn.com/ web.sas.upenn.edu/dist/1/313/files/2017/04/Modeling_strategic_behavior_9_16-2020.pdf

[15] McAfee, R P. (1993), "Mechanism Design by Competing Sellers," *Econometrica*, 61(6), 1281-1312.

[16] Mertens, J. and S. Zamir (1985), "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1-29.

[17] Ortner, J. and S. Chassang (2018), "Making Corruption Harder: Asymmetric Information, Collusion and Crime," *Journal of Political Economy*, 126, 2108-2133.

[18] Pavan, A., I. Segal, and J. Toikka (2014), "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82(2), 601-653.

[19] Sugaya, T. and A. Wolitzky (2020), "The Revelation Principle in Multistage Games," *Review of Economic Studies*, forthcoming.

[20] Watson, J. (2017), "A General, Practicable Definition of Perfect Bayesian Equilibrium," University of California San Diego, Working paper

A Proof of Theorem 2

Fix any (correlated) SCF $f^* \in \mathcal{F}^C(\mu)$ that yields v_j as player j's expected payoff for $j \in \mathcal{I} \cup \mathcal{J}$. This is the target equilibrium payoff for player j. Following Fudenberg and Maskin (1986), we choose a vector of payoffs

$\left(v'_1, \ldots, v'_J\right)$, with $\underline{w}_j^C < v'_j < v_j$ for all $j \in \mathcal{I} \cup \mathcal{J}$. The full dimensionality assumption ensures that for each j, there exists $\epsilon > 0$ and a SCF in $\mathcal{F}^C(\mu)$ that yields expected payoffs

$$\beta^j := \left(\beta_i^j\right)_{i=1}^{i=I+J} = \left(v'_1 + \epsilon, \ldots, v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, \ldots, v'_{J+I} + \epsilon\right) \tag{21}$$

such that ϵ is small enough to satisfy $\beta_i^j < v_j$ for $i \neq j$. Strategies are defined by the following rules.

1. Play starts in phase I. Suppose that $f = (f_1, \ldots, f_J)$ in the support of $f^* \in \mathcal{F}^C(\mu)$ is a SCF that needs to be supported given the realization of the PCD. Each principal j offers the DM $\pi_j = f_j$. Agents report the actual type θ_i^t to all principals at time t. If principal j deviates unilaterally (offers a mechanism other than π_j), agents play a one-shot continuation equilibrium in the current period; play moves to phase II$_j$ from the next period. If agent i's deviation from truthful type reporting is detected, move to phase II$_i$ from the next period.

2. Let us explain phase II. Phase II$_j$ proceeds as follows for $j \in \mathcal{J}$. For any $\gamma_j \in \Gamma_j$ offered by principal j, agents send messages to j to induce the action $g_j\left(\gamma_j\right)$ irrespective of their types. Each principal $k \neq j$ offers the EDM ξ_k^j that assigns the DM $\varphi_k^j(\alpha_j)$ if a majority of agents report α_j. Agents report the true types and $g_j\left(\gamma_j\right)$ to each principal $k \neq j$ at time t.

   Phase II$_i$ proceeds as follows for $i \in \mathcal{I}$. Principals offer the profile of DMs $\pi^i = \left(\pi_1^i, \ldots \pi_J^i\right)$ that attains $\underline{w}_i^C$ of agent i.

   If any player ℓ deviates and he/she is detected as the unique deviator while in phase II$_i$ for $i \in \mathcal{I} \cup \mathcal{J}$, start phase II$_\ell$. If an agent deviates but no agent can be identified as the unique deviator, start phase II$_\ell$ for all $\ell \in \mathcal{I}$ with probability $1/I$ according to the realization of the PCD (See footnote 11). If there is no deviation in II$_i$ for $i \in \mathcal{I} \cup \mathcal{J}$, switch to phase III$_i$ with probability $1 - q \in (0, 1)$ independently across time after each period spent in phase II_i, where q is determined based on the realization of the PCD.

3. In phase III_i, for $i \in \mathcal{I} \cup \mathcal{J}$, pick a (correlated) SCF $\widetilde{f} \in \triangle\mathcal{F}_0^C$ based on the realization of the PCD, which yields expected payoff vector β^i; principal j offers the DM $\pi_i = f_i$ when a SCF $f = (f_1, \ldots, f_J)$ in the support of $\widetilde{f}$ is supposed to be implemented give the realization of the PCD. Agents report their types truthfully. Remain forever in this phase, unless any player ℓ deviates unilaterally and triggers phase II_ℓ.

VERIFICATION OF EQUILIBRIUM:

By the one-shot deviation principle, it suffices to show that the proposed strategy is unimprovable, i.e. no one-shot deviation by any player i from any phase is profitable. Let L^i_j denote player j's expected utility from the beginning of phase II_i without deviation. First, with $j = i$, player i's lifetime (discounted average) payoff in phase II_i is defined recursively as $L^i_i = (1-\delta)\underline{w}^C_i + \delta(qL^i_i + (1-q)\beta^i_i)$, so that

$$L^i_i = \frac{(1-\delta)\underline{w}^C_i + \delta(1-q)\beta^i_i}{1-\delta q}. \tag{22}$$

Note that $L^i_i \to \beta^i_i = v'_i$ as $\delta \to 1$. When calculating L^i_j for $j \neq i$, note that it is bounded on both sides as follows:

$$(1-\delta)\,(-\overline{u}) + \delta(qL^i_j + (1-q)\beta^i_j) \leq L^i_j \leq (1-\delta)\,\overline{u} + \delta(qL^i_j + (1-q)\beta^i_j).$$

As $\delta \to 1$, it is easy to check that $L^i_j \to \beta^i_j = v'_j + \epsilon$.

1. Phase III_i for $i \in \mathcal{I} \cup \mathcal{J}$: From the definitions, it is clear that the difference in the lifetime payoffs to one-shot deviation and conformity cannot be greater than

$$(1-\delta)\,\overline{u} + \delta[(1-p)\beta^i_i + pL^i_i] - \beta^i_i = (1-\delta)\left[\overline{u} - \frac{(1+\delta p - \delta q)\beta^i_i - \delta p \underline{w}^C_i}{1-\delta q}\right] \tag{23}$$

using (22). An immediate implication of inequality (19) defining q is that (23) is strictly negative for all δ close to 1, so that i cannot profitably deviate from Phase III$_i$. Since $\beta^i_j > \beta^j_j \; \forall j \neq i$, it is immediate that, for such δ satisfying $(1-\delta)\,\overline{u} + \delta[(1-p)\beta^i_i + pL^i_i] - \beta^i_i < 0$, players $j \neq i$ do not have a profitable one-shot deviation either from phase III$_i$.

2. Phase II$_j$ for $j \in \mathcal{J}$: This is the phase where principal j is punished. Given j's mechanism γ_j, suppose that agent i deviates to induce an action other than $g_j(\gamma_j)$. Then, move to phase II$_k$ for all $k \in \mathcal{I}$ with probability $1/I$. Agent i will not deviate to induce an action from γ_j other than $g_j(\gamma_j)$ if the following condition is satisfied

$$(1-\delta)\,\overline{u} + \delta\frac{1}{I}\sum_{k=J+1}^{J+I} L^k_i \leq L^j_i.$$

This condition is satisfied with strict inequality as $\delta \to 1$ because the left-hand and right hand-side approach $v'_i + \frac{I-1}{I}\epsilon$ and $v'_i + \epsilon$ respectively as $\delta \to 1$. Agent i also will not deviate from truthful type reporting to non-deviating principals who offer EDMs if

$$(1-\delta)\,\overline{u} + \delta[(1-p)L^j_i + pL^i_i] \leq L^j_i. \tag{24}$$

This condition is satisfied with strict inequality as $\delta \to 1$ because the left-hand and right hand-side approach $v'_i + (1-p)\epsilon$ and $v'_i + \epsilon$ respectively as $\delta \to 1$. Given the majority rule employed in non-deviating principals' EDMs, agent i also has no incentive to deviate from truthful reporting of principal j's action to non-deviating principals when all agents report principal j's true action. Principal $i \neq j$ will not deviate if

$$(1-\delta)\,\overline{u} + \delta L^i_i \leq L^j_i, \tag{25}$$

which is clearly satisfied as $\delta \to 1$ because the left-hand and right hand-side approach v'_i and $v'_i + \epsilon$ respectively as $\delta \to 1$. Principal j

will choose his own mechanism that best respond to others' EDMs and agents' communication protocol.

3. Phase II_j for $j \in \mathcal{I}$: This is the phase where principals offer a profile of DMs that are UIC for agent j but CIC for the other agents in order to punish agent j. Since UIC is imposed for agent j, there is no need for deviation by agent j. Agent i ($i \neq j$) will not deviate if $(1-\delta)\,\overline{u} + \delta[(1-p)L_i^j + pL_i^i] \leq L_i^j$. Similar to (24), this is satisfied as $\delta \to 1$. Principal $i \neq j$ will not deviate if $(1-\delta)\,\overline{u} + \delta L_i^i \leq L_i^j$. Similar to (25), this is also satisfied as $\delta \to 1$.

4. Phase I: This is on the equilibrium path where principals offer a profile of CIC DMs given the realization of the PCD. Principal j will not deviate if $(1-\delta)\,\overline{u} + \delta L_j^j \leq v_j$, which is satisfied because, as $\delta \to 1$, the left-hand approaches v_j' that is less than v_j. Agent j will not deviate if $(1-\delta)\,\overline{u} + \delta[(1-p)v_j + pL_j^j] \leq v_j$, which is also satisfied because, as $\delta \to 1$, the left-hand approaches $(1-p)v_j + pv_j'$ that is less than v_j.

In sum, for high δ, the posited strategy profile is unimprovable after all histories, and hence is an equilibrium.