

# Repeated Oligopoly with Contracts\*

Sambuddha Ghosh<sup>†</sup>      Seungjin Han<sup>‡</sup>

December 14, 2023

## Abstract

We model dynamic contracting competition, where multiple principals (manufacturers) can simultaneously and repeatedly offer complex short-term contracts to multiple agents (retailers). As each principal's contract can incorporate other contracts as inputs, the complexity of contracts escalates, leading to the "infinite regress problem" (McAfee, 1993). We demonstrate that, when players exhibit sufficient patience, the limit equilibrium utility set is determined in terms of actions and direct mechanisms, even though any kind of complex contracts are allowed. We identify simple contracts that principals can use without loss of generality in any path of the game to support any equilibrium payoff profile.

---

\*This paper subsumes and expands on earlier working papers by the same authors on this broad area. We thank Madhav Chandrasekher, Ani Dasgupta, Yuk-fai Fong, Ying Lei, Maxim Ivanov, Rummen Kostadinov, Bart Lipman, Albert Ma, 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.

<sup>†</sup>School of Economics, Shanghai University of Finance and Economics, China Email: sghosh@mail.shufe.edu.cn

<sup>‡</sup>Department of Economics, McMaster University, Canada. Email: hansj@mcmaster.ca

# 1 Introduction

In a repeated duopoly, each of two manufacturers produces a good; these are distinct (or differentiated) goods. These manufacturers must sell their goods to retailers in wholesale markets, and then retailers sell to consumers. Each manufacturer offers to sell to one or more retailers at certain prices, which may differ across retailers. Then each retailer charges a predetermined markup to sell to consumers. The demands for the goods are unknown to the manufacturers, but observed by retailers. The scenario is repeated over time.

Clearly each manufacturer would like to have the retailers' demand information. In the standard textbook version the price vector offered by each manufacturer cannot respond to the demand shock. But this seems like an odd limitation. It seems natural to allow each manufacturer to ask the retailers to send a message, and condition the selling prices on it. Since all retailers know the demand shock, it should be possible to elicit it—everyone reports truthfully because everyone else does. This seemingly natural model can be quite intractable without *a priori* restrictions on what selling contracts are allowed.

In his paper on competing auctions in a dynamic setting, McAfee (1993) says “It is also quite difficult to define the large strategy space since mechanisms must map the set of mechanisms into outcomes, ...” Unlike single-principal models, private information is partly *endogenous* in models with multiple principals because, in addition to any exogenous private information, agents also observe the contract or terms of trade offered by the various principals, each of whom knows only his own contract. It seems natural to then allow a principal to write contracts that ask agents to report all their private information including the information about other principals' contracts. However, the crux is that reporting about a complicated contract requires one that is yet more

complicated; the complexity of such contracts starts to blow up quickly—this is what McAfee refers to as the ‘infinite regress problem’.

We study such contracting competition when it is repeated frequently. Multiple principals compete over the infinite horizon by repeatedly offering non-exclusive short-term contracts to multiple agents who possess information that the principals want. In every period each principal offers a contract (or mechanism) specifying that principal’s current action as a function of the profile of contemporaneous private messages she receives from the agents. Thus principals can write contracts and take actions, while agents have information that they can choose to reveal through the use of private messages. When all players are long-lived, we are interested in identifying the class of contracts that principals can use without loss of generality, the profiles of utilities that can be supported in equilibrium, and the strategies needed to do so. Our message is that it is possible to study these tractably, when the game is repeated frequently enough. There is no need for exogenous restrictions on the complexity of communication that short-term contracts permit, a noteworthy feature that distinguishes our work from several pioneering works such as McAfee (1993).

This paper shows that the infinite regress problem of McAfee does not arise when principals and agents are patient and interact repeatedly. The equilibrium utility set is determined *in terms of model primitives* (actions and direct mechanisms) (See Theorem 1.1 (Equilibrium Characterization)). Although we allow a broad range of contracts, principals can without loss of generality use simple contracts on and off the equilibrium path, to support any equilibrium payoff profile (See Theorem 1.2 (Extended Revelation Principle)); this is ‘without loss of generality’ in the sense that a principal cannot gain by unilaterally deviating to richer contracts, nor can he punish another player more severely with richer contracts.

Our key result shows that the minmax payoff of manufacturer  $j$  in this game equals

the solution to a simpler problem: Find the maxmin payoff of  $j$  in a simpler game where  $j$  first picks a price vector and then the other manufacturer picks her price to lower  $j$ 's profit as far as possible, given retailers' truthful reports on the demand parameters and  $j$ 's price vector. As players become patient, any allocation that generates more than this cutoff for each manufacturer (and zero payoffs for the retailers) can be supported in a perfect Bayesian equilibrium of the discounted repeated game.

Our equilibrium characterization shares standard elements in the literature on repeated games, including the full dimensionality assumption due to Fudenberg and Maskin (1986)). However, we must grapple with additional issues: (i) calculating the worst equilibrium utility for principals (which is non-standard because the principal's ability to offer any complex short-term contracts (mechanisms) gives him some commitment power and potentially raises his utility when he is being punished), and (ii) preventing deviations by agents that cannot be identified as coming from a specific agent. The solution to the first problem is to have agents give no information to a principal who is being punished. The solution to the second is to restrict attention to constrained incentive compatible (CIC) mechanisms (i.e., mechanisms where undetectable agent-deviations are not profitable).

We show that all utilities 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 utility cutoffs (or minmax values) in our world are not easily specified in closed form. By the virtue of the extended revelation principle, we manage to do so by developing an algorithm to compute the unrestricted minmax value<sup>1</sup> as the maxmin value of a much simpler game where the only contracts are constant-action contracts, direct mechanisms (DMs) and a slight extension of DMs. No such

---

<sup>1</sup>That is, without exogenous restrictions on the complexity of mechanisms.

algorithm exists in the literature. Second, while folk theorems are about utilities, 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.

Section 2 shows the key ideas of our main results and why the Revelation Principle fails when all information is commonly known to agents. Section 3 sets up the model. Sections 4 and 5 formulate the notion of CIC and provide the main results respectively when agents have private information about their payoff type. Our main result, Theorem 1 is presented in Section 5.1 for the model with three or more agents when mechanisms and actions are *observable* by every player at the end of each period.

Corollary 2 in Section 5.2 extends Theorem 1 for the model with two agents. Corollaries 3 and 4 in Section 5.3 extend the results when principals do not observe the other principals' mechanisms and actions whereas agents observe mechanisms with probability one and actions with some positive probability. Section 5.4 discuss the possibility of relaxing CIC.

**Related literature** There has been a growing interest in designing dynamic mechanisms, with a flurry of research in the last decade. These models study mechanism design by a single principal (i.e., mechanism designer or mediator) who has full commitment and can thus offer a contract spanning the entire horizon of the game—an infinite horizon in Bergemann and Välimäki (2010), Athey and Segal (2013), Pavan, Segal, and Toikka (2014), and Guo and Hörner (2018); and a finite one in Battaglini and Lamba (2019), and Sugaya and Wolitzky (2021).

Baker, Gibbons, Murphy (1994) and Pearce and Stacchetti (1998) are pioneering works studying the interaction of explicit (say specifying the wage as a function of output) and implicit (through incentives provided by repetition) short-term contracts. Their models do not consider competition among principals in the space of contracts

and the contracts they consider are restricted to a very simple class. Much of the difficulty we encounter comes from the fact that we do not limit how many ‘levels’ a contract may have.

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, and Vielle (2015). The object of their study is not the structure of actions because these are simple (unlike contracts, which are the ‘actions’ that principals choose in our model); rather, they are interested in the limiting utility set of truthful public perfect equilibria given a fixed action space, abstracting away from contracting possibilities. We bring contracting processes between multiple principals and multiple agents to the forefront of equilibrium analysis.

## 2 Common information: Repeated duopoly

This section presents the key ideas in the context of a model of repeated duopoly, when all information is commonly known among agents. The structure, although simple, suffices to bring out the complexity of competition when we do not restrict the contracting space in an *ad hoc* way.

At each  $t \in \mathbb{N}$  two manufacturers, 1 and 2, can each produce a differentiated non-storable product at zero marginal cost; each period’s product is sold in wholesale markets to retailers indexed by  $i \in \mathcal{I}$ . In this application, we assume that there are three or more retailers, i.e. the index set  $\mathcal{I}$  has at least size 3. Manufacturers and retailers have zero reservation profits. Retailers sell the products in retail markets operating *à 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

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

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

Manufacturers and retailers are all long lived and they maximize the discounted sum of their respective per-period profits using a common discount factor  $\delta$ . We assume that there is a public correlation device available at the start of each period to permit (independent) randomization over mechanisms.

We now describe sales in the wholesale market from manufacturers to retailers. At the beginning of each period, each manufacturer  $j$  offers a one-period contract. Any contract comprises a set  $M_{ij}$  of messages that retailer  $i$  can send to manufacturer  $j$ , with  $M_j$  defined as  $\times_{i \in \mathcal{I}} M_{ij}$ . No restrictions are imposed on the complexity of the message spaces and what information it can contain. Let  $\gamma_{ij} : M_j \rightarrow \mathbb{R}_+$  be the price paid by retailer  $i$  to manufacturer  $j$  as a function of the messages sent by all agents to  $j$ . Let  $\Gamma_j$  be the set of all possible contracts for manufacturer  $j$ . Any contract in this set 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 retailers can buy any amount at the offered prices. Motivated by industry practices, we also assume that retailers sell the product at the price that incorporates a standard industry *mark-*

up  $\Delta$ , say  $\Delta = 0.1$  (i.e., 10%), on the wholesale price.<sup>2</sup> Here is our assumption on information: *Contracts and actions (i.e., prices) become public information at the end of each period, but messages and states are never revealed.*

## 2.1 Equilibrium characterization

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  $j$ ? We need to know this to derive  $j$ 's worst equilibrium utility  $\underline{w}_j$ , which in turn is key to deterring deviations by  $j$ . 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 utility is equal to the maximum utility 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 *à la* Bertrand. Given  $\mathbf{p}_j$ , let

$$p_j^\circ(\mathbf{p}_j) := \min \{p_{ij} : i \in \mathcal{I}\}.$$

Suppose that each manufacturer  $\ell \neq j$  asks each agent to report just two pieces of information—(i)  $j$ 's action, and (ii) the state of the local market, i.e. the value of  $\theta$ .

---

<sup>2</sup>For simplicity, we take the standard industry mark up as given but this too can be incorporated into the equilibrium outcome.



If more than a half of retailers report the same action  $\mathbf{p}_j$  and state  $\theta$ , manufacturer  $\ell$  takes  $\{\mathbf{p}_j, \theta\}$  as manufacturer  $j$ 's true action and state.<sup>3</sup> Since  $\mathbf{p}_j$  and  $\theta$  are fully observable to retailers and there are three or more retailers, reporting truthfully is an equilibrium with majority rule. We refer to these as **extended direct mechanisms** with majority rule.

Then, for every  $\{\mathbf{p}_j, \theta\}$ , the worst punishment manufacturer  $\ell$  can inflict on  $j$  is to choose  $p_{i\ell} = p_\ell^*(\mathbf{p}_j, \theta)$  for all  $i \in \mathcal{I}^4$  to minimize the demand for product  $j$ , i.e.,

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

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  $\ell$  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 price, then the maximum utility he could achieve when he was punished would be

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

---

<sup>3</sup>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  $\ell$ 's extended direct mechanism can ask retailers to report  $p_j^\circ(\mathbf{p}_j) = \min \{p_{ij} : i \in \mathcal{I}\}$ .

<sup>4</sup>Since every retailer can purchase product  $\ell$  at the same price from manufacturer  $\ell$ , they supply equal quantities to the market for product  $\ell$ .

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

$$\underline{w}_j = \max_{p_j} \left\{ p_j \times \mathbb{E}_\theta \left[ \min_{p_\ell} D_j((1 + \Delta) p_\ell, (1 + \Delta) p_j, \theta) \right] \right\} \quad (2)$$

$$= \max_{p_j} \left\{ p_j \times \left[ 1 - 2.2p_j - \frac{3.63}{8} p_j^2 \right] \right\} \quad (3)$$

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, manufacturer  $j$ 's utility cannot go below  $\underline{w}_j$  in a phase where he is punished when he can use any arbitrary contract in  $\Gamma_j$ . How much higher can it be? This is a hard question to answer directly as we cannot simply list all possible contracts. Our Theorem 1 shows 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.

In this application, we highlight some of the key ideas that permit such a simplification. Suppose that manufacturer  $j$  offers any arbitrary contract at the beginning of a period when he is being punished. Given such a contract  $\gamma_j = \{(\gamma_{ij})_{i \in \mathcal{I}}\}$ , let

$$\gamma_j(M_j) := \{(\gamma_{ij}(m_j))_{i \in \mathcal{I}} : m_j \in M_j\}$$

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

where

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

Such communication behavior *always* induces  $\mathbf{p}_j^\times(\gamma_j)$  from  $\gamma_j$ , so for manufacturer  $j$ , offering contract  $\gamma_j$  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  $\theta$  to manufacturer  $\ell$  whose extended direct mechanism then chooses  $p_{i\ell} = p_\ell^*(\mathbf{p}_j^\times(\gamma_j), \theta)$  for all  $i \in \mathcal{I}$ .

Such a communication is enforceable if retailers are sufficiently patient ( $\delta \rightarrow 1$ ). Because the contract and action are observable at the end of the period, the other manufacturers observe manufacturer  $j$ 's contract  $\gamma_j$  and action  $p_j$  at the end of period. If  $p_j \neq p_j^\times(\gamma_j)$ , then they know that at least one retailer has deviated from such communication. This retailer can then be punished in the repeated game: manufacturers charge so high a price to this retailer that she cannot sell any products. This suffices to deter deviation by retailers from a phase in which a manufacturer is being punished. Even when the identity of the deviating retailer is unknown, manufacturers can join forces to punish each retailer with equal probability; each retailer faces a positive probability of loss of the opportunity to make positive profits based on mark-up.

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 utility can be lowered to  $\underline{w}_j$  in a phase where he is punished even if he can use any complex contracts. Importantly retailers do not need to describe  $\gamma_j$  directly but only  $j$ 's action to manufacturer  $\ell$ , so extended DMs are free from the infinite regress problem.

## 2.2 Failure of the Revelation Principle

What would be each manufacturer's worst equilibrium payoff when manufacturers are restricted to offer only direct mechanisms as in McAfee (1993)? Manufacturer  $j$ 's direct mechanism is  $\pi_j := \{(\pi_{ij})_{i \in \mathcal{I}}\}$  with  $\pi_{ij} : \{1, 2\}^3 \rightarrow \mathbb{R}_+$  specifying the price for retailer  $i$  as a function of all retailers' messages on the state.

Suppose that manufacturer  $j$  offers a direct mechanism in the phase where he is punished for his past deviation. Applying the logic described earlier, if retailers are sufficiently patient, they can be enforced to block the transmission of any information on the state so that they always induce the same price vector from manufacturer  $j$ 's direct mechanism regardless of the state. Therefore, manufacturer  $j$  cannot do any better with a direct mechanism than he does with a fixed price.

Since there are three retailers, the incentive compatibility has no bite and hence the state can be revealed truthfully when the strict majority rule is applied to a direct mechanism. Therefore, manufacturer  $\ell$  can implement any price vector  $\mathbf{p}_\ell = (p_\ell^1, p_\ell^2)$  for every retailer with a direct mechanism, where  $p_\ell^\theta$  is  $\ell$ 's price in state  $\theta \in \{1, 2\}$ .

Given manufacturer  $\ell$ 's price vector  $\mathbf{p}_\ell = (p_\ell^1, p_\ell^2)$ , manufacturer  $j$  chooses a fixed price  $p_j(\mathbf{p}_\ell)$  that maximizes his expected profit in the phase where he is punished:

$$p_j(\mathbf{p}_\ell) \in \arg \max_{p_j \in \mathbb{R}_+} p_j \times \mathbb{E}_\theta [D_j((1 + \Delta) p_\ell^\theta, (1 + \Delta) p_j, \theta)]$$

Then, manufacturer  $\ell$  chooses  $\mathbf{p}_\ell = (p_\ell^1, p_\ell^2)$  that minimizes

$$p_j(\mathbf{p}_\ell) \times \mathbb{E}_\theta [D_j((1 + \Delta) p_\ell^\theta, (1 + \Delta) p_j(\mathbf{p}_\ell), \theta)] .$$

In this way, we can derive the worst equilibrium profit for principal  $j$ :

$$\underline{w}_j^d = \min_{(p_\ell^1, p_\ell^2) \in \mathbb{R}_+^2} \max_{p_j \in \mathbb{R}_+} p_j \times \mathbb{E}_\theta [D_j((1 + \Delta) p_\ell^\theta, (1 + \Delta) p_j, \theta)] = 0.1111$$

$\underline{w}_j^d$  is reached when manufacturer  $\ell$ 's choose the direct mechanism that implements  $(p_\ell^1, p_\ell^2) = (0.12089, 0.060449)$  by retailers' truth telling, and manufacturer  $j$  best respond to it by choosing  $p_j = 0.2198$ .

Using the standard argument of the folk theorem, one can show that any SCF  $f \in F$  is the outcome of a PBE of  $(G, \Pi)^\infty(\delta)$  for high  $\delta$  if, for all  $j \in \{1, 2\}$  and  $\ell \neq j$ ,

$$\mathbb{E}_\theta [D_j((1 + \Delta) f_j(\theta), (1 + \Delta) f_\ell(\theta), \theta) \times f_j(\theta)] > \underline{w}_j^d.$$

As one can see, manufacturer  $j$ 's worst equilibrium payoff  $\underline{w}_j^d = 0.1111$  in competing direct mechanisms is strictly higher than his worst equilibrium payoff  $\underline{w}_j = 0.1088$  in competing complex mechanisms. This implies that the Revelation Principle does not hold. The failure of the Revelation Principle is not caused by the restriction to direct mechanisms off the equilibrium path because a deviating manufacturer cannot do any better with a direct mechanism or any more complex mechanism than he does with a fixed action. It fails because the non-deviating manufacturer cannot punish the deviating manufacturer with a direct mechanism as severely as he does with an extended direct mechanism. Why? With a direct mechanism, the non-deviating manufacturer condition the implementation of his price only on the state but not on the deviating manufacturer's price.

If manufacturers are restricted to offer only constant contracts (i.e., fixed prices),

then manufacturer  $j$ 's worst equilibrium payoff  $w_j^\circ$  is even higher than  $\underline{w}_j^d = 0.1111$ :

$$\underline{w}_j^\circ = \min_{p_\ell \in \mathbb{R}_+} \max_{p_j \in \mathbb{R}_+} p_j \mathbb{E}_\theta [D_j((1 + \Delta)p_\ell, (1 + \Delta)p_j, \theta)] = 0.1114. \quad (4)$$

When manufacturer  $\ell$  uses a constant contract, he does not use any information retailers have when punishing manufacturer  $j$ . The more information the non-deviating manufacturer can exploit from retailers, the more severely he can punish the deviating manufacturer ( $\underline{w}_j^\circ > \underline{w}_j^d > \underline{w}_j$ ). However,  $\underline{w}_j$  is the lowest possible equilibrium payoff even if there are no restrictions on mechanisms manufacturers can use.

### 3 A model of repeated contracting competition

Motivated by the duopoly model, we now offer a three-stage description of a more general model of contracting competition over time. First, we describe the **underlying game**, which doesn't include contracts. Second, we add contracts to it to obtain the **stage game**. Third, we describe the **repeated game** and, in particular, what is observable.

**Underlying game.** Principals and agents, respectively, comprise the sets  $\mathcal{J} := \{1, \dots, J\}$  and  $\mathcal{I} := \{J + 1, \dots, J + I\}$ . We assume that there are multiple principals and multiple agents (i.e.,  $J \geq 2$  and  $I \geq 2$ ). In most parts we assume that  $I \geq 3$ , but Section 5.2 shows that our results go through with  $I = 2$ . Each principal  $j$  has a finite set<sup>5</sup>  $\mathcal{A}_j$  of actions, with typical action  $\alpha_j$ . A profile of actions is  $\alpha = (\alpha_1, \dots, \alpha_J) \in \mathcal{A} := \times_{j \in \mathcal{J}} \mathcal{A}_j$  and  $\mathcal{A}_{-j} := \times_{k \neq j} \mathcal{A}_k$ .

Each agent  $i$  is informed about her type  $\theta_i$  drawn from a finite set  $\Theta_i$  according to

---

<sup>5</sup>Finite type and action spaces are not critical for our results: With a modicum of technicalities we can deal with a compact set of actions and a countable type-space.

a distribution  $\mu_i$ ; the profile of types  $\theta = (\theta_{J+1}, \dots, \theta_{J+I})$  is drawn from  $\Theta := \times_{i \in \mathcal{I}} \Theta_i$  according to the joint distribution  $\mu = \times_{i \in \mathcal{I}} \mu_i$ . If types are perfectly correlated (i.e.,  $\theta_i = \theta_{i'}$  for all  $i, i' \in \mathcal{I}$ ), then it is the case of **common information**, i.e. complete information among agents; otherwise we say the model is one of **private information**. In both models principals have no information about the type(s). In our example with manufacturers and retailers,  $\theta_i$  is the private information that retailer  $i$  knows about the state in the retail market.

The utility function for player  $\ell$  (principal or agent) is  $u_\ell : \mathcal{A} \times \Theta \rightarrow \mathbb{R}$ ; utilities are uniformly bounded by  $\bar{u} < \infty$ , i.e.  $|u_\ell(\alpha, \theta)| < \bar{u}$  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:

$$G := \left( \mathcal{J}; \mathcal{I}; (\mathcal{A}_j)_{j \in \mathcal{J}}; (\Theta_i)_{i \in \mathcal{I}}; (\mu_i)_{i \in \mathcal{I}}; (u_\ell)_{\ell \in \mathcal{I} \cup \mathcal{J}} \right). \quad (5)$$

**Stage game.** Fix an underlying game  $G$  as in (5), and for each  $j \in \mathcal{J}$  fix a collection of sets  $\{M_{ij} | i \in \mathcal{I}\}$  and a set  $\Gamma_j$  that comprises continuous mappings  $\gamma_j$  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  $j$ . The set of mechanisms available to principal  $j$  is  $\Gamma_j$ , with typical element  $\gamma_j$ . We assume that  $\Gamma_j$  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: Mechanisms in  $\Gamma_j$  may allow agents to report not only their own types, but also mechanisms offered by the other principals, and so on.

The **stage game**  $(G, \Gamma)$  is the game with the following timing of moves:

1. Each principal  $j$  simultaneously offers a mechanism  $\gamma_j$  from  $\Gamma_j$ .
2. After observing the profile of mechanisms  $\gamma = (\gamma_1, \dots, \gamma_J)$  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 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 utility  $u_\ell(\gamma_1(m_1), \dots, \gamma_J(m_J), \theta)$ .

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.

A **direct mechanism (DM)**  $\pi_j : \Theta \rightarrow \mathcal{A}_j$  for principal  $j$  is a special mechanism where the set of messages that each agent  $i$  can send to principal  $j$  is simply  $i$ 's type space  $\Theta_i$ . Let  $\Pi_j$  be the set of all possible direct mechanisms for principal  $j$  and  $\Pi := \times_{j \in \mathcal{J}} \Pi_j$ . We assume that  $\Gamma_j$  is bigger than  $\Pi_j$  ( $\Gamma_j \succ \Pi_j$ ) for all  $j \in \mathcal{J}$ .<sup>6</sup>

**Repeated game.** We now describe the infinitely repeated game  $(G, \Gamma)^\infty(\delta)$ .<sup>7</sup> It involves playing the stage-game  $(G, \Gamma)$  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  $\mu_i$  independently of all past types and the current types of the other agents (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 correlation device (PCD) produces a continuous signal  $\omega^t \in [0, 1]$  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.<sup>8</sup> So, principal  $j$  can offer

---

<sup>6</sup>Formally,  $\Gamma_j \succ \Pi_j$  if there exists an embedding  $\eta_j : \Pi_j \rightarrow \Gamma_j$ . It implies that there are more mechanisms in  $\Gamma_j$  than in  $\Pi_j$ . We will also assume later that  $\Gamma_j$  is bigger than the set of principal  $j$ 's extended DMs that ask agents to report their types and the punished principal's actions.

<sup>7</sup>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. When it does not create confusion, we drop the superscript  $t$  for notational simplicity.

<sup>8</sup>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)).



a mechanism  $\gamma_j^t \in \Gamma_j$  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 5.3 for details). Therefore, the only substantive assumption is that agents learn the actions taken given a profile of mechanisms offered by principals.

Starting with the null history  $h^0$ , agent  $i$ 's period- $t$  history  $h_i^t$  is constructed from her period- $(t-1)$  history  $h_i^{t-1}$  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 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 utility of player  $\ell \in \mathcal{I} \cup \mathcal{J}$  from period  $\tau$  onwards is  $(1 - \delta) \sum_{t \geq \tau} \delta^{t-\tau} u_\ell(\alpha^t, \theta^t)$ .

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.<sup>9</sup> Furthermore all players share a common belief about any other player, and no player signals what he does not know.

### 3.1 Some comments on the model

First, why do we consider only short-term contracts? From a technical standpoint nothing would change if we were to allow finite multi-period contracts. So effectively what we are doing is to rule out infinite-horizon contracts. Since contracts are ultimately

---

<sup>9</sup>Mailath (2020) calls this "almost perfect Bayesian equilibrium".

based on laws, which themselves change in response to other forces (political, socio-logical 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.

Second, what are actions in our model? This depends on the 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 competing auctions, where each seller  $j$  is endowed with one unit of a good in each period, seller  $j$ 's action is a profile of pairs of monetary transfers and the probability of winning the object  $a_j = [(t_{ij})_{i \in \mathcal{I}}, (q_{ij})_{i \in \mathcal{I}}]$ , one for each bidder  $i$  with  $\sum_{i \in \mathcal{I}} q_{ij} = 1$ . In a 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 pairs of payments and quantities, one for each supplier  $i$ . In loan contracting, the bank  $j$ 's action is  $a_j = [(p_{ij})_{i \in \mathcal{I}}, (q_{ij})_{i \in \mathcal{I}}]$ , a profile of pairs of loans and repayments conditional on a project's success, one for each entrepreneur  $i$ .

Note that we do not explicitly model agent  $i$ 's effort. However, depending on the application, it is included in the notion of the principal's actions. For example, if an agent's effort  $x_i$  is decomposable with respect to principals, i.e.,  $x_i = (x_{i1}, \dots, x_{iJ})$  and is also contractible, then  $(x_{ij})_{i \in \mathcal{I}}$  can be incorporated into the model as part of principal  $j$ 's action. For example, in the problem with final producers and intermediate good suppliers or in loan contracting,  $(q_{ij})_{i \in \mathcal{I}}$  can be thought of as the profile of principal  $j$  specific components of agents' effort. This saves us the trouble of explicitly modelling an agent's effort

## 4 Incentive compatibility

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  $\pi = (\pi_1, \dots, \pi_J)$ , the expected stage-game utility of agent  $i$  of type  $\theta_i$  who reports  $\theta_{ij}$  to each principal  $j$ , when the other agents report truthfully, is

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

where  $\mathbb{E}_{\mu_{-i}}$  is the expectation operator with respect to the probability distribution  $\mu_{-i}$  over  $\Theta_{-i}$ .

We say that a profile of DMs  $\pi = (\pi_1, \dots, \pi_J)$  satisfies **unconstrained incentive compatibility (UIC)** if for all  $i \in \mathcal{I}$  and all  $\theta = (\theta_i, \theta_{-i}) \in \Theta$ , we have

$$\begin{aligned} \mathbb{E}_{\mu_{-i}} [u_i(\pi(\theta), \theta)] &\geq \\ \mathbb{E}_{\mu_{-i}} [u_i(\pi_1(\theta_{i1}, \theta_{-i}), \dots, \pi_J(\theta_{iJ}, \theta_{-i}), \theta)], \forall (\theta_{i1}, \dots, \theta_{iJ}) \in (\Theta_i)^J, \end{aligned} \quad (6)$$

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

Given a profile of DMs, agents' type messages induce a profile of actions that carries some information about the messages sent by agents. In particular, some type messages where only a single agent lies may be detected when they trigger an off-path action profile. However, in the one-shot game this detection comes too late to punish the deviating agent; this is why UIC is used in the one-shot game to prevent all possible lies. However, in the repeated game punishments for such lies can be deferred. For any agent  $i$  in the repeated game, incentive compatibility thus does not need to be

imposed over type messages by an agent  $i$  that with positive probability lead to a profile of actions that can arise only when  $i$  lies. We illustrate this weaker notion, called Constrained Incentive Compatibility (CIC), using the example below.

**Example 1** *There are four players. Players 1 and 2 are principals and players 3 and 4 are agents. The first agent's type is either  $\theta_3$  or  $\theta'_3$  and the second agent's type is either  $\theta_4$  or  $\theta'_4$ . Suppose that principals 1 and 2 offer DMs  $\pi_1$  and  $\pi_2$  respectively:*

$$\begin{aligned}\pi_1(\theta_3, \theta_4) &= \alpha_1, & \pi_2(\theta_3, \theta_4) &= \alpha'_2, \\ \pi_1(\theta_3, \theta'_4) &= \alpha_1, & \pi_2(\theta_3, \theta'_4) &= \alpha_2, \\ \pi_1(\theta'_3, \theta_4) &= \alpha'_1, & \pi_2(\theta'_3, \theta_4) &= \alpha'_2, \\ \pi_1(\theta'_3, \theta'_4) &= \alpha''_1, & \pi_2(\theta'_3, \theta'_4) &= \alpha'_2.\end{aligned}$$

The table below shows the action profiles induced by agents' type messages. Rows correspond to agent 1's message profiles (one for each principal) whereas columns correspond to agent 2's message profiles.

	$(\theta_4, \theta_4)$	$(\theta_4, \theta'_4)$	$(\theta'_4, \theta_4)$	$(\theta'_4, \theta'_4)$
$(\theta_3, \theta_3)$	$(\alpha_1, \alpha'_2)$	$(\alpha_1, \alpha_2)$	$(\alpha_1, \alpha'_2)$	$(\alpha_1, \alpha_2)$
$(\theta_3, \theta'_3)$	$(\alpha_1, \alpha'_2)$	$(\alpha_1, \alpha_2)$	$(\alpha_1, \alpha'_2)$	$(\alpha_1, \alpha'_2)$
$(\theta'_3, \theta_3)$	$(\alpha'_1, \alpha'_2)$	$(\alpha_1, \alpha_2)$	$(\alpha''_1, \alpha'_2)$	$(\alpha''_1, \alpha_2)$
$(\theta'_3, \theta'_3)$	$(\alpha'_1, \alpha'_2)$	$(\alpha'_1, \alpha'_2)$	$(\alpha''_1, \alpha'_2)$	$(\alpha''_1, \alpha'_2)$

The set of action profiles induced by truthful reports is

$$\mathcal{A}((\pi_1, \pi_2), \emptyset) := \{(\alpha_1, \alpha'_2), (\alpha_1, \alpha_2), (\alpha'_1, \alpha'_2), (\alpha''_1, \alpha'_2)\}.$$

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

Given the truthful type reporting by the first agent (player 3), the second agent's

inconsistent type report  $(\theta_4, \theta'_4)$  induces  $(\alpha_1, \alpha_2)$  and  $(\alpha'_1, \alpha'_2)$  when the first agent's truthful type reports are  $(\theta_3, \theta_3)$  and  $(\theta'_3, \theta'_3)$  respectively. Because  $(\alpha_1, \alpha_2), (\alpha'_1, \alpha'_2) \in \mathcal{A}((\pi_1, \pi_2), \emptyset)$ , the second agent's inconsistent type report  $(\theta_4, \theta'_4)$  cannot be detected by observing the action profiles. Similarly, the second agent's inconsistent type report  $(\theta'_4, \theta_4)$  cannot be detected by observing the action profiles.

Given the truthful type reporting by the second agent (player 4), the first agent's inconsistent type report  $(\theta_3, \theta'_3)$  induces  $(\alpha_1, \alpha'_2)$  for both possible profiles of the second agent's truthful reports  $(\theta_4, \theta_4)$  and  $(\theta'_4, \theta'_4)$ . Because  $(\alpha_1, \alpha'_2) \in \mathcal{A}((\pi_1, \pi_2), \emptyset)$ , the other players cannot detect the first agent's inconsistent type reporting by observing  $(\alpha_1, \alpha'_2)$  at the end of the period. On the other hand, the first agent's inconsistent type report  $(\theta'_3, \theta_3)$  induces  $(\alpha'_1, \alpha'_2)$  and  $(\alpha''_1, \alpha_2)$  when the second agent's truthful type reports are  $(\theta_4, \theta_4)$  and  $(\theta'_4, \theta'_4)$  respectively. Note that  $(\alpha''_1, \alpha_2) \notin \mathcal{A}((\pi_1, \pi_2), \emptyset)$  and it cannot be induced by the second agent's inconsistent type reports. Therefore, if  $(\alpha''_1, \alpha_2)$  is observed at the end of the period, players know that it is the first agent who lied. Then, the first agent can be punished from the next period so that incentive compatibility does not need to be imposed for the first agent's inconsistent type report  $(\theta'_3, \theta_3)$ .

We now crystallize the above example into a more general definition. Given a profile of DMs  $\pi$  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 action 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

$$\mathcal{A}(\pi, i) \setminus (\cup_{k \neq i} \mathcal{A}(\pi, k)) =: \mathcal{A}_L^i(\pi) \quad (7)$$

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

(7). Formally,

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

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  $\pi$ . Note that any consistent message profile  $(\theta_{i1}, \dots, \theta_{iJ})$  of agent  $i$  (i.e.,  $\theta_{i1} = \theta_{i2} = \dots = \theta_{iJ}$ ) is not in  $L_i(\pi)$  even though it is a lie. If a message profile  $(\theta_{i1}, \dots, \theta_{iJ})$  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 show in our equilibrium characterization, 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  $\pi$  satisfies **constrained incentive compatibility (CIC)** under a type distribution  $\mu$  if for all  $i \in \mathcal{I}$  and all  $\theta = (\theta_i, \theta_{-i}) \in \Theta$ , we have*

$$\mathbb{E}_{\mu_{-i}} [u_i(\pi(\theta), \theta)] \geq \mathbb{E}_{\mu_{-i}} [u_i(\pi_1(\theta_{i1}, \theta_{-i}), \dots, \pi_J(\theta_{iJ}, \theta_{-i}), \theta)] \quad \forall (\theta_{i1}, \dots, \theta_{iJ}) \in (\Theta_i)^J \setminus L_i(\pi). \quad (8)$$

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:

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

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

$$\pi_j^t := \gamma_j^t \circ (m_{ij}^t)_{i \in \mathcal{I}}. \quad (9)$$

**Definition 2** *A PBE is called a constrained PBE (cPBE) if the profile of agents' communication strategies induces CIC DMs from the mechanisms offered by principals given any histories of agents.*

At the end of Section 5.1, we explain why CIC is generally needed in the phase where a player is punished. For that reason, we are interested in characterizing the set of cPBE allocations in the model with two or more agents.

Section 5.4 discusses how to dispense with incentive compatibility on the equilibrium path where no one has previously deviated, 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. It also discusses the relation to the approach that provides agent with the incentive compatibility by checking the frequency of messages over a block of periods (e.g., Jackson and Sonnenschein (2007), Renault, Solan, and Vielle (2013)).

## 5 Equilibrium in the general model

This section presents the main results, offering an extended revelation principle as well as an equilibrium characterization.

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  $\gamma$  offered by principals in the phase where principal  $j$  is punished for his past deviation.<sup>10</sup> Given  $\gamma$ ,

---

<sup>10</sup>The equilibrium characterization in Theorem 1 shows that  $\Pi(\gamma)$  is non-empty.

principal  $j$ 's lowest possible stage utility is

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

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

$$\underline{w}_j^C := \min_{\gamma_{-j} \in \Gamma_{-j}} \max_{\gamma_j \in \Gamma_j} \underline{u}_j(\gamma_{-j}, \gamma_j). \quad (10)$$

For agent  $i$ 's worst cPBE utility, note that 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 utility. 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  $\ell$  ( $\neq i$ ) 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  $(\Theta_i)^J$ . Agent  $i$ 's worst cPBE utility is

$$\underline{w}_i^C := \min_{\pi \in \Pi^C(i)} \mathbb{E}_\mu [u_i(\pi(\theta), \theta)]. \quad (11)$$

A (stage) allocation or social choice function (SCF) is a mapping  $f : \Theta \rightarrow \Delta(\mathcal{A}_1 \times \cdots \times \mathcal{A}_J)$  from type profiles to probability distributions over actions. The set of SCFs is denoted by  $\mathcal{F}$ . The set of deterministic SCFs is denoted by

$$\mathcal{F}_0 := \{f \in \mathcal{F} | f(\Theta) \subset \mathcal{A}\} \subset \mathcal{F}. \quad (12)$$



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

$$\mathbb{E}_\mu[u_\ell(f(\theta), \theta)] > \underline{w}_\ell^C \text{ for all } \ell \in \mathcal{I} \cup \mathcal{J}. \quad (13)$$

$\mathcal{F}_0$  is the set of deterministic SCFs. 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:

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

The theorem below shows that any SCF  $f^* \in \mathcal{F}^C(\mu)$  is supportable in a cPBE of  $(G, \Gamma)^\infty(\delta)$ , provided players are sufficiently patient. Note that any  $f^* \in \mathcal{F}^C(\mu)$  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  $f^* \in \mathcal{F}^C(\mu)$  does not need to be SIR. As in any repeated game, what matters is that the expected utility is above the threshold even if the current period's utility (after observing the PCD) isn't.

## 5.1 Three or more agents

Suppose that principal  $j$ 's action is fixed at  $\alpha_j$ . A fixed action  $\alpha_j$  can be thought of as a constant mechanism that always assigns  $\alpha_j$  regardless of agents' messages. Given  $\alpha_j \in \mathcal{A}_j$ , let  $\Pi_{-j}^C(\alpha_j)$  be the set of all profiles of DMs for principals except  $j$  that are CIC conditional on  $\alpha_j$ :

$$\Pi_{-j}^C(\alpha_j) := \{\pi_{-j} \in \Pi_{-j} \mid (\pi_{-j}, \alpha_j) \in \Pi^C\}.$$

Lemma 1 below shows that  $\underline{w}_j^C$  for all  $j \in \mathcal{J}$  cannot be lower than

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

which is principal  $j$ 's maximin utility when  $j$  strategically chooses a fixed action given that the other principals will respond with CIC direct mechanisms conditional on  $j$ 's action to punish  $j$  most severely.

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

$$\max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} \mathbb{E}_\mu [u_j(\pi_{-j}(\theta), \alpha_j, \theta)] \leq \underline{w}_j^C \quad (14)$$

**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

$$\min_{\gamma_{-j} \in \Gamma_{-j}} \max_{\alpha_j \in \mathcal{A}_j} u_j(\gamma_{-j}, \alpha_j, \theta) \leq \underline{w}_j^C. \quad (15)$$

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

$$\min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} \mathbb{E}_\mu [u_j(\pi_{-j}(\theta), \alpha_j, \theta)].$$

It implies that

$$\max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} \mathbb{E}_\mu [u_j(\pi_{-j}(\theta), \alpha_j, \theta)] \leq \max_{\alpha_j \in \mathcal{A}_j} \min_{\gamma_{-j} \in \Gamma_{-j}} u_j(\gamma_{-j}, \alpha_j, \theta). \quad (16)$$

In addition, it is clear that

$$\max_{\alpha_j \in \mathcal{A}_j} \min_{\gamma_{-j} \in \Gamma_{-j}} \underline{u}_j(\gamma_{-j}, \alpha_j, \theta) \leq \min_{\gamma_{-j} \in \Gamma_{-j}} \max_{\alpha_j \in \mathcal{A}_j} \underline{u}_j(\gamma_{-j}, \alpha_j, \theta) \quad (17)$$

(15), (16), and (17) lead to (14). ■

The value  $\underline{w}_j^C$  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. The proof in Appendix A shows how to construct a profile of equilibrium communication strategies that makes principal  $j$  no better off with any mechanism in  $\Gamma_j$  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. As established in Theorem 1, this makes  $\underline{w}_j^C$  equal to the left hand side of (14). To show this, we utilize the class of extended direct mechanisms that is only slightly more general than DMs.

**Definition 3 (EDM)** *A mechanism  $\zeta_\ell^j : T_\ell \rightarrow \mathcal{A}_\ell$  offered by principal  $\ell$  is said to be an 'extended direct mechanism' (EDM) if*

1. *for some  $j \neq \ell$ , we have  $T_{i\ell} = \mathcal{A}_j \times \Theta_i$  and  $T_\ell := \times_{i \in \mathcal{I}} T_{i\ell}$  and*
2. *for any  $(\tilde{\alpha}_{i\ell}^j, \tilde{\theta}_{i\ell}) \in T_{i\ell}$  for all  $i \in \mathcal{I}$ , let  $\tilde{\alpha}_\ell^j := (\tilde{\alpha}_{i\ell}^j)_{i \in \mathcal{I}}$  be a profile of all agents' messages on principal  $j$ 's action and  $\tilde{\theta}_\ell := (\tilde{\theta}_{i\ell})_{i \in \mathcal{I}}$  a profile of all agents' messages on their types. Then,*

$$\zeta_\ell^j(\tilde{\alpha}_\ell^j, \tilde{\theta}_\ell) := \begin{cases} \varphi_\ell^j(\alpha_j)(\tilde{\theta}_\ell) & \text{if } \exists \alpha_j \text{ s.t. } \#\{i : \tilde{\alpha}_{i\ell}^j = \alpha_j\} > I/2, \\ \bar{\alpha}_\ell & \text{otherwise} \end{cases}, \quad (18)$$

where  $\varphi_\ell^j(\alpha_j) \in \Pi_j$  and  $\bar{\alpha}_\ell$  is some arbitrary action in  $\mathcal{A}_\ell$

In principal  $\ell$ 's EDM  $\zeta_\ell^j$  offered off the path following principal  $j$ 's deviation, an agent is asked to report (i) principal  $j$ 's action that she thinks he would take and (ii)

her type. If a strict majority of agents report  $\alpha_j$ ,  $\zeta_\ell^j$  assigns a direct mechanism  $\varphi_\ell^j(\alpha_j)$ , which then assigns principal  $\ell$ 's action  $\varphi_\ell^j(\alpha_j) \left( \tilde{\theta}_\ell \right)$  when the profile of type messages is  $\tilde{\theta}_\ell = (\tilde{\theta}_{i\ell})_{i \in \mathcal{I}}$ . When there is no action reported by a strict majority of agents,  $\zeta_\ell^j$  assigns a fixed action regardless of the type messages reported by agents.

**Definition 4** A profile of EDMs  $\zeta_{-j}^j = (\zeta_\ell^j)_{\ell \neq j}$  is CIC if  $\varphi_{-j}^j(\alpha_j) = (\varphi_\ell^j(\alpha_j))_{\ell \neq j} \in \Pi_{-j}^C(\alpha_j)$  for all  $\alpha_j \in \mathcal{A}_j$ .

We now present our main theorem, Theorem 1. The standard full dimensionality assumption (FD) that the set of expected utilities has the same dimension as the number of players, i.e.  $\dim[u(\mathcal{F}^C(\mu))] = J + I$ , allows us to design player-specific punishments (See Fudenberg and Maskin, 1986). It packs three results in one. First, there is the characterization of the CIC minmax value of a principal as the maxmin of a restricted space. Second, there is the equilibrium characterization. Third, we offer an extended revelation principle, which speaks to the kind of mechanisms that need to be used; it is worth noting that very simple mechanisms are needed, ones that are closer in dimension to the types spaces than to arbitrary mechanisms.

**Theorem 1** Consider i.i.d. types with distribution  $\mu \in \Delta\Theta$ . Under the standard full dimensionality assumption on the expected utilities, we have that

$$\underline{w}_j^C = \max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} \mathbb{E}_\mu[u_j(\pi_{-j}(\theta), \alpha_j, \theta)] \quad \forall j \in \mathcal{J}. \quad (19)$$

Furthermore,

1. **(Equilibrium Characterization)** Any SCF  $f^* \in \mathcal{F}^C(\mu)$  is the outcome of a cPBE of  $(G, \Gamma)^\infty(\delta)$  for high  $\delta$ .
2. **(Extended Revelation Principle)** To punish principal  $j$  all other principals employ EDMs satisfying CIC, whereas  $j$  employs a fixed action, i.e. a constant

*mechanism; to punish agent  $i$  all principals offer DM satisfying UIC for agent  $i$  and CIC for the other agents; in other phases, there is no loss of generality if principals offer DM satisfying CIC.*

**Proof.** See Appendix A. ■

The two intermediate results are combined to lead up to Theorem 1, showing that  $\underline{w}_j^C$  and  $\max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} E_\mu [u_j(\pi_{-j}(\theta), \alpha_j, \theta)]$  are equal. The first intermediate result is to show how and what type of information principals can extract from agents. They can freely extract any common information such as (i) what contract principal  $j$  who is being punished has offered, (ii) what action agents are supposed to induce from  $j$ 's contract, etc. This is because there are three or more agents. If all agents report  $\alpha_j$ , any single agent's deviation from  $\alpha_j$  cannot change principal  $\ell$ 's DM away from  $\varphi_\ell^j(\alpha_j)$ . Therefore, truthful action reporting can be sustained. On the other hand, they use CIC to extract agents' private information. The second result is to show that agents can be incentivized to not give any information to a principal who is being punished.

To see how this is done, note that in the phase where principal  $j$  is punished, each principal  $\ell \neq j$  offers an EDM  $\zeta_\ell^j$  that implements  $\varphi_\ell^j(\alpha_j)$  when a majority of agents report  $\alpha_j$  as principal  $j$ 's expected action such that  $\varphi_{-j}^j(\alpha_j) = (\varphi_\ell^j(\alpha_j))_{\ell \neq j}$  satisfies

$$\varphi_{-j}^j(\alpha_j) \in \arg \min_{\pi_{-j} \in \Pi_{-j}^C(\alpha_j)} \mathbb{E}_\mu [u_j(\pi_{-j}(\theta), \alpha_j, \theta)]. \quad (20)$$

If principal  $j$  offers any complex mechanism  $\gamma_j$  in the phase where he is punished, let agents send messages to principal  $j$  that induce  $g_j(\gamma_j)$  such that

$$g_j(\gamma_j) \in \min_{\alpha_j \in \gamma_j(M_j)} \mathbb{E}_\mu [u_j(\varphi_{-j}^j(\alpha_j)(\theta), \alpha_j, \theta)]. \quad (21)$$

Agents report  $\alpha_j = g_j(\gamma_j)$ , together with their types, to non-deviating principals who offer the EDMs. The crux in the proof of (19) 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  $\gamma_j$ , principals know that at least one agent deviated. They may not know the deviating agent's identity. Nonetheless, 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.<sup>11</sup>

If agent  $i$  conforms to induce  $g_j(\gamma_j)$ , 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.

(19) in 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  $j$  is punished,  $\underline{w}_j^C$  is reached without loss of generality when all principals except for  $j$  offers the EDMs  $\zeta_{-j}^j = (\zeta_\ell^j)_{\ell \neq j}$  satisfying (20) 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.

1 and 2 in Theorem 1 is established as follows. Suppose that  $f = (f_1, \dots, f_J)$  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

---

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

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.

A natural question arises as to whether we can further relax CIC. For example, it is tempting to remove IC over profiles 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.

### 5.1.1 Example

There are two principals and three agents -  $\mathcal{A}_1 = \{\alpha_1, \alpha'_1\}$ ,  $\mathcal{A}_2 = \{\alpha_2, \alpha'_2\}$ ,  $\Theta_3 = \{\theta_3, \theta'_3\}$ ,  $\Theta_4 = \{\theta_4\}$ ,  $\Theta_5 = \{\theta_5\}$ . Players 1 and 2 are principals and players 3, 4, and 5 are agents. The first agent's type is either  $\theta_3$  or  $\theta'_3$  with equal probability, whereas the second and third agents have no private information about their types because  $\Theta_4$  and  $\Theta_5$  are singletons. Players' payoffs are given by the following tables, one for each possible type of the first agent. The numbers in each cell represents players' payoffs in the order of players 1, 2, 3, 4, and 5.

$\theta = (\theta_3, \theta_4, \theta_5)$		
	$\alpha_2$	$\alpha'_2$
$\alpha_1$	4,2,2,1,1	3,5,3,1,1
$\alpha'_1$	6,8,4,1,1	9,9,2,1,1

$\theta' = (\theta'_3, \theta_4, \theta_5)$		
	$\alpha_2$	$\alpha'_2$
$\alpha_1$	8,6,4,1,1	7,9,2,1,1
$\alpha'_1$	2,3,1,1,1	5,4,3,1,1

For ease of exposition, we consider only deterministic SCFs. For each principal  $j$ , there are four possible DMs:  $\Pi_j = \{\bar{\pi}_j, \bar{\pi}'_j, \pi_j, \pi'_j\}$ , where the four DMs are defined as follows.

$$\begin{aligned}
\bar{\pi}_j(\theta_3, \theta_4, \theta_5) &= \alpha_j, & \bar{\pi}_j(\theta'_3, \theta_4, \theta_5) &= \alpha_j, \\
\bar{\pi}'_j(\theta_3, \theta_4, \theta_5) &= \alpha'_j, & \bar{\pi}'_j(\theta'_3, \theta_4, \theta_5) &= \alpha'_j, \\
\pi_j(\theta_3, \theta_4, \theta_5) &= \alpha_j, & \pi_j(\theta'_3, \theta_4, \theta_5) &= \alpha'_j, \\
\pi'_j(\theta_3, \theta_4, \theta_5) &= \alpha'_j, & \pi'_j(\theta'_3, \theta_4, \theta_5) &= \alpha_j.
\end{aligned}$$

Because each principal has four DMs, there are sixteen profiles of DMs that principals can offer. Given any profile of mechanisms, the first agent can report one of four different type profiles in  $\Theta_3 \times \Theta_3 = \{(\theta_3, \theta_3), (\theta_3, \theta'_3), (\theta'_3, \theta_3), (\theta'_3, \theta'_3)\}$ . The notion



of UIC imposes incentive compatibility over all possible type reports in  $\Theta_3 \times \Theta_3$  and hence the set of profiles of UIC DMs is

$$\Pi^U = \{(\bar{\pi}_1, \bar{\pi}_2), (\bar{\pi}_1, \bar{\pi}'_2), (\bar{\pi}_1, \pi'_2), (\bar{\pi}'_1, \bar{\pi}_2), (\bar{\pi}'_1, \bar{\pi}'_2), (\bar{\pi}'_1, \pi_2), (\pi_1, \bar{\pi}'_2), (\pi'_1, \bar{\pi}_2)\}$$

Although  $(\pi'_1, \pi_2)$  and  $(\pi'_1, \pi'_2)$  are not UIC, they are CIC, and the set of profiles of CIC DMs is

$$\Pi^C = \{(\pi'_1, \pi_2), (\pi'_1, \pi'_2)\} \cup \Pi^U$$

For example, the table below show the the action profile induced by the first agent's type reports given  $(\pi'_1, \pi_2)$ , where rows and columns correspond to the messages sent to principals 1 and 2 respectively.

	$\theta_3$	$\theta'_3$
$\theta_3$	$\alpha'_1, \alpha_2$	$\alpha'_1, \alpha'_2$
$\theta'_3$	$\alpha_1, \alpha_2$	$\alpha_1, \alpha'_2$

$(\pi'_1, \pi_2)$  is not UIC but CIC. To see why it is CIC, note that the set of action profiles induced by consistent type reports is  $\mathcal{A}((\pi'_1, \pi_2), \emptyset) = \{(\alpha'_1, \alpha_2), (\alpha_1, \alpha'_2)\}$ . Even though IC is satisfied over consistent type reporting,  $(\pi'_1, \pi_2)$  is not UIC: If the first agent of type  $\theta'_3$  reports  $\theta'_{33}$  to principal 1 but  $\theta_3$  to principal 2 so that  $(\alpha_1, \alpha_2)$  is induced, her payoff is 4 whereas her payoff is 2 when  $(\alpha_1, \alpha'_2)$  is assigned by truthful type reporting to both principals. However, if we adopt the notion of CIC, we do not worry about such an inconsistent type reports because  $(\alpha_1, \alpha_2) \notin \mathcal{A}((\pi'_1, \pi_2), \emptyset)$ . In fact, any inconsistent type report results in an action not in  $\mathcal{A}((\pi'_1, \pi_2), \emptyset)$ . We can also show that  $(\pi'_1, \pi'_2)$  is CIC but not UIC.

The table below shows the expected payoffs for players 1, 2, and 3 (principal 1, principal 2, and the first agent) given each profile of CIC DMs. The first eight profiles

of DMs are UIC (and therefore CIC) and the last two profiles are only CIC but not UIC.

	$\bar{\pi}_1, \bar{\pi}_2$	$\bar{\pi}_1, \bar{\pi}'_2$	$\bar{\pi}_1, \pi'_2$	$\bar{\pi}'_1, \bar{\pi}_2$	$\bar{\pi}'_1, \bar{\pi}'_2$	$\bar{\pi}'_1, \pi_2$	$\pi_1, \bar{\pi}'_2$	$\pi'_1, \bar{\pi}_2$	$\pi'_1, \pi_2$	$\pi'_1, \pi'_2$
P1	6	5	<b>5.5</b>	4	7	<b>7.5</b>	6	<b>7</b>	<b>6.5</b>	8.5
P2	3	7	<b>5.5</b>	5.5	6.5	<b>6</b>	4.5	<b>7</b>	<b>8.5</b>	7.5
P3	3	2.5	<b>3.5</b>	2.5	2.5	<b>3.5</b>	3	<b>4</b>	<b>3</b>	2.5

Since the second and third agents have a single type message and their payoffs are always one, there is no need to consider their incentive compatibility. Which profiles of CIC in  $\Pi^C$  can be supported as an equilibrium allocation? To identify, we need to compute the minmax value for principals 1 and 2, and the first agent.

Let us first compute principal 1's pure minmax value with respect to complex mechanism via his action-DM pure minmax value. Fixing principal 1's action at  $\alpha_1$ , the set of principal 2's DMs that are CIC conditional on  $\alpha_1$  is  $\Pi_{-1}^C(\alpha_1) = \{\bar{\pi}_2, \bar{\pi}'_2, \pi'_2\}$ .<sup>12</sup> Given  $\alpha_1$ , DMs in  $\Pi_{-1}^C(\alpha_1)$  generate the expected payoffs for principal 1 as follows:  $\mathbb{E}_\mu[u_1(\bar{\pi}_2(\theta), \alpha_1, \theta)] = 6$ ,  $\mathbb{E}_\mu[u_1(\bar{\pi}'_2(\theta), \alpha_1, \theta)] = 5$ ,  $\mathbb{E}_\mu[u_1(\pi'_2(\theta), \alpha_1, \theta)] = 5.5$ . Therefore, if principal 1 plays  $\alpha_1$ ,  $\bar{\pi}'_2$  minimizes principal 1's expected payoff among all DMs in  $\Pi_{-1}^C(\alpha_1)$ . Therefore, in principal 2's EDM,  $\varphi_2^1(\alpha_1) = \bar{\pi}'_2$ .

Similarly,  $\Pi_{-1}^C(\alpha'_1) = \{\bar{\pi}_2, \bar{\pi}'_2, \pi'_2\}$ .  $\bar{\pi}_2$  minimizes principal 1's expected payoff among all DMs in  $\Pi_{-1}^C(\alpha'_1)$  and it is  $\mathbb{E}_\mu[u_1(\bar{\pi}_2(\theta), \alpha'_1, \theta)] = 4$ . Therefore, in principal 2's EDM,  $\varphi_2^1(\alpha'_1) = \bar{\pi}_2$ . Because  $\mathbb{E}_\mu[u_1(\bar{\pi}_2(\theta), \alpha'_1, \theta)] = 4 < \mathbb{E}_\mu[u_1(\bar{\pi}'_2(\theta), \alpha_1, \theta)] = 5$ , it follows that

$$\underline{w}_1^C = \mathbb{E}_\mu[u_1(\bar{\pi}'_2(\theta), \alpha_1, \theta)] = 5.$$

---

<sup>12</sup>With two principals in this example, principal 2 is only one non-deviating principal when principal 1 has deviated. In this case, the set of principal 2's CIC DMs conditional on principal 1's action  $\alpha_1$  is the same as the set of principal 2's UIC DMs conditional on principal 1's action  $\alpha_1$ .

No such algorithm exists in the one-shot model. Similarly, for principal 2,

$$\underline{w}_2^C = 4.5.$$

When the first agent 1 is punished, we need to consider UIC for her. Since the second and third agents have a single message, this implies that the minmax value for agent 1 is

$$\underline{w}_3^C := \min_{\pi \in \Pi^C(3)} \mathbb{E}_\mu [u_3(\pi(\theta), \theta)] = \min_{\pi \in \Pi^U} \mathbb{E}_\mu [u_3(\pi(\theta), \theta)] = 2.5$$

With a slight abuse of notation, the set of strictly individually rational (deterministic) SCFs is

$$\{(\bar{\pi}_1, \pi'_2), (\bar{\pi}'_1, \pi_2), (\pi'_1, \bar{\pi}_2), (\pi'_1, \pi_2)\}$$

Each of them can be supported as an equilibrium allocation since players 1, 2, and 3 receive expected payoffs higher than their minmax value,  $\underline{w}_1^C$ ,  $\underline{w}_2^C$ , and  $\underline{w}_3^C$  respectively.

## 5.2 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  $\ell$  ( $\neq j$ ) offers the EDM  $\zeta_\ell^j$  defined in (18) to punish principal  $j$ . If two agents' reports on  $j$ 's action to principal  $\ell$  are not consistent, principal  $\ell$  knows that at least one agent has deviated. However, the other principals do not know that because agents' action

reports to principal  $\ell$  are not observable by them. In order to punish an agent together with the other principals, principal  $\ell$  needs to let them know that at least one agent has sent to him a false message. The following corollary shows that Theorem 1 holds with  $I = 2$  by identifying a sufficient condition for that.

**Corollary 2** *If*

$$\mathcal{A}_\ell \setminus \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\} \neq \emptyset. \quad (22)$$

*is satisfied for every principal  $\ell \in \mathcal{J}$  and every  $j \in \mathcal{J} \setminus \{\ell\}$  then Theorem 1 holds with  $I = 2$ .*

**Proof.** See Appendix B. ■

We believe that the sufficient condition is very weak because it is satisfied as long as each principal's action set is large so that principal  $\ell$  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  $\ell$  charges the price of 0.21321 to every retailer in a phase where manufacturer  $\ell$  is punished. Therefore, (22) is satisfied. Manufacturer  $\ell$  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.

### 5.3 Observability

We have assume that, at the end of each period, mechanisms and actions are observable to both principals and agents. However, it is not necessary.

First, consider the case with three or more agents. For characterization and implementation of equilibrium allocations without imposing the observability of mechanisms

and actions for principals, we modify the model so that agents are allowed to send *privately observed cheap talk messages* from  $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$  (the power set of  $\mathcal{I} \cup \mathcal{J}$ ) to principals at the end of every period.

**Corollary 3** *Suppose that principals do not observe the other principals' mechanisms and their actions, whereas agents observe mechanisms with probability one but actions with positive probability  $\lambda > 0$ . Given the availability of privately observed cheap talk messages from  $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$  at the end of each period, Theorem 1 holds with  $I \geq 3$ .*

**Proof.** See Appendix C. ■

We can also extend Theorem 1 to the case with two agent without imposing the observability of mechanisms and actions for principals.

**Corollary 4** *Suppose that principals do not observe the other principals' mechanisms and their actions, whereas agents observe mechanisms with probability one but actions with positive probability  $\lambda > 0$ . If (22) is satisfied for every principal  $\ell \in \mathcal{J}$  and every  $j \in \mathcal{J} \setminus \{\ell\}$  and each agent can announce a publicly observable cheap talk message from  $\mathcal{P}(\mathcal{I} \cup \mathcal{J})$  at the end of each period, Theorem 1 holds with  $I = 2$*

**Proof.** See Appendix D. ■

## 5.4 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 path, we cannot dispense with incentive compatibility off the equilibrium path over message profiles that must be imposed over message 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.

The reason is that we need to deter an agent to further deviate from her 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 where no one has previously deviated, 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. Suppose that given the realization of the PCD, principals offer a profile of DMs  $\pi$  that impose incentive compatibility only over  $B_i(\pi)$  for all  $i$  on the equilibrium path. Recall that  $\mathcal{A}(\pi, \emptyset)$  is the set of actions that can be induced from  $\pi$  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  $\mathcal{A}(\pi, \emptyset)$  irrespective of the types of the other agents as long as the others report truthfully:

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

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  $(1 - \delta) \bar{u} + \delta \left[ (1 - p^*) v_i + \frac{p^*}{I} \sum_{k=J+1}^{J+I} L_i^k \right]$ , where  $p^*$  is the maximum probability that an agent's lie is detected and  $\bar{u}$  is the maximum utility. As  $\delta \rightarrow 1$ , 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.

It worth mentioning that Jackson and Sonnenschein (2007) have shown how a single mechanism designer can 'link' several decisions together by giving each agent a 'budget' of messages that encompasses all decisions, and thereby impose some 'statistical checks' on messages, instead of directly providing incentives through payoffs of each individual period. Similar intuition is found in the dynamic cheap-talk game between a single sender and single receiver (Renault, Solan, and Vielle (2013)).

The key idea is that in a fully dynamic setting, we divide equilibrium play into several 'blocks' - within each block the mechanism designer or principal follows the sender's suggestion until the sender exhausts her budget; thereafter the message is replaced by a fictitious message in a deterministic fashion so that the budget holds exactly at the end of the block. This logic does not directly carry over to our model: the crux is that messages are private between an agent and a principal, and neither the other agents nor the other principals know whether agent  $i$  has exhausted her quota of messages to principal  $j$ .

## 6 Conclusion

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 allo-

cations, 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 phenomenon underlies the extended revelation principle and facilitates the characterization and implementation of equilibrium allocations using straightforward contracts, effectively circumventing the infinite regress problem.



## 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] Baker, G., Gibbons R. and K. J. Murphy (1994), “Subjective Performance Measures in Optimal Incentive Contracts,” *Quarterly Journal of Economics*, 109(4), 1125-1156.
- [4] Battaglini, M. and R. Lamba (2019), “Optimal Dynamic Contracting: The First-Order Approach and Beyond,” *Theoretical Economics*, 14, 1435-1482.
- [5] Bergemann, D. and J. Välimäki (2010), “The Dynamic Pivot Mechanism,” *Econometrica*, 78(2), 771-789.
- [6] Dasgupta A. and S. Ghosh (2016), “Repeated Games with Public Randomisation: A Self-accessible approach,” Boston University working paper.
- [7] Fudenberg, D. and E. Maskin (1986), “The Folk Theorem in Repeated Games with Discounting Or with Incomplete Information,” *Econometrica*, 54(3), 533-554.
- [8] ————— (1991), “On the Dispensability of Public Randomization in Discounted Repeated Games,” *Journal of Economic Theory* 53(2), 428-438.
- [9] Guo, Y. and J. Hörner (2018), “Dynamic Allocation without Money,” Working paper, Northwestern University.
- [10] Hörner, J., Takahashi, S., and N. Vielle (2015), “Truthful Equilibria in Dynamic Bayesian Games,” *Econometrica*, 83, 1795-1848.

- [11] Jackson, M. O. and H. F. Sonnenschein (2007), “Overcoming Incentive Constraints by Linking Decisions,” *Econometrica*, 75, 241-257
- [12] Levin, J. (2003), “Relational Incentive Contracts,” *American Economic Review*, 93, 835-857.
- [13] 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](https://cpb-us-w2.wpmucdn.com/web.sas.upenn.edu/dist/1/313/files/2017/04/Modeling_strategic_behavior_9_16-2020.pdf).
- [14] McAfee, R. P. (1993), “Mechanism Design by Competing Sellers,” *Econometrica*, 61(6), 1281-1312.
- [15] Pavan, A., I. Segal, and J. Toikka (2014), “Dynamic Mechanism Design: A Myersonian Approach,” *Econometrica*, 82(2), 601-653.
- [16] Pearce, D. G. and E. Stacchetti (1998), “The Interaction of Implicit and Explicit Contracts in Repeated Agency,” *Games and Economic Behavior*, 23(1), 75-96.
- [17] Renault, J., Solan, E. and N. Vieille (2013), “Dynamic Sender-Receiver Game,” *Journal of Economic Theory*, 148, 502-534
- [18] Sugaya, T. and A. Wolitzky (2021), “The Revelation Principle in Multistage Games,” *Review of Economic Studies*, 88, 1503-1540.
- [19] Watson, J. (2017), “A General, Practicable Definition of Perfect Bayesian Equilibrium,” University of California San Diego, Working paper

# Appendix

## A Proof of Theorem 1

We first prove (19) in Theorem 1. 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  $\gamma_j(M_j)$  only conditional on  $\gamma_j$ , say  $g_j(\gamma_j) \in \gamma_j(M_j)$  irrespective of their own types. This means that offering  $\gamma_j$  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  $\alpha_j$  that principal  $j$  takes, for the other principals, punishing principal  $j$  is equivalent to choosing a profile of their CIC DMs conditional on  $\alpha_j$ . The reason is that agents' equilibrium communication with them, given the mechanisms that they choose, induces CIC DMs. Therefore, principal  $j$ 's lowest possible utility conditional on  $\alpha_j$  can be realized if other principals can implement  $\varphi_{-j}^j(\alpha_j) = (\varphi_\ell^j(\alpha_j))_{\ell \neq j} \in \Pi_{-j}^C(\alpha_j)$ , where  $\varphi_\ell^j(\alpha_j)$  is principal  $\ell$ 's DM and  $\varphi_{-j}^j(\alpha_j)$  is defined in (20). In the phase where principal  $j$  is punished, all principals offers the EDMs  $\zeta_{-j}^j = (\zeta_\ell^j)_{\ell \neq j}$  satisfying (20).

Now we choose the mapping  $g_j : \Gamma_j \rightarrow \mathcal{A}_j$  as follows. If principal  $j$  offers  $\gamma_j \in \Gamma_j$  in the phase where he is punished, let agents send messages to principal  $j$  that induce  $g_j(\gamma_j)$ , which is defined in (21). At the same time, agents send  $g_j(\gamma_j)$  and their types to non-deviating principals who offer the EDMs. The proofs of 1 and 2 in Theorem 1 below show 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  $\Gamma_j$  than he does by offering an action in  $\mathcal{A}_j$ . Therefore, even when principal  $j$  can choose any mechanism in  $\Gamma_j$ , the maximum utility principal  $j$  can

achieve is

$$\max_{\alpha_j \in \mathcal{A}_j} \mathbb{E}_\mu [u_j (\varphi_{-j}^j (\alpha_j) (\theta), \alpha_j, \theta)] = \max_{\alpha_j \in \mathcal{A}_j} \min_{\pi_{-j} \in \Pi_{-j}^C (\alpha_j)} \mathbb{E}_\mu [u_j (\pi_{-j} (\theta), \alpha_j, \theta)], \quad (\text{A1})$$

which is the left-hand side of (14). Together with Lemma 1, this implies that (19) is true for all  $j \in \mathcal{J}$ .

Now we prove 1 and 2 in Theorem 1.

**Notation:** In discussions related to our results, 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 utility for  $j \in \mathcal{I} \cup \mathcal{J}$ . This is the target equilibrium utility for player  $j$ . Following Fudenberg and Maskin (1986), we choose a vector of utilities  $(v'_1, \dots, v'_J)$ , 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 utilities

$$\beta^j := (\beta_i^j)_{i=1}^{i=I+J} = (v'_1 + \epsilon, \dots, v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, \dots, v'_{J+I} + \epsilon) \quad (\text{A2})$$

such that  $\epsilon$  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  $f = (f_1, \dots, f_J)$  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  $\theta_i^t$  to all principals at time  $t$ . If principal  $j$  deviates unilaterally (offers a mechanism other than  $\pi_j$ ), agents play a one-shot continuation equilibrium in the current period; play moves to phase II <sub>$j$</sub>  from the next period. If agent  $i$ 's deviation from truthful type reporting is detected, move to phase II <sub>$i$</sub>  from the next period.

2. Let us explain phase II. Phase  $\text{II}_j$  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  $g_j(\gamma_j)$  irrespective of their types. Each principal  $k \neq j$  offers the EDM  $\zeta_k^j$  that assigns the DM  $\varphi_k^j(\alpha_j)$  if a majority of agents report  $\alpha_j$ . Agents report the true types and  $g_j(\gamma_j)$  to each principal  $k \neq j$  at time  $t$ .

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

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

3. In phase  $\text{III}_i$ , for  $i \in \mathcal{I} \cup \mathcal{J}$ , pick a (correlated) SCF  $\tilde{f} \in \Delta\mathcal{F}_0^C$  based on the realization of the PCD, which yields expected utility vector  $\beta^i$ ; principal  $j$  offers the DM  $\pi_i = f_i$  when a SCF  $f = (f_1, \dots, f_J)$  in the support of  $\tilde{f}$  is supposed to be implemented give the realization of the PCD. Agents report their types truthfully. Remain forever in this phase, unless any player  $\ell$  deviates unilaterally and triggers phase  $\text{II}_\ell$ .

#### 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  $L_j^i$  denote player  $j$ 's expected utility from the beginning of phase  $\text{II}_i$  without deviation. First, with  $j = i$ , player  $i$ 's lifetime (discounted average) utility in phase  $\text{II}_i$

is defined recursively as  $L_i^i = (1 - \delta)\underline{w}_i^C + \delta(qL_i^i + (1 - q)\beta_i^i)$ , so that

$$L_i^i = \frac{(1 - \delta)\underline{w}_i^C + \delta(1 - q)\beta_i^i}{1 - \delta q}. \quad (\text{A3})$$

Note that  $L_i^i \rightarrow \beta_i^i = v_i'$  as  $\delta \rightarrow 1$ . When calculating  $L_j^i$  for  $j \neq i$ , note that it is bounded on both sides as follows:

$$(1 - \delta)(-\bar{u}) + \delta(qL_j^i + (1 - q)\beta_j^i) \leq L_j^i \leq (1 - \delta)\bar{u} + \delta(qL_j^i + (1 - q)\beta_j^i).$$

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

$$\bar{u}(1 - q) < \beta_j^j(1 + p - q) - p\underline{w}_j^C \text{ for all } j \in \mathcal{I} \cup \mathcal{J}. \quad (\text{A4})$$

Such a  $q$  exists because at  $q = 1$  this inequality becomes  $0 < p(\beta_j^j - \underline{w}_j^C)$ , which is satisfied because  $\beta_j^j = v_j' > \underline{w}_j^C$ .

As  $\delta \rightarrow 1$ , it is easy to check that  $L_j^i \rightarrow \beta_j^i = v_j' + \epsilon$ .

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

$$(1 - \delta)\bar{u} + \delta[(1 - p)\beta_i^i + pL_i^i] - \beta_i^i = (1 - \delta) \left[ \bar{u} - \frac{(1 + \delta p - \delta q)\beta_i^i - \delta p\underline{w}_i^C}{1 - \delta q} \right] \quad (\text{A5})$$

using (A3). An immediate implication of inequality (A4) defining  $q$  is that (A5) is strictly negative for all  $\delta$  close to 1, so that  $i$  cannot profitably deviate from Phase III<sub>*i*</sub>. Since  $\beta_j^i > \beta_j^j \forall j \neq i$ , it is immediate that, for such  $\delta$  satisfying  $(1 - \delta)\bar{u} + \delta[(1 - p)\beta_i^i + pL_i^i] - \beta_i^i < 0$ , players  $j \neq i$  do not have a profitable one-shot deviation either from phase III<sub>*i*</sub>.

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

$$(1 - \delta) \bar{u} + \delta \frac{1}{I} \sum_{k=J+1}^{J+I} L_i^k \leq L_i^j.$$

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

$$(1 - \delta) \bar{u} + \delta[(1 - p)L_i^j + pL_i^i] \leq L_i^j. \quad (\text{A6})$$

This condition is satisfied with strict inequality as  $\delta \rightarrow 1$  because the left-hand and right hand-side approach  $v'_i + (1 - p)\epsilon$  and  $v'_i + \epsilon$  respectively as  $\delta \rightarrow 1$ . 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

$$(1 - \delta) \bar{u} + \delta L_i^i \leq L_i^j, \quad (\text{A7})$$

which is clearly satisfied as  $\delta \rightarrow 1$  because the left-hand and right hand-side approach  $v'_i$  and  $v'_i + \epsilon$  respectively as  $\delta \rightarrow 1$ . Principal  $j$  will choose his own mechanism that best respond to others' EDMs and agents' communication protocol.

3. Phase  $\Pi_j$  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  $(1 - \delta) \bar{u} + \delta[(1 - p)L_i^j + pL_i^i] \leq L_i^j$ . Similar to (A6), this is satisfied as  $\delta \rightarrow 1$ . Principal  $i \neq j$  will not deviate if  $(1 - \delta) \bar{u} + \delta L_i^i \leq L_i^j$ . Similar to (A7), this is also satisfied as  $\delta \rightarrow 1$ .

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  $(1 - \delta) \bar{u} + \delta L_j^j \leq v_j$ , which is satisfied because, as  $\delta \rightarrow 1$ , the left-hand approaches  $v_j'$  that is less than  $v_j$ . Agent  $j$  will not deviate if  $(1 - \delta) \bar{u} + \delta[(1 - p)v_j + pL_j^j] \leq v_j$ , which is also satisfied because, as  $\delta \rightarrow 1$ , the left-hand approaches  $(1 - p)v_j + pv_j'$  that is less than  $v_j$ .

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

## B Proof of Corollary 2

Note that given the two agents, principal  $\ell$ 's EDM  $\zeta_\ell^j$  assigns  $\bar{\alpha}_\ell$  when the two agents send inconsistent messages on  $j$ 's action. The crux of the proof is how to pick  $\bar{\alpha}_\ell$  that is a perfect signal to the other principals on agents' inconsistent messages on  $j$ 's action to principal  $\ell$ .

Principal  $\ell$ 's EDM assigns an action  $\varphi_\ell^j(\alpha_j)(\theta)$  when  $\theta$  is a profile of type messages that agents send to principal  $\ell$  and more than a half of agents (both agents in the model with two agents) send  $\alpha_j$  as  $j$ 's action to principal  $\ell$ . Recall that  $\varphi_{-j}^j(\alpha_j)$  is the profile of CIC DMs for the principals except for  $j$  that minimizes principal  $j$ 's utility conditional on  $\alpha_j$ . 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  $\ell$ '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  $\mathcal{A}_\ell \setminus \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\}$  is non-empty. For  $\bar{\alpha}_\ell$ , principal  $\ell$  can choose an arbitrary action in  $\mathcal{A}_\ell \setminus \{\varphi_\ell^j(\alpha_j)(\theta) \in \mathcal{A}_\ell : \alpha_j \in \mathcal{A}_j, \theta \in \Theta\}$ . Upon observing  $\bar{\alpha}_\ell$  at the end of the period, the other principals know that two agents sent inconsistent messages on principal  $j$ 's action to principal  $\ell$ . After observing  $\bar{\alpha}_\ell$ , 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, (22). The remaining proof to establish Theorem 1 with  $I = 2$  is omitted since it is analogous to the proof of Theorem 1 with  $I \geq 3$ .

## C Proof of Corollary 3

The crucial part of the proof is how agents monitor themselves and let principals know a player's deviation when they detect it. The other parts of the proof can be done analogous to the proof of Theorem 1.

In any period, 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  $\{j\} \subset \mathcal{J}$  to each principal  $\ell$  then those principals believe that  $j$  has deviated and it is what principal  $j$  expects when he deviates. If more than a half of agents report  $\{i\} \subset \mathcal{I}$  to every principal, principals believe that  $i$  has deviated. If more than a half of agents report  $\mathcal{I}$  to each principal, principals believe that there is at least one agent has deviated.

Fix a period. If principal  $j$  deviates, each agent sends  $\{j\}$  as her cheap talk message to all principals. Note that a principal's deviation is detected by agents with probability one because mechanisms are observable.

If actions become observable for agents and the observed action profile uniquely identifies the deviating agent, say  $i$ , then each agent sends  $\{i\}$  as her cheap talk message to all principal.

If the current period is in the phase where principal  $j$  is punished actions, actions become observable for agents, and the observed action profile reveals that there is a deviating agent but not the identity of a deviating agent, then each agent sends  $\mathcal{I}$  to each principal. Principals then punish each agent with equal probability. In all other cases, agents remain silent.

Given principals' beliefs, the cheap talk message strategy described above is optimal for each agent when everyone else follows it. The only thing we need to be careful about is that the choice of  $q$ . Recall that, after player  $j \in \mathcal{I} \cup \mathcal{J}$  deviates, players stay in phase  $\text{II}_j$  where the deviator is punished with probability  $q$  and move to the phase where each player  $i (\neq j)$  is rewarded with  $\beta_i^j$  with probability  $1 - q$ . Analogous to (A4), we choose  $q \in (0, 1)$  such that

$$\bar{u}(1 - q) < \beta_j^j(1 + \lambda p - q) - p\underline{w}_j^C \text{ for all } j \in \mathcal{I} \cup \mathcal{J}. \quad (\text{A8})$$

Such a  $q$  exists because at  $q = 1$  this inequality becomes  $0 < \lambda p(\beta_j^j - \underline{w}_j^C)$ , which is satisfied because  $\beta_j^j = v_j' > \underline{w}_j^C$ .

## D Proof of Corollary 4

When both agents' cheap talk messages are the same, principals believe that it is the set of deviating players. If agents' cheap talk messages are inconsistent, principals believe that at least one agent's cheap talk message is not true. In this case, both agents are punished with equal probability.

An agent's cheap talk message is truthful except for the case where an agent, say  $i$ ,

deviates from her communication strategy for principals' mechanisms, actions become observable for agents and the observed action profile reveals her deviation. In this case, she announces  $i' \in \mathcal{I} \setminus \{i\}$  as her cheap talk. Since the other agent announces  $i$  in this case, each agent is punished with equal probability.

The remaining parts of the proof can be done analogous to the proof of Theorem 1 given  $q \in (0, 1)$  that satisfies (A8).