Sequential Common Agency: The Revelation Principle

Anastasia Kartasheva*

July 18, 2006

Abstract

The paper extends the Revelation Principle to sequential common agency games under asymmetric information. Each period a principal contracts with a common agent. An implemented allocation is observed by other principals. Depending on whether the message reported by the agent to a principal is observed by other principals, we distinguish between private and public communication. Under private communication, the Revelation Principle applies, but optimal contracts are stochastic. However, the dimension of the support of an equilibrium contract does not exceed the number of types that achieve this stage with a positive probability. Under public communication, the reporting strategy of agent is stochastic, but the true type is reported with a positive probability. We demonstrate that the two regimes are equivalent in that they result in the same distribution of allocations. The results hold when the agent's type is not persistent, or the outcome of the contract is observed with noise.

Keywords: common agency, sequential mechanisms, dymanic contracts, adverse selection. Revelation Principle.

JEL Codes: C73, D82.

1 Introduction

In this paper we derive the Revelation Principle for sequential common agency games. We analyze a situation in which a number of principals contract sequentially with an agent

*This paper is adapted from Chapter 1 of my doctoral dissertation at the University of Toulouse. I am indebted to Jean-Jacques Laffont and Patrick Rey for their encouragement and supervision of my work. I am grateful to Bruno Jullien, David Martimort, Ilya Segal and Jean Tirole, and seminar participants in Toulouse, ECARES, Brussels, EEA Meeting 2005 for helpful comments. *Correspondence address:* Department of Risk Management and Insurance, J.Mack Robinson College of Business, Georgia State University, P.O. Box 4036, Atlanta, GA 30302 - 4036, USA. *Email:* kartasheva@gsu.edu. Home page: http://www.rmi.gsu.edu/rmi/faculty/kartasheva.htm

under asymmetric information about agent's preferences. At each stage one principal offers a contract to the agent. The implementation of the contract results in allocation which is observed by the other principals. An allocation is payoff relevant for all the principals and the agent, and may affect the feasible sets of allocations of subsequent principals.

The framework we study applies to many economic environments. Examples include: (i) non-exclusive credit, where information sharing between creditors affects the amount of credit and the probability of debt repayment (Padilla and Pagano (1997), Pagano and Jappelli (1993), Sharpe (1990)); (ii) interaction of public and private health insurance programs, which affects the choice of the insured and the terms of contract of the private insurer (Culter and Gruber (1996)); (iii) retail industries such as supermarkets, airlines, credit cards, where the information about the purchase history of customers allows sellers to offer personalized deals (Acquisti and Varian (2002), Chen and Zhang (2001), Taylor (2002), Villas-Boas (1999)); and (iv) certification intermediaries, where information disclosure by an intermediary affects the size and the distribution of the surplus between the buyer and the seller (Lizzeri (1999), Peyrache and Quesada (2004)), (v) interaction between firm's financing and production decisions, where the choice of financial structure affects firm's position vis-a-vis its competitors (Gertner, Gibbons and Scharfstein (1988), Bhattacharya and Ritter (1983)).

Although the above applications provide useful insights, in most of them either there is no strategic role for principals, due to competition assumption, for example, or the analysis is restricted to very particular institutional arrangements such as linear contracts. As a result, the predictions are very sensitive to assumptions of a particular model.

One possible reason for the lack of a unified, general approach is that the Revelation Principle[1] widely used to study contractual relationships under asymmetric information, may not be valid in the environments with more than one principal. The standard Revelation Principle states that when a principal contracts with agents under asymmetric information about agents' preferences, any contract can be described by a direct incentive compatible mechanism in which the terms of the contracts are based on the agent's report on its private information, and the agent has incentives to report the information thruth-

[1] The revelation principle for the environment in which a principal contracts with agents who have private information has been developed by Gibbard (1973), Green and Laffont (1977), Dasgupta, Hammond and Maskin (1979), and Myerson (1979). Myerson (1982) extends the revelation principle to the situations when the principal also faces moral hazard problems, in addition to asymmetries of information.

fully. The practical application of this result is that the optimal contract can be found by the means of optimal programming subject to incentive compatibility constraints.

Extending the Revelation Principle to games with many principals can be problematic. The literature started by analyzing common agency games in which the principals offer contracts simultaneously to the agent(s). Epstein and Peters (1999) characterize the universal message space that can be used to characterize any indirect mechanism. However, this message space may not be practical. The reason is that in such situation each principal would like to make its contract dependent on the contracts offered by the other principals, leading to the problem of infinite regress. As a result the simplest message space needed to describe the mechanisms must be rather rich, and is hard to work with. When principals offer contracts in a sequential manner, the problem of infinite regress does not arise: Once a principal and an agent implement a contract, the principal cannot improve her payoff by requesting information about subsequent offers.

This paper is related to two other recent papers in the literature. When the outcome of contracting between a principal and an agent is not observed by the other principals, Pavan and Calzolari (2006) show that the equilibria can be described within the message space that includes agent's types and the allocations implemented with the preceding principals. On the contrary, we consider sequential common agency games with public contracts in which the contract and the outcome of the contract, that is, the allocation implemented at each stage, are observable by the other principals. This framework is a better description of economic environments mentioned above. Also it leads to a different type of externality that a contract between a principal and an agent exerts on the other principals: The implemented allocation itself becomes a signal about agent's private information. In this respect the paper is closely related to work on dynamic principal - agent relationships under imperfect commitment of Bester and Strausz (2001). We show that techniques developed in Bester and Strausz can be applied to study a dynamic contracting problem with many principals, or any combination of single principal - agent relationships under imperfect commitment and multiprincipal relationships. In this respect we provide a generalization of Bester and Strausz result.

We study general communication mechanisms in which, without loss of generality, a contract is composed of a message space and a decision rule. The agent sends a message from the specified message set and, based on the message, a principal commits to a (possibly stochastic) contract that implements a feasible allocation. All principals are free to choose the message spaces and decision rules, which may in particular be state dependant.

In general there are three elements that can signal agent's private information to subsequent principals: the mechanism, its outcome, and the message reported by the agent. Thus we distinguish between the cases of public and private communication. Under private communication, the message reported by the agent to one principal is not observed by the other principals. Under public communication, the principals also observe the information submitted by the agent to preceding principals.

The main result of the paper is that the set of equilibrium allocations of the sequential common agency game can be characterized within the type space. Naturally, the agent's incentives to report private information to a principal depend crucially on whether this report is observed by the other principals.

Under private communication, the agent is not concerned that his report to one principal may affect the contracting choices with the other principals. As a result, the standard version of the Revelation Principle applies at each stage game. At the same time, deterministic contracts are suboptimal. Indeed, assigning a distinct allocation to each type implies that the outcome of the contract is a perfect signal about agent's type to the other principals. By offering a lottery a principal can limit the information about the agent's type revealed by the outcome of the contract. Thus, an optimal contract of a principal is a menu of lotteries designed for each type of agent. However, this result does not imply any restrictions on the structure of the lottery. Next, we study the structure of an optimal lottery and show that the dimension of its support does not exceed the number of types that reach a given stage with a positive probability. Therefore, an optimal contract can be characterized as a menu of lotteries over a finite support. Consequently, an optimal contract can be found as a solution to an optimization problem.

Under public communication, the revelation of private information by the agent may be costly for both the principal and the agent. The reason is that this information can be used by other principals in the subsequent stages. A similar problem arises when a single principal contracts with an agent over a number of periods. As new information about the agent becomes available during the relationship, the principal and the agent may find it mutually beneficial to renegotiate the initial long term contract. However, anticipating the renegotiation of the initial contract, the agent may become less prompt to reveal its private information. This ultimately increases the cost for the principal of inducing a truthful report. As a result, the principal may prefer to decrease the informativeness of the agent's report about its private information. Bester and Strausz (2001) analyze this situation and establish that the equilibria of the game can be characterized using direct mechanisms in which it is an optimal strategy for the agent to report its type. In contrast with the standard Revelation Principal, however, the agent does not necessarily reveal its private information with probability one: It may be beneficial for both parties that the agent randomizes and misreports its information with some probability. In this paper we show that the technique developed by Bester and Strausz (2001) can be extended to sequential common agency games with public communication. This is because at each contracting stage, the principal, be it the same or a different one, with possibly a different objective at each state, is constrained to offer allocations that belong to the Perfect Bayesian equilibria of the continuation game. The main difference between our framework and that of imperfect commitment is that a principal does not internalize the impact of its contract on subsequent stages, and out-of-equilibrium messages may be needed to preserve the equilibrium outcome. However, we show that these messages can be preserved by the means of a direct mechanism.

We compare the set of equilibria under private and public communication. We show that the two sets are equivalent in a sense that for each type of agent they induce the same probability distribution of allocations. The main idea of this result is that a principal can generate the same belief about agent's type for subsequent principals either by designing a lottery on the set of the messages, as under public communication, or the set of implemented allocation, as under private communication.

The characterization results of the paper also apply to situations when the type of agent is not persistent over time, and when the messages or allocations are observed with some exogenous noise. In these case this information must be incorporated in definition of the Bayes rule, but the equilibria can still be studied within the type space.

Our result is weaker than the Revelation Principle of Epstein and Peters as it holds only for equilibrium mechanisms while the later paper constructs a language to characterize all the mechanisms of the game. However, we believe that it provides a useful tool to study many applications. A useful practical feature of our result is that, like in single principal-agent mechanism design problems under asymmetric information, it allows to state the sequential contracting problem as a sequence of programming problems in each of which a principal maximizes its expected payoff under incentive compatibility constraints.

In the next section we present an example that illustrates the main features of the result. In section 3 we set up the model of the sequential common agency. Then in sections 5 and 4 we establish the Revelation Principle for the case of private and public communication. Section 8 concludes.

2 Example

In this section we present a simple example of sequential contracting that illustrates the issues addressed in the paper. Two manufacturers, P_1 and P_2, contract sequentially with a common supplier A for provision of an essential input. The manufacturers produce two goods that they sell on different downstream markets. The contracting game lasts for two periods. Each period $i = 1, 2$ P_i contracts with A for the provision of a quantity of input $q_i \geq 0$ at price t_i that allows him to produce at most q_i units of a final good. Denote $C_i = (q_i, t_i)$ the contract between the P_i and A. The outcome of the contract between P_1 and A is observed by P_2 who makes an offer in the beginning of second stage. The inverse demand function in P_i's downstream market is $P(q_i) = 1 - q_i$.

The supplier produces an input at constant marginal cost θ which is her private information. It may be a low cost $\underline{\theta}$ with probability ν or a high cost $\overline{\theta}$ with probability $1 - \nu$, where $0 < \nu < 1$ and $\Delta\theta \equiv \overline{\theta} - \underline{\theta} > 0$.

The profit made with P_i equals $u_i = t_i - \theta q_i$, so A's total profit from serving the two manufactures is $u = u_1 + u_2$. The profit of each manufacturer is $v_i = (1 - q_i)q_i - t_i$.

In the absence of information asymmetries about the cost the contracting decisions of manufacturers are independent. Each P_i captures all the joint surplus with A by offering a contract $t_i = \theta q_i$ and producing an efficient quantity

$$q^*(\theta) = \tfrac{1}{2}(1 - \theta).$$

P_i thus obtains $v_i(\theta) = \frac{1}{4}(1 - \theta)^2$, and A gets no rent, $u_1 = u_2 = 0$.

When the cost of the supplier is not known to the manufacturers, each manufacturer would like to screen the supplier in order to base his market strategy on the cost. However, in an otherwise symmetric situation, by observing the outcome of contracting between P_1 and A, P_2 receives an additional signal about the supplier resulting in updated beliefs $\mu = \Pr(\underline{\theta}| C_1)$. This information is valuable for P_2 but may be disadvantageous for A. By decreasing the uncertainty of P_2 about A, P_1 allows P_2 to extract a bigger share of their joint surplus. Thus, to reveal any information to P_1, A has to be compensated for the loss of the information advantage with P_2. This information externality between the two contracts increases the cost of information revelation for P_1. It thus affects P_1's trade-off between efficiency and the informational rents. This leads to the question: How much information P_1 would like to acquire from A?

For example, P_1 may abstain from revealing (and learning) any information by offering a single contract for both types of cost. As a result, P_1 does not incur the cost of information revelation, but ignores the supplier's cost. To be accepted by both types, P_1 must pay at least $t = \overline{\theta} q$. Then, regardless the type, P_1 produces the efficient quantity for high cost $\overline{q}^*$ and leaves to a low cost A a positive rent $u^* = \Delta\theta\ \overline{q}^*$. The profit of P_1 under this contract is $v_1^P = \frac{1}{4}(1-\overline{\theta})^2$, which corresponds to the profit of dealing with a high cost A under full information. Obviously, this contract has high efficiency costs due to the low production level asked from a low cost supplier.

Under this contract, the outcome of the first stage provides no new information to P_2, who thus contracts with A under the prior belief $\mu = \nu$. The best contract for P_2 is a screening mechanism that makes the production contingent on the value of the cost. To induce the low cost type to reveal her information, P_2 has to leave her the rent that she can obtain by overstating the cost. The contract of P_2 must therefore satisfy the following incentive compatibility condition:

$$\underline{u}_2 \geq \overline{u}_2 + \Delta\theta\overline{q}_2. \tag{1}$$

The optimal trade off between rent extraction and efficiency (see Laffont and Martimort (2002)) results in a downward distortion in the output $\overline{q}_2$:

$$\overline{q}_2 = \overline{q}^{**} = \tfrac{1}{2}(1 - \overline{\theta} - \frac{\nu}{1-\nu}\Delta\theta),$$

and a positive rent for the low cost A: $u^{**} = \Delta\theta\overline{q}^{**}$. The total rent of the low cost A is thus $u^P = u^* + u^{**}$, where P stands for pooling.

Alternatively, P_1 can design a contract with distinct outcomes for each type of A. Then the outcome of the first stage allows P_2 to infer perfectly the type of A, implying that A gains no rent at the second stage. To induce her to reveal the information, P_1 must therefore compensate A for the rent she could obtain by overstating the cost in each of the two stages. If a low cost A selects the contract designed for the high cost type, P_2 is persuaded that he is facing a high cost supplier. This strategy allows A to gain u^* at the second stage, and u^{**} at the first stage. Therefore, the total cost of information revelation to P_1 under separating contract equals to $u^S = u^* + u^{**}$.

The benefit of information for P_1 is the efficiency gains of production: P_1 produces an efficient quantity $\underline{q}^*$ when dealing with a low cost supplier, and a conditionally efficient quantity $\overline{q}^{**}$ when dealing with a high cost supplier. The profit of P_1 is $v_1^S = \nu[\frac{1}{4}(1-\underline{\theta})^2 - (u^* + u^{**})] + (1-\nu)[\frac{1}{4}(1-\overline{\theta})^2 - (\frac{\nu}{1-\nu}\Delta\theta)^2]$.

The two examples of contracts presented illustrate how the information externality affects the equilibria of the game. Also it suggest that the ability of P_1 to alter the information transmitted to P_2 affects the incentives of A to reveal it to P_1. Then the natural questions are: What is the optimal degree of revelation? and What is the best strategy for P_1, in the absence of any ad hoc restrictions on the class of contracts from which he can choose.

3 The model

We consider a dynamic game between N principals, $P_1, ..., P_N$, and a single agent, A. There are N stages. At each stage P_i contracts with A over an allocation $x_i \in X_i$. The outcome of the contracting game defines an allocation $x = (x_1, ..., x_N) \in X = X_1 \times ... \times X_N$, where all X_i, $i = 1, ..., N$ are assumed to be metric spaces. Denote $x_i^- \equiv (x_1, ..., x_i)$ the outcomes of contracting up to period i and $x_{i+1}^+ \equiv (x_{i+1}, ..., x_N)$ the outcomes of contracting from period $i+1$ till period N. The decisions of principals $\mathrm{P}_1, ..., \mathrm{P}_i$ may restrict the feasible choice of principal P_{i+1}. To account for this feature we assume that once P_i implements allocation x_i, the feasible choice of P_{i+1} is restricted to $F_{i+1}(x_i^-)$, where $F_{i+1}(\cdot)$ is a correspondence $F_{i+1} : X_i \Rightarrow X_{i+1}$. The agent has ex-ante private information about its type $\theta \in \Theta = (\theta_1, ..., \theta_T)$, where $2 \leq T < \infty$ that is persistent through N stages. In Section 7 we show how the results extend to the case of non-persistent private information. The prior distribution of types, $\gamma = (\gamma_1, ..., \gamma_T)$, with $\gamma_t > 0$ for $t = 1, ..., T$ and $\sum_t \gamma_t = 1$, is common knowledge. Denote Θ_i the set of types that play at stage i with a positive probability.

We consider communication mechanisms (contracts) which are functions from messages to probability distributions over feasible allocations. A mechanism of principal P_i, Γ_i, consists of a message space M_i and a decision rule $\beta_i(\cdot)$: For each message $m_i \in M_i$, P_i commits to a decision $\beta_i(m) \in \Delta_i$, where Δ_i is the set of probability distributions over $F_i(x_{i-1}^-)$. M_i is assumed to be a metric space, and $\mathcal{M}_i$ denotes the Borel $\sigma-$ algebra on M_i. The decision is a measurable mapping $\beta_i : M_i \to \Delta_i$. Denote $\beta_i^- \equiv (\beta_1, ..., \beta_i)$ the mechanisms proposed by principals $\mathrm{P}_1, ..., \mathrm{P}_i$.

The strategy of P_i is the choice of a mechanism Γ_i. Contracts are incomplete in that P_i cannot contingent its contract on the decisions taken by the other principals

P_{-i}. The agent's strategy at stage i is a message to P_i. Formally it is described by a mapping from the type-contract space to the space S_i of probability measures over $\mathcal{M}_i$, $\sigma_i : \Theta \times \Gamma_1 \times ... \times \Gamma_{i-1} \to S_i$. Denote $\overline{\sigma}_i \equiv \Sigma\gamma_t\sigma_{i,t}$ and note that $\overline{\sigma}_i \in S_i$. Also denote $\overline{\beta}_i = \Sigma\gamma_t\beta_{i,t}$ and note that $\beta_i \in \Delta_i$.

The payoff of P_i depends on the allocation x and on the type of agent θ_t. Denote $v_{i,t}(x_j^-)$ the payoff of P_i at stage j when the agent is of type θ_t. It should be emphasized that P_i controls directly only the allocation x_i. However, x_i may have an indirect impact on allocations x_{i+1}^+ through two channels: by affecting the feasible choice of P_{i+1}^+ and by changing the perception of P_{i+1}^+ about the agent's type. The payoff of P_i at stage i equals to $v_{i,t}(x_i^-)$, and his overall payoff is

$$v_t(x) = \sum_{i=1}^{N} v_{i,t}(x_i^-).$$

Similarly, the payoff of the agent θ_t at stage i is $u_{i,t}(x_i^-)$, so its overall payoff is given by

$$u_t(x) = \sum_{i=1}^{N} u_{i,t}(x_i^-).$$

The functions $v_i(\cdot)$ and $u_i(\cdot)$, $i = 1, ..., N$ are continuous and bounded on their domains.

The timing of the game is the following:

- The agent learns its type $\theta \in \Theta$.
- At each stage i, $i = 1, ..., N$, P_i offers A to play Γ_i, that results in allocation x_i.
- Once Γ_i is played, principal P_{i+1} observes information I_i on past contracting activities (I_i is specified below) and updates beliefs about the agent's private information to p_i.
- At stage $N + 1$ the game ends.

The information of the agent at stage i consists of its type θ, the mechanisms β_i^- offered by P_i^-, the profile of messages $m_{i-1}^- \equiv (m_1, ..., m_{i-1})$ sent to P_{i-1}^-, and the profile of outcomes x_{i-1}^- realized at stages $1, ..., i-1$. Denote h_i^A the history of the game for the agent, where

$$h_i^A = (m_{i-1}^-, \beta_{i-1}^-, x_{i-1}^-, \beta_i, \theta).$$

For example, at $i = 2$, $h_i = (m_1, \beta_1(m_1), x_1, \beta_2, \theta)$.

The information of P_i at stage i, I_i, has at most three components. The first one is the sequence of mechanisms $\Gamma_i^- \equiv (\Gamma_1, ..., \Gamma_i)$ that were offered by the preceding principals P_i^-. The second one is the sequence of allocations that resulted from these mechanisms, x_i^-. Finally, the third one is the sequence of messages that were communicated by the agent to the preceding principals P_{i-1}^-, m_i^-. We assume that Γ_i^- and x_i^- become common knowledge at stage i. For the sequence of messages we distinguish between *public* and *private* communication. Under *public communication*, P_i observes all the messages that has been sent by the agent to the preceding principals P_i^-. In this case the history of the game for P_i is

$$h_i(P_i,\ public) = (m_{i-1}^-, \Gamma_{i-1}^-, x_{i-1}^-).$$

Given that P_i observes the sequence of messages m_{i-1}^-, it can also infer the decisions β_{i-1}^- that were implemented by P_{i-1}^-. Under *private communication*, the message reported by A to P_i is their private information. Then the history of P_i is

$$h_i(P_i, private) = (\Gamma_{i-1}^-, x_{i-1}^-),$$

that is, it is composed only of the profile of mechanisms offered by the preceding principals and the profile of the realized allocations. Denote $(\beta, \sigma) \equiv (\beta_i, \sigma_i)_{i=1}^N$ the strategy profile for the principals and the agent in the game $\Gamma = \{\Gamma_1, ..., \Gamma_N\}$, $(\beta_i^+, \sigma_i^+) \equiv (\beta_k, \sigma_k)_{k=i}^N$ the strategy profile at stages $k = i, ..., N$, and $(\beta_i^-, \sigma_i^-) \equiv (\beta_k, \sigma_k)_{k=1}^i$ the strategy profile at stages $k = 1, ..., i$.

The observed history h_i results in updating of beliefs concerning the type of the agent. The posterior belief of P_i is a measurable mapping $p_i : I_{i-1} \to P_i$, where $P_i = \{p \in \mathbf{R}_+^{|\Theta_i|} \mid p_j > 0, \underset{j}{\Sigma} p_j = 1\}$ is the set of probability distributions over Θ_i.

At stage i, P_i faces a history $h_i(P_i)$ and a state (x_{i-1}^-, p_i), and offers a mechanism Γ_i. The outcome of the mechanism Γ_i determines the history $h_{i+1}(P_{i+1})$ and the beliefs $p_{i+1} : I_i \to P$, resulting in the subsequent state (x_i^-, p_{i+1}).

We study the Perfect Bayesian Equilibria of the game.

Definition 1 *A Perfect Bayesian Equilibrium of the game Γ consists of a profile (β, σ) and beliefs $(\gamma, p_1, ..., p_{N-1})$ that satisfy the following three conditions:*

1. ***Optimality of*** β*. The mechanism of P_i, $i = 1, ..., N$ is optimal, anticipating its*

impact on the continuation game Γ^+_{i+1}, that is, for every $m \in M_i$,

$$\sum_{\Theta_i} p_{i,t} \int_{\Delta_i} \int_{M_i} v_{i,t}(x^-_{i-1}, x_i(m), x^+_{i+1}(x_i(m), p_i(m))) d\sigma_{i,t}(m) d\beta_i(m)$$

$$\geq \sum_{\Theta_i} p_{i,t} \int_{\Delta_i} \int_{M_i} v_{i,t}(x^-_{i-1}, x_i(m), x^+_{i+1}(x_i(m), p_i(m))) d\sigma_{i,t}(m) d\beta'_i(m), \text{ for all } \beta'_i \in \Delta_i.$$

2. ***Optimality of*** σ. *The reporting strategy of A is optimal at each stage i, anticipating its impact on the continuation game Γ^+_{i+1}, described below in (3).*

3. ***Bayes rule.*** *The posterior belief of P_i is consistent with the Bayes rule, described below in (4) and (5).*

Let us consider a given (possibly state dependant) sequence of message spaces M. For all types θ_t that play at stage i with a positive probability, $p_{i,t} > 0$, the expected payoffs of P_i contracting with type θ_t, and the payoff of A are, respectively,

$$V_{i,t}(\sigma_{i,t}, p_i, \beta_i, x^-_{i-1} | M_i) \equiv \sum_{j=1}^{i-1} v_{i,t}(x^-_j)$$

$$+ \sum_{j=i}^{N} \int_{\Delta_i} \int_{M_i} v_{i,t}(x^-_{j-1}, x_i(m), x^+_{j+1}(x_i(m), p_i(m))) d\sigma_{i,t}(m) d\beta_i(m),$$

$$U_{i,t}(\sigma_{i,t}, p_i, \beta_i, x^-_{i-1} | M_i) \equiv \sum_{j=1}^{i-1} u_{i,t}(x^-_j)$$

$$+ \sum_{j=i}^{N} \int_{\Delta_i} \int_{M_i} u_{j,t}(x^-_{j-1}, x_i(m), x^+_{j+1}(x_i(m), p_i(m))) d\sigma_{i,t}(m) d\beta_i(m).$$

Then the expected payoff of P_i at state (x^-_{i-1}, p_i) is

$$\sum_{\Theta_i} p_{i,t} V_{i,t}(\sigma_{i,t}, p_i, \beta_i, x^-_{i-1} | M_i). \tag{2}$$

The objective of P_i is to choose β_i, p_i and σ_i to maximize her expected payoff (2) subject to three constraints: First, the agent's reporting strategy is optimal, anticipating its impact on the subsequent states:

$$U_{i,t}(\sigma_{i,t}, p_i, \beta_i, x^-_{i-1} | M_i) \geq U_{i,t}(\sigma'_{i,t}, p_i, \beta_i, x^-_{i-1} | M_i) \text{ for all } \sigma_i \in S_i. \tag{3}$$

Second, the mechanisms offered by subsequent principals belong to PBE of the game, that is, β^+_{j+1} is such that each β_j, $j \geq i+1$, maximizes

$$\sum_{\Theta_j} p_{j,t} V_{j,t}(\sigma_{j,t}, p_j, \beta_j, x^-_{j-1} \big| M_j).$$

Third, for all types in Θ_i, the belief of P_{i+1} is consistent with the Bayes rule. The definition of the Bayes rule depends on the information available to P_{i+1}.

Under private communication, the belief of P_{i+1} is derived from the mechanisms Γ_i, the behavioral strategy of the agent σ_i and the allocations implemented at previous stages, x^-_i. The Bayes rule in this case writes:

$$\int_H p_{i,t}(x) d\overline{\beta}_i(x) = p_{i-1,t}\beta_{i,t}(H) \text{ for all } H \in \Delta_i \text{ with } \overline{\beta}_i(H) > 0. \qquad (4)$$

Under public communication, in addition to Γ_i, x^-_i and σ_i, P_{i+1} observes the message m_i reported by the agent to P_i. Thus the Bayes rule writes:

$$\int_H p_{i,t}(m) d\overline{\sigma}_i(m) = p_{i-1,t}\sigma_{i,t}(H) \text{ for all } H \in \mathcal{M}_i \text{ with } \overline{\sigma}_i(H) > 0. \qquad (5)$$

To interpret (5), divide each side of the expression by $\overline{\sigma}_i(H) > 0$. Then the left hand side represents the belief of P_{i+1} to face a type θ_t upon receiving a message from the set H. The right hand side is the conditional probability that A is of type θ_t when A follows the reporting strategy σ_i and the message from the set H is realized[2]. The only difference in (4) is that a principal updates beliefs upon observing realized allocations instead of messages.

Our objective is to construct a set of tractable mechanisms that provide the same payoff for all the players as the original game Γ. For this reason, we introduce some further definitions that permit to order mechanisms in terms of payoffs obtained by P_i and A. We say that $(x^-_{i-1}, \sigma_i, p_{i+1}, \beta_i, \beta^+_{i+1} \big| M_i)$ is *incentive feasible* if σ_i is an optimal strategy of the agent, so it satisfies (3), and p_{i+1} is derived from the Bayes rule (4) or (5), depending on the communication mode. $(x^-_{i-1}, \sigma_i, p_{i+1}, \beta_i, \beta^+_{i+1} \big| M_i)$ is *incentive efficient* if there is no incentive feasible $(x^-_{i-1}, \sigma'_i, p'_{i+1}, \beta'_i, \beta^+_{i+1} \big| M_i)$ such that

$$\sum_{\Theta_i} p_{i,t}[V_{i,t}(x^-_{i-1}, \sigma'_i, p'_{i+1}, \beta'_i, \beta^+_{i+1} \big| M_i) - V_{i,t}(x^-_{i-1}, \sigma_i, p_{i+1}, \beta_i, \beta^+_{i+1} \big| M_i)] > 0$$
$$\text{and } U_{i,t}(x^-_{i-1}, \sigma'_i, p'_{i+1}, \beta'_i, \beta^+_{i+1} \big| M_i) = U_{i,t}(x^-_{i-1}, \sigma_i, p_{i+1}, \beta_i, \beta^+_{i+1} \big| M_i).$$

[2] As noted in BS, by Radon-Nikodym's Theorem (see, e.g. Stockey and Lucas (1989)), equation (4) defines p_{i-1} uniquely $\overline{\sigma}_i$- almost everywhere.

Finally, $(x_{i-1}^{-}, \sigma_i, p_{i+1}, \beta_i, \beta_{i+1}^{+} \big| M_i)$ and $(x_{i-1}^{-}, \sigma_i', p_{i+1}', \beta_i', \beta_{i+1}^{+} \big| M_i)$ are *payoff equivalent* if

$$\sum_{\Theta_i} p_{i,t}[V_{i,t}(x_{i-1}^{-}, \sigma_i', p_{i+1}', \beta_i', \beta_{i+1}^{+} \big| M_i) - V_{i,t}(x_{i-1}^{-}, \sigma_i, p_{i+1}, \beta_i, \beta_{i+1}^{+} \big| M_i)] = 0$$
$$\text{and } U_{i,t}(x_{i-1}^{-}, \sigma_i', p_{i+1}', \beta_i', \beta_{i+1}^{+} \big| M_i) = U_{i,t}(x_{i-1}^{-}, \sigma_i, p_{i+1}, \beta_i, \beta_{i+1}^{+} \big| M_i).$$

In the rest of the paper we will analyze how a given equilibrium profile (β, σ) of the game with unrestricted message spaces relates to an equilibrium profile of the game in which the principals are restricted to use messages form the type space Θ. The classical result of the mechanism design is the Revelation Principle which states that the two equilibria sets are equivalent. In terms of above definitions, the standard Revelation Principle can be stated as follows.

Revelation Principle. *Suppose $N = 1$. If $(\sigma, \beta | M)$ is incentive feasible, then there exists a direct mechanism $\Gamma^d = (\Theta, \beta^d)$ and an incentive feasible $(\sigma^d, \beta^d \big| \Theta)$ such that $(\sigma, \beta | M)$ and $(\sigma^d, \beta^d \big| \Theta)$ are payoff-equivalent. Moreover, it is an optimal strategy for the agent to reveal its type, $\sigma_t(\theta_t) = 1$ for all $\theta_t \in \Theta$.*

The intuition behind the standard Revelation Principle is that the principal can replicate the behavior of the agent by combining two functions $\sigma : \Theta \to M$ and $\beta : M \to \Delta(X)$ to a single function $\beta^d \equiv \beta \circ \sigma : \Theta \to \Delta(X)$. Then the mechanism β^d induces the same probability distribution over allocations as the original mechanism β with the reporting strategy σ. The reason why the Revelation Principle cannot be applied directly to the game with many principals is that each principal P_i can commit only to its own mechanism β_i, and not to the whole game $\beta = (\beta_1, ..., \beta_N)$. (At the same time, when P_i contracts with the agent, its contract induces a continuation game with an outcome that can be characterized as a PBE.) In the following two sections we establish the Revelation Principle for the cases of public and private communication.

4 Public Communication

4.1 The Revelation Principle

In this section we characterize the equilibria of the game under public communication. When the agent contracts with a principal under public communication, the report sent to P_i is observed by all P_i^+. Therefore, when selecting its reporting strategy, the agent is concerned not only with the impact of the report on the decision of P_i, but also of all the subsequent principals. It implies that communication at each stage cannot be considered independently as under private communication.

Another feature of this setting is that all the information transmitted by the agent to P_i by the means of message m_i is revealed directly to P_i^+. If P_i finds it beneficial to preserve some uncertainty about the agent's type for the subsequent principals, this strategy cannot be achieved when the agent reports its type truthfully. In the example of Section 2, if the supplier submits a truthful report about her type, P_2 becomes perfectly informed. Note that offering a stochastic contract does not provide the remedy against full revelation because the lottery is assigned after A submits the report to P_1, and therefore, after P_2 learns the type. Hence, if P_1 prefers to control the information revealed to P_2, there must be at least some types that follow a non-degenerate stochastic reporting strategy.

It may seem that the characterization of the implementable allocations faces a serious problem because there can be mechanisms that are not supported by the direct incentive compatible mechanism. However, the information that a principal aims to infer from the agent through the communication mechanism is only the agent's type. The reason why the direct revelation mechanism may not support the principal's optimal mechanism is that a truthful deterministic report on the type can be suboptimal. One way to circumvent this problem is to reduce the informativeness of the report about the type. In other words, the agent must be allowed to misreport its private information with some probability. Then communication strategy itself becomes stochastic. However, randomizing over the type space is sufficient to generate any belief for P_{i+1}^+, and the equilibria of the game can still be characterized within the type space.

To provide a formal proof of the argument, we apply the technique developed by Bester and Strausz (2001) who study a general problem of contracting under imperfect commitment in a long term principal - agent relationship. They show that any optimal mechanism can be characterized within the class of direct mechanisms in which an agent reports its true type with a positive probability. To reduce the cost of information revelation, the principal commits to a gradual learning policy.

The major difference between contracting under imperfect commitment and sequential contracting with many principals is that the objectives of different principals are not aligned. Each principal does not internalize the externality that its contract imposes on the other principals. However, this difference is not crucial for considering the equilibria within the type space. Under imperfect commitment, when a principal contracts with an agent in period i, he is constrained to allocations that arise as Perfect Bayesian equilibria of the continuation game. Basically, when contracting in period i, the principal anticipates that he will react to the outcome of today's contract by offering the best contract in period $i+1$. Each period the principal anticipates that his different selves will behave optimally in the subsequent periods.

A similar situation arises when a number of principals contract sequentially with an agent. At each period i a principal P_i anticipates that the subsequent principals P_{i+1}^{+} will behave optimally given the state and information induced by its contract. Thus P_i is constrained to allocations that arise as Perfect Bayesian equilibria of the game between P_{i+1}^{+} and A. One potential obstacle with replacing a general mechanism with a direct one is that out-of-equilibrium messages may be important to support the optimal mechanism. However, we show that this problem can be circumvented by preserving the allocations that arise out-of-equilibrium by the means of a direct mechanism.

To characterize the equilibria of the game, we proceed in the following steps. First we note that at stage N the standard Revelation Principle applies. At stage $N-1$ we apply the technique developed in Bester and Strausz (2001) to replace an original mechanism $(\beta_{N-1}, \sigma_{N-1})$ with an incentive feasible and payoff equivalent direct mechanism $(\beta_{N-1}^{d}, \sigma_{N-1}^{d})$ with $\sigma_{N-1,t}^{d}(\theta_t) > 0$ for all $\theta_t \in \Theta$. Also we show that when P_{N-1} offers a direct mechanism, P_{N-2}^{-} and A do not deviate from the original equilibrium profile $(\beta_{N-2}^{-}, \sigma_{N-2}^{-})$ at stages $1, ..., N-2$. By iterating the argument for all $i = 1, ..., N$, we conclude that the equilibria allocations of the original game with unrestricted message spaces can be characterized within a class of direct mechanisms.

It is straightforward to show that the standard Revelation Principle applies at stage N.

Lemma 1 *For any equilibrium profile (β_N, σ_N) there exists an incentive feasible (β_N^d, σ_N^d) with a message space Θ and $\sigma_{N,t}^{d}(\theta_t) = 1$ for all $\theta_t \in \Theta$. Moreover, anticipating the direct mechanism of P_N, P_{N-1}^{-} and A do not deviate from the original profile $(\beta_{N-1}^{-}, \sigma_{N-1}^{-})$.*

At stage $N-1$, P_{N-1} faces a state (x_{N-2}^{-}, p_{N-1}) and anticipates that P_N offers a direct mechanism β_N^d. To construct a direct mechanism at stage $N-1$, we first construct a direct mechanism for all types $\theta_t \in \Theta_{N-1}$ that reach stage $N-1$ with a positive probability. Then in Proposition 3 we extend it to the original message space Θ.

Let us first focus on the set of types $\theta_t \in \Theta_{N-1}$ that reach the stage $N-1$ with a positive probability. Following Bester and Strausz (2001), the direct mechanism (β_{N-1},σ_{N-1}) can be constructed in two steps. First step is to show that there exists an incentive feasible profile with a message space M'_{N-1} that contains at most $T_{N-1} = |\Theta_{N-1}|$ messages and is payoff equivalent to the original profile. Second step is to apply the Marriage Theorem to the reduced message space and construct a direct mechanism (β^d_{N-1},σ^d_{N-1}).

Proposition 1 *Let* $(x^-_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1}, \beta_N | M_{N-1})$ *be incentive efficient. Then there exists an incentive feasible* $(x^-_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1}, \beta_N | M'_{N-1})$ *and a finite set* $M'_{N-1} = \{m_1, ..., m_{T_{N-1}}\} \in \mathcal{M}_{N-1}$ *with* $|M'_{N-1}| \leq T_{N-1}$ *and* $\overline{\sigma}'_{N-1}(M'_{N-1}) = 1$ *such that*

$(x^-_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1}, \beta_N | M_{N-1})$ and $(x^-_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1}, \beta_N | M'_{N-1})$ are payoff-equivalent. Moreover, the vectors $\sigma'_{N-1}(m_h) = (\sigma'_{N-1}(m_h), ..., \sigma'_{N-1}(m_h))$, $h = 1, ..., |M'_{N-1}|$ are linearly independent.

The basic idea of Proposition 1 is that when P_{N-1} uses more messages than there are types, the vectors of the agent's reporting strategy under the original mechanism $\{\sigma_h\}$ must be linear dependent. Since the agent is indifferent between all the messages sent with a positive probability, without affecting the incentives of A, P_{N-1} can distribute the weight from some messages so as to reduce the dimension of the message space to the type space. An important implication of Proposition 1 is that P_{N-1} does not need to use a message space of a dimension higher than the type space.

As an illustration, suppose that A can be one of two types, $\theta_t \in \{\overline{\theta}, \underline{\theta}\}$, and the original mechanism of P_{N-1} uses three messages $\{m_1, m_2, m_3\}$. Proposition 1 states that there exists a reporting strategy σ' that supports a perfect Bayesian equilibrium and contains only two messages. The result trivially holds if some message in $\{m_1, m_2, m_3\}$ is sent with zero probability by both types. Suppose now that each message is in the support of at least some type. Consider vectors

$$\sigma_1 = \begin{pmatrix} \overline{\sigma}_1 \\ \underline{\sigma}_1 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} \overline{\sigma}_2 \\ \underline{\sigma}_2 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} \overline{\sigma}_3 \\ \underline{\sigma}_3 \end{pmatrix},$$

where $\overline{\sigma}_h$ (respectively, $\underline{\sigma}_h$), $h = 1, 2, 3$ represents the probability that type $\overline{\theta}$ (respectively, $\underline{\theta}$) sends a message h. Since there are only two types, the vectors $\sigma_1, \sigma_2, \sigma_3$ are linearly dependent. So, there exists a non-zero vector $\mu = (\mu_1, \mu_2, \mu_3) \neq 0$ such that

$$\sum_h \mu_h \sigma_h = 0. \tag{6}$$

For any given μ let us define

$$\lambda^t \equiv -\frac{1}{\min_h \mu_h} \text{ and } \lambda^b \equiv -\frac{1}{\max_h \mu_h}. \tag{7}$$

Consider a new reporting strategy of the A to P_{N-1} such that

$$\sigma'_h(\theta_t;\lambda) = (1+\lambda\mu_h)\sigma_h(\theta_t) \text{ with } \lambda \in [\lambda^b, \lambda^t]. \tag{8}$$

$\sigma'_h(\theta_t;\lambda)$ indeed constitutes a strategy of type θ_t: $\sigma'_h(\theta_t;\lambda) \geq 0$ and from condition (6) we obtain $\sum_h \sigma'_h(\theta_t;\lambda) = \sum_h (1+\lambda\mu_h)\sigma_h(\theta_t) = 1$.

If a profile (p_N, σ, β_N) is an equilibrium under the mechanism β_{N-1} with a message space $\{m_1, m_2, m_3\}$, then the profile $(p_N, \sigma'(\lambda), \beta_N)$ is also an equilibrium under this mechanism. Indeed, a new reporting strategy $\sigma'(\lambda)$ induces the same posterior beliefs for P_N,

$$p_{t,N}(m_h;\lambda) = \frac{p_{t,N-1}(1+\lambda\mu_h)\sigma_h(\theta_t)}{\sum_j p_{j,N-1}(1+\lambda\mu_h)\sigma_h(\theta_j)} = p_{t,N}(m_h),$$

whenever $\sum_j p_{j,N-1}(1+\lambda\mu_h)\sigma_h(\theta_j) > 0$. It implies that the choice of β_N remains optimal for P_N. Furthermore, since any message m_h results in the same allocation of P_{N-1} and P_N, the agent is indifferent between the two strategies.

The payoff of P_{N-1} under the new reporting strategy is maximized when A uses at most two messages. The payoff of P_{N-1} writes

$$V_{N-1}(p_N, \sigma'(\lambda), \beta_N | \Gamma_{N-1}) = \sum_t p_{t,N-1} \sum_h \sigma'_h(\theta_t;\lambda) V_{N-1,t}(x^-_{N-2}, \beta_{N-1}, \beta_N)$$

and is linear in λ. Therefore, it is maximized by some $\lambda^* \in \{\lambda^t, \lambda^b\}$. Note that the strategy $\sigma'(\lambda^*)$ is distinct from the original strategy σ since $\lambda^b < 0 < \lambda^t$. And by construction, under λ^* at least one message is sent with zero probability. Hence there exists a reporting strategy $\sigma'(\lambda^*)$ that supports the same equilibrium as under the original mechanism and uses at most two messages.

The next step is to use the reduced message space M'_{N-1} to construct a direct mechanism. The idea is to associate each message in M'_{N-1} with some type. We rely on the following theorem attributed to Hall (1935).

Lemma 2 (Marriage theorem) *Let H be a finite non-empty set and K be a non-empty set, possibly infinite. Further, let $D : H \rightrightarrows K$ be a correspondence and for any set $G \subsetneq H$ define*

$$D(G) = \underset{h\in G}{\cup} D(h).$$

Then there exists a mapping $d : H \to K$ *such that:*

(i) $d(h) = d(k)$ *implies* $h = k$, *and*

(ii) $d(h) \in D(h)$ *for all* $h \in H$,

if and only if $|D(G)| \geq |G|$ *for all* $G \subseteq H$.

This combinatorial results can be applied to our setting as follows. Take the set H to be the reduced message space M'_{N-1}. For each subset G of messages in H, $D(G)$ denotes the set of types that send messages from G with a positive probability. The Marriage Theorem asserts that if the number of messages in G does not exceed the number of types that send messages from G with a positive probability, then there is a way to assign distinctively to each message a type that sends this message with a positive probability. Applying the results of Proposition 1 and Lemma 2 leads to the following proposition.

Proposition 2 *If* $(x^-_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ *is incentive efficient, then there exists a direct mechanism* $\Gamma^d_{N-1} = (\Theta_{N-1}, \beta^d_{N-1})$ *and an incentive feasible* $(x^-_{N-1}, \sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \Theta_{N-1})$ *such that* $(x^-_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ *and* $(x^-_{N-1}, \sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \Theta_{N-1})$ *are payoff equivalent. Moreover, with the direct mechanism each type reports its private information with a positive probability,* $\sigma'_{N-1,t}(\theta_t) > 0$ *for all* $\theta_t \in \Theta_{N-1}$.

The result of Proposition 2 is obtained in two steps. First, it can be verified that the conditions of the Marriage Theorem are satisfied for the reduced message space. Therefore, there exists a mapping $d : M'_{N-1} \to \Theta_{N-1}$ that assigns a type to each message. Then this mapping can be inverted to associate each type with a message that it sends with positive probability. The inverted mapping can be used to construct a direct mechanism in which the reporting strategy is a probability that type θ_i sends message θ_j, with the property that each type reports its private information with positive probability.

To complete the characterization, one has to extend the message space Θ_{N-1} to the set of types that play at stage $N-1$ with zero probability. The mechanisms offered to these types may be necessary to sustain the original equilibrium profile (β, σ). In the following proposition we extend the direct mechanism to the original message space Θ and verify that this extension preserves the out-of-equilibrium allocations of (β, σ). Hence we conclude that the behavior of P^-_{N-2} and A does not change when they anticipate that P_{N-1} will offer a direct mechanism.

Proposition 3 *For any incentive feasible* $(\sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ *there exists a direct mechanism* $\Gamma_{N-1} = (\Theta, \beta^d_{N-1})$ *and an incentive feasible* $(\sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \Theta)$ *such that*

$(\sigma_{N-1}, p_{N-1}, \beta_{N-1} \big| M_{N-1})$ *and* $(\sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} \big| \Theta)$ *are payoff equivalent.*

Iterating the argument of Propositions 2 and 3 leads to the main theorem of this section.

Theorem 1 *In the game with public communication, for any equilibrium profile* $(\beta, \sigma) \in PBE(\Gamma)$ *there exists an equilibrium profile* (β^d, σ^d)*, where* β^d *is a direct mechanism and* $\sigma^d_t(\theta_t) > 0$ *which is payoff equivalent to* (β, σ)*.*

The result of Theorem 1 asserts that the equilibria of a game under public communication can be characterized within the class of direct mechanisms in which A reports its type with a positive probability. It reduces substantially the complexity of the problem. The game can be solved backwards. At each stage i P_i offers a direct mechanism that maximizes its payoff subject to (*i*) the behavior of the subsequent principals, (*ii*) the incentive compatibility constraints of the agent, and (*iii*) complementarity conditions on the reporting strategy that guarantee that the agent is indifferent among all the messages it sends with a positive probability.

4.2 Structure of the optimal mechanism

In general an optimal mechanism of P_i is a lottery over F_i. In this section we show that if a principal does not offer a stochastic contract in an isolated game, then an optimal contract in a multistage game is also deterministic. By an isolated stage game we mean a game played between a principal P_i and an agent in the absence of the other principals P_{-i}. For any given x_{-i}, an optimal contract of this game can be characterized by a direct mechanism and it is a solution to the following program.

$$\max \sum_{\Theta_i} p_{i,t} \int_{F_i} v_{i,t}(x^-_{i-1}, x, x^+_{i+1}) d\beta_t(x) \tag{9}$$
$$s.t.\ \beta_t \in \arg\max_{\{\beta_j\}} \int_{F_i} u_{i,t}(x^-_{i-1}, x, x^+_{i+1}) d\beta_j(x).$$

Proposition 4 *If a principal* P_i *offers a deterministic contract in an isolated game 9, then for each message sent with a positive probability at stage* i *of a multistage game,* P_i *assigns a deterministic allocation.*

Proof. Consider a multistage game. An optimal contract of P_i at stage i is a solution to the program

$$\max \sum_{t\in\Theta_i} p_{i,t} \sum_{j\in\Theta_i} \sigma_{j,t} \int_{F_i} v_{i,t}(x_{i-1}^{-}, x, x_{i+1}^{+}(x, p_{i+1}))d\beta_t(x)$$
$$s.t.\ \sigma_{j,t} \in \arg\max_{\{\sigma'_{j,t}\}} \sum_{j\in\Theta_i} \sigma'_{j,t} \int_{F_i} u_{i,t}(x_{i-1}^{-}, x, x_{i+1}^{+}(x, p_{i+1}))d\beta_j(x).$$

Since an isolated game an optimal contract is deterministic, for each message reported with a positive probability, an optimal contract of P_i must also be deterministic. Indeed, it is an optimal choice of P_i. Also it does not affect beliefs of P_{i+1}^{+}, and therefore the reporting strategy of the agent at stage i. ■

This simple result is very useful in applications. It states that at each stage a principal does not need to offer a contract that contains more allocations than the dimension of $|\Theta_i|$. Strausz (2004) analyzes the conditions under which deterministic mechanism is optimal in an isolated stage game. He demonstrates that if an optimal deterministic mechanism does not involve bunching, then it is also optimal within a general class of stochastic mechanisms.

We apply the results of this section to characterize an optimal contract of the example of Section 2

Example. Under public communication, the report of A to P_1 is observed by P_2. So, if the low cost supplier reports truthfully its type to P_1, P_2 infers perfectly the type of A and offers a contract under full information. This policy is costly to P_1 who has to leave the low cost type the rent $\Delta\theta\overline{q}_2^*$, where $\overline{q}_2^*$ is the full information quantity offered to high cost type at the second stage. By reducing the informativeness of the report of high cost supplier, P_1 will induce P_2 to introduce downward distortion of the output of this type, and consequently, decrease the rent paid to the low cost supplier. The optimal contract of P_1 thus consists of a menu of two allocations $(\underline{t}_1, \underline{q}_1)$ and $(\overline{t}_1, \overline{q}_1)$ and the following reporting strategies. A high cost type reports its true type with probability one. A low cost type reveals the true type with probability σ, $\sigma = \Pr(\widetilde{\theta} = \underline{\theta} \big| \underline{\theta})$. P_1 assigns a contract $(\underline{t}_1, \underline{q}_1)$ when A reports $\underline{\theta}$, and $(\overline{t}_1, \overline{q}_1)$ when A reports $\overline{\theta}$. When P_2 observes message $\underline{\theta}$, he infers that the supplier has low cost. When he observe the message $\overline{\theta}$, he holds beliefs

$$\mu = \frac{\nu(1-\sigma)}{\nu(1-\sigma) + 1 - \nu}.$$

In the second stage, the low cost supplier produces the efficient quantity $\underline{q}^*$. The quantity of the high cost supplier is distorted downwards, but now the value of the distortion depends on the reporting strategy σ.

$$\begin{aligned} \underline{q}_2 &= \underline{q}^*, \\ \overline{q}_2(\sigma) &= \tfrac{1}{2}(1-\overline{\theta}-\Delta\theta\frac{\nu}{1-\nu}(1-\sigma)). \end{aligned} \tag{10}$$

The lower is the σ, the lower is the informativeness of message $\overline{\theta}$ that the agent has high cost. As a result, the higher is the distortion of the quantity of the high cost type at the second stage.

The optimal contract of P_1 solves the program:

$$\begin{aligned} \max_{(t_1,q_1,\sigma)} \quad & \nu\sigma[(1-\underline{q}_1)\underline{q}_1-\underline{t}_1]+\nu(1-\sigma)[(1-\overline{q}_1)\overline{q}_1-\overline{t}_1] \\ & +(1-\nu)[(1-\overline{q}_1)\overline{q}_1-\overline{t}_1] \\ & \underline{IC}: \quad \underline{t}_1-\underline{\theta}\underline{q}_1 \geq \overline{t}_1-\underline{\theta}\overline{q}_1+\Delta\theta\overline{q}_2, \\ & \overline{IC}: \quad \overline{t}_1-\overline{\theta}\overline{q}_1 \geq \underline{t}_1-\overline{\theta}\underline{q}_1, \\ & \underline{PC}: \quad \underline{t}_1-\underline{\theta}\underline{q}_1 \geq 0, \\ & \overline{PC}: \quad \overline{t}_1-\overline{\theta}\overline{q}_1 \geq 0, \\ & \underline{CC}: \quad (1-\sigma)[\underline{t}_1-\underline{\theta}\underline{q}_1-(\overline{t}_1-\underline{\theta}\overline{q}_1+\Delta\theta\overline{q}_2)]=0. \end{aligned}$$

The last constraint $\underline{CC}$ is the complementarity condition. It states that when a low cost type is has a non-degenerate reporting strategy $\sigma \neq 1$, it must be indifferent between messages $\underline{\theta}$ and $\overline{\theta}$.

Compared to the incentive constraint of under full separation (1), the incentive constraint $\underline{IC}$ of the above problem is relaxed because $\Delta\theta\overline{q}_2(\sigma) < \Delta\theta\overline{q}^*$. However, introducing this noise is not costless. With probability $1-\sigma$ P_1 assigns an inefficient production quantity $\underline{q}_1$ to the efficient supplier. An optimal contract trade-offs the benefits of a reduced rent with the efficiency costs of the lottery. The incentive constraint $\underline{IC}_1$ and the participation constraint $\overline{PC}_1$ are the binding constraints. Then, the first order (sufficient) conditions with respect to $\underline{q}_1$, $\overline{q}_1$ and σ imply:

$$\begin{gathered} \underline{q}_1 = \underline{q}^*, \\ \overline{q}_1 = \tfrac{1}{2}(1-\overline{\theta}-\frac{\nu\sigma}{1-\nu\sigma}\Delta\theta), \\ [(1-\underline{\theta}-\underline{q}_1)\underline{q}_1-\Delta\theta(\overline{q}_1+\overline{q}_1(\sigma))]-(1-\overline{\theta}-\overline{q}_1)\overline{q}_1 = \frac{1}{2}\sigma\frac{\nu}{1-\nu}(\Delta\theta)^2. \end{gathered}$$

The optimal lottery is determined by the last condition. The marginal cost of the lottery is the reduction in profits of P_1 due to assigning an inefficient production to the low cost supplier. The marginal benefit is the decrease in the rent left to this type. Note that by offering a lottery with $\sigma < 1$ P_1 increases the efficiency of his contract compared to the full separation: $\overline{q}_1 > \overline{q}^{**}$. The optimal contract of P_1 is an intermediate case between no disclosure and full disclosure. It has a pooling feature in that observing the message $\overline{\theta}$ leaves P_1 uncertain about the type of the supplier.

5 Private Communication

5.1 The Revelation Principle

The basic idea of the Revelation Principle is that a direct mechanism can replicate the distribution of outcomes of any indirect mechanism. This argument extends easily to the game with many principals under private communication. Indeed, when the message reported to a principal is not observed by the other principals, it cannot affect the beliefs of these principals about agent's type. From the point of view of the agent, the whole game can be considered as a sequence of N games with N independent reporting strategies $\sigma_1, ..., \sigma_N$. The link between these games is provided through the beliefs that a principal P_i derives from observing the state x_{i-1}^{-}. Since replacing the original mechanism with a direct one in each single principal - agent relationship results in the same probability distribution over allocations, it leads to the same state structure. Therefore, the beliefs of P_{i+1} when P_i plays a direct mechanism are the same as under the original mechanism. Each principal can thus innocuously replace its original mechanism with a direct one without affecting the distribution over allocations and the information structure. Hence, a unilateral deviation of P_i to a direct mechanism does not affect the behavior of P_{-i} and the reporting strategy of A to P_{-i}. Iterating this argument for all $i = 1, ..., N$ leads to the following result.

Theorem 2 *In the game with private communication, for any equilibrium profile $(\beta, \sigma) \in PBE(\Gamma)$ there exists an incentive feasible direct mechanism β^d which is payoff equivalent to β. Moreover, $\sigma_{i,t}^d(\theta_t) = 1$.*

One interesting implication of Theorem 2 is that, even though the agent's private information at each stage i consists of the type θ and the messages sent to principals

P_{i-1}^{-}, a principal P_i does not need the information about these messages. In fact all the information that a principal needs for assigning a contract is the agent's report on its type.

5.2 Structure of the optimal mechanism

In contrast with most static principal - agent problems in which the "no bunching" property is a sufficient condition for the optimal mechanism to be deterministic (Strausz (2004)), an optimal mechanism in the game with many principals is stochastic. The reason is that to control the information revealed to subsequent principals, the outcome of stage i must contain some noise about the agent's type. As shown in the simple example of Section 2, if P_i offers a deterministic contract, the principals P_{i+1}^{+} contract with the agent under full information about the type. When revelation of information is costly, designing a stochastic contract will be optimal. In this section we show that the equilibria of the game under private communication can be characterized by a contract with a surprisingly simple structure: At each stage i the support of a stochastic contract contains at most $|\Theta_i|$ allocations.

In equilibrium, the optimal mechanism designed by principal P_i, β_i, the mechanisms offered by the subsequent principals, β_{i+1}^{+}, the posterior beliefs, p_{i+1} must satisfy the following three conditions: (i) optimality of β_{i+1}^{+}, (ii) optimality of agent's behavioral strategy, $\sigma_{i,t}(\theta_t) = 1$ and

$$\begin{aligned} & \sum_{j=i}^{N} \int_{\Delta(F_{i-1})} u_{j,t}(x_{j-1}^{-}, x, x_{j+1}^{+}(x, p_i(x))) d\beta_{i,t}(x) \\ \geq & \sum_{j=i}^{N} \int_{\Delta(F_{i-1})} u_{j,t}(x_{j-1}^{-}, x, x_{j+1}^{+}(x, p_i(x))) d\beta_i'(x), \text{ for all } \beta_i' \in \Delta(F_i), \end{aligned}$$

and (iii) Bayes rule (4) whenever possible. Then one can define incentive feasible and incentive efficient profiles with respect to the support F_i of the principal's optimal mechanism. We say that $(x_{i-1}^{-}\beta_i, p_{i+1}, \beta_{i+1}^{+} \big| F_i)$ is incentive feasible if $(\beta_i, p_{i+1}, \beta_{i+1}^{+})$ is a Perfect Bayesian equilibrium given mechanism (β_i, F_i). It is incentive efficient if it is incentive feasible and there is no other incentive feasible mechanism $(x_{i-1}^{-}\beta_i', p_{i+1}', \beta_{i+1}^{+} \big| F_i)$ such that

$$\sum_{\Theta_i} p_{i,t}[V_{i,t}(x_{i-1}^{-}\beta_i', p_{i+1}', \beta_{i+1}^{+} \big| F_i) - V_{i,t}(x_{i-1}^{-}\beta_i, p_{i+1}, \beta_{i+1}^{+} \big| F_i)] > 0 \tag{11}$$

$$\text{and } U_{i,t}(x_{i-1}^{-}\beta_i, p_{i+1}, \beta_{i+1}^{+} \big| F_i) = U_{i,t}(x_{i-1}^{-}\beta_i', p_{i+1}', \beta_{i+1}^{+} \big| F_i). \tag{12}$$

Finally, the two profiles are payoff equivalent when condition (12) holds and (11) is satisfied as an equality.

The logic of Bester and Strausz (2001) can now be applied to analyze the support F_i of an incentive efficient mechanism $(x_{i-1}^{-}\beta_i, p_{i+1}, \beta_{i+1}^{+} \big| F_i)$. We show that an incentive efficient profile $(x_{i-1}^{-}\beta_i, p_{i+1}, \beta_{i+1}^{+} \big| F_i)$ can be replaced with a payoff equivalent $(x_{i-1}^{-}\beta_i', p_{i+1}, \beta_{i+1}^{+} \big| F_i)$ in which the support of β_i' consists of at most Θ_i allocations. We employ this result to establish that any PBE can be characterized by a payoff equivalent mechanism that uses at most $|\Theta_i|$ allocations at stage i, and for each type there is one distinct allocation that is assigned to this type with a positive probability.

Proposition 5 *In a game with private communication, for any equilibrium mechanism β_i with the support F_i there exists a payoff equivalent mechanism β_i^* with the support $F_i^* \subset F_i$ that contains at most $|\Theta_i|$ elements. When type θ_t is assigned a non-degenerate lottery, it is indifferent among the allocations that are assigned to this type with a positive probability.*

The structure of the proof is identical to the Revelation Principle under public communication. There, for any given equilibrium profile we considered a reporting strategy of the agent over a general message space. We have shown that replacing it with a strategy over the type space does not affect the payoffs and the beliefs of principals P_{i+1}^{+}. Here, instead of studying the reporting strategy of the agent (which is a deterministic truthful revelation strategy), we focus on the mechanism offered by the principal. We show that any mechanism that is part of PBE can be replaced with a payoff equivalent mechanism with finite support. Moreover, this mechanism does not affect optimal mechanisms of P_{i+1}^{+}, incentives of the agent and beliefs of subsequent principals. This is a general result that does not rely on any assumptions about the structure of payoff functions, like Spence-Mirelees condition, for example. It is useful for applications because it provides a concrete way to characterize an optimal contract of a principal as a solution of an optimization problem under incentive constraints. We apply the results of this section to characterize the optimal mechanism under private communication for the example of Section 2.

Example: As in the case of public communication, the outcome of the contract of P_1 should be an imperfect signal when the agent has high cost. This way P_1 induces P_2 to reduce the quantity offered to the high cost type, and ultimately decreases the rent paid to the low cost type in the first period. Therefore, the contract of P_1 has the following structure. Type $\overline{\theta}$ is assigned a deterministic allocation $(\bar{t}_1, \overline{q}_1)$; type $\underline{\theta}$ is assigned a lottery between $(\underline{t}_1, \underline{q}_1)$ and $(\bar{t}_1, \overline{q}_1)$ with probabilities σ and $1-\sigma$, respectively. When P_2 observes a contract $(\underline{t}_1, \underline{q}_1)$, he infers that A is $\underline{\theta}$ type. When P_2 observes the contract $(\bar{t}_1, \overline{q}_1)$, it updates the beliefs to

$$\mu = \frac{\nu(1-\sigma)}{\nu(1-\sigma)+1-\nu}.$$

In the latter case the output schedule of P_2 is the same as in the case of public communication (10).

At the first stage, P_1chooses the contracts $(\underline{t}_1, \underline{q}_1)$ and $(\bar{t}_1, \overline{q}_1)$, and the lottery σ that solve the program.

$$\begin{aligned}
\max_{(t_1,q_1,\sigma)} \quad & \nu\sigma[(1-\underline{q}_1)\underline{q}_1 - \underline{t}_1] + \nu(1-\sigma)[(1-\overline{q}_1)\overline{q}_1 - \bar{t}_1] \\
& +(1-\nu)[(1-\overline{q}_1)\overline{q}_1 - \bar{t}_1] \\
& \underline{IC}_1: \quad \underline{t}_1 - \underline{\theta}\underline{q}_1 \geq \bar{t}_1 - \underline{\theta}\overline{q}_1 + \Delta\theta\overline{q}_2(\sigma), \\
& \overline{IC}_1: \quad \bar{t}_1 - \overline{\theta}\overline{q}_1 \geq \underline{t}_1 - \overline{\theta}\underline{q}_1, \\
& \underline{PC}_1: \quad \underline{t}_1 - \underline{\theta}\underline{q}_1 \geq 0, \\
& \overline{PC}_1: \quad \bar{t}_1 - \overline{\theta}\overline{q}_1 \geq 0.
\end{aligned}$$

The striking feature of the above program is that it is identical to the one under public communication. It turns out that in this example the observability of communication does not have any additional strategic effect on the behavior of the players. Under either communication regime, P_1 controls the information that it transmitted to P_2. Under private communication, the uncertainty about the type of agent is preserved by the stochastic structure of the contract offered by P_1. Under public communication, the uncertainty is preserved by the stochastic structure of the agent's reporting strategy. In the next section we show that this feature of equilibria is a general property that holds for any sequential common agency game.

6 Equivalence

In this section we note that the expected payoff of players and the distribution of allocations in equilibrium do not depend on communication mode. The basic idea of the result is that a principal can generate the same beliefs either by inducing a stochastic reporting strategy or by offering a stochastic contract.

Consider a contract of principal P_i under private communication. It consists of at most $|\Theta_i|$ distinct allocations $\{x_{i,1}, ..., x_{i,|\Theta_i|}\}$ and a distribution $\beta_{i,t}(x_{i,j}) = \Pr(x_{i,j}|\,\theta_t)$. Denote $x_{i,t}$ the allocation that is assigned to type θ_t with a positive probability. Then an optimal contract of P_i solves

$$\max \ \sum_t p_{i,t} \sum_j \beta_{i,t}(x_{i,j}) v_{i,t}(x_{i-1}^-, x_{i,j}, x_{i+1}^+(x_{i,j}, p_{i+1})) \tag{13}$$
$$s.t. \ u_{i,t}(x_{i-1}^-, x_{i,t}, x_{i+1}^+(x_{i,j}, p_{i+1})) \geq u_{i,t}(x_{i-1}^-, x_{i,k}, x_{i+1}^+(x_{i,j}, p_{i+1})) \text{ for } \forall t, k \in \Theta_i,$$
$$if \ \beta_{i,t}(x_{i,j}) > 0 \ then \ u_{i,t}(x_{i-1}^-, x_{i,t}, x_{i+1}^+(x_{i,j}, p_{i+1})) = u_{i,t}(x_{i-1}^-, x_{i,j}, x_{i+1}^+(x_{i,j}, p_{i+1})).$$

The first constraint says that each type must weakly prefer an allocation associated with this type. The second constraint guarantees that an agent is indifferent among allocations in the support of the optimal stochastic contract.

Under public communications, a contract of P_i consists of the set of allocations $\{x_{i,1}, ..., x_{i,|\Theta_i|}\}$ and reporting strategies for the agent $\sigma_{i,t}(\theta_j) = \Pr(\theta_j|\,\theta_t)$. Denote $x_{i,j}$ an allocation assigned when the agent reports message θ_j. An optimal contract solves

$$\max \ \sum_t p_{i,t} \sum_j \sigma_{i,t}(\theta_j) v_{i,t}(x_{i-1}^-, x_{i,j}, x_{i+1}^+(x_{i,j}, p_{i+1})) \tag{14}$$
$$s.t. \ u_{i,t}(x_{i-1}^-, x_{i,t}, x_{i+1}^+(x_{i,j}, p_{i+1})) \geq u_{i,t}(x_{i-1}^-, x_{i,k}, x_{i+1}^+(x_{i,j}, p_{i+1})) \text{ for } \forall t, k \in \Theta_i,$$
$$if \ \sigma_{i,t}(\theta_j) > 0 \ then \ u_{i,t}(x_{i-1}^-, x_{i,t}, x_{i+1}^+(x_{i,j}, p_{i+1})) = u_{i,t}(x_{i-1}^-, x_{i,j}, x_{i+1}^+(x_{i,j}, p_{i+1})).$$

The first constraint guarantees that an agent weakly prefers to reveal his type. The second constraint states that an agent is indifferent among the messages sent with a positive probability.

It is straightforward to see that the two programs (13) and (14) are equivalent. Thus we obtain the following result.

Proposition 6 *The equilibrium distribution of allocations and the expected payoff of the principals and the agent are independent of communication mode.*

The implication of this result is that when the principal can offer stochastic contracts, the disclosure of information reported by a principal to the agent does not affect the outcome of the game.

7 Extensions and Discussion

Non-persistent private information. The characterization results presented in the paper can also be extended to the environment where it is common knowledge that private information of an agent changes over time. When types are independent over time, the externality of the contract of one principal on the feasible set of the subsequent principals is conducted only through the choice of allocation. In this case under both communication modes the information is revealed with probability one at each contracting stage. A more interesting situation arises when private information of the agent is correlated over time. Technically, it is tedious but straightforward to show that in this case the equilibria can still be characterized within the type space. The only new feature is that the information about correlation of types should be incorporated into the definition of the Bayes rule.

Noisy observation of contract outcomes. The results can also be extended to the case where the outcome of contracting is observed with some exogenous noise. For example, one can assume that the subsequent principals observe a signal either about the message or about the implemented allocation, depending on communication mode. If information about the signal and precision of the signal is common knowledge, again, the results presented in the paper hold, but the Bayes rule should be adjusted for this information.

One interesting question that presents an avenue for further research is whether it is possible to obtain closed form solutions for a game with an arbitrary number of periods. So far the applications in the literature on imperfect commitment or sequential common agency mostly focused on two period models. One major obstacle that prevented extension of this literature to games with many periods is that it is not straightforward to understand which constraints are binding at each contracting stage. On the other side, control of information revealed by the contract to the other principals is the major reason why principals decide to implement stochastic contracts. The example studied in the paper provides good intuition about the structure of the pooling contract. In the example, it is costly for the first principal to implement a contract under which an efficient agent's type is revealed with probability one. Thus, this type is partially pooled with an inefficient one. This result also translates in the necessary condition on the set of binding incentive constraints: To be indifferent between the two contracts, the downward incentive constraint of the efficient type must be binding. Thus, if all principals and the agent have the same ordering of types, the constraints will be binding downwards. Another observation is that this structure implies that following each implemented contract the set of types assigned a positive probability is weakly decreasing. These observation suggest that recursive methods developed by Abreu, Pearce and Stacchetti (1990), Phelan and Townsend (1991) and Marcet and Marimon (1998) could be extended to study adverse selection problem of the type we addressed in the paper.

8 Conclusion

The paper characterizes the direct mechanisms for sequential common agency games. We show that when the outcome of contracting with one principal is observed by the other principals, the equilibria of the game can be characterized within the type space. We distinguish between the cases of private and public communication. Under private communication, the message that the agent submits to one principal is not observed by the other principals. In this case the standard version of the Revelation Principle holds: The equilibrium of the game can be characterized within the class of direct mechanisms in which the agents reports truthfully its type to each principal. However, the contract is in general stochastic. We also show that the size of the support of the stochastic contract does not need to exceed the number of types that occur with a positive probability at this stage. When communication between the principal and the agent is public, that is, observed by the other principals, the equilibria of the game can also be characterized within the class of direct mechanisms. However, the requirement on the revelation of private information by the agent is weaker than in the classical case, and occurs with a positive probability. We also show that in equilibrium the two communication modes lead to the same distribution of allocations.

The characterization results that we present in the paper allow to formulate the contracting problem as a solution of an optimization problem where a principal needs to select a finite number of allocations and a lottery over this support of allocations. The results also extend to the cases when the private information of the agent is not persistent or the subsequent principals observes the outcomes of previous stages with some noise.

There are some questions that need further investigation. The most interesting one is to understand whether under some ordering assumptions on the type space, like Spence-Mirelees condition, an optimal contract can be formulated as a recursive problem.

References

[1] Abreu, D., D. Pearce and E. Stacchetti, 1990, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica,* 58: 1041-1063.

[2] Acquisti, A. and H. Varian, 2002, "Conditioning Prices on Purchase History", mimeo, University of California, Berkeley.

[3] Bester, H. and R. Strausz, 2001, "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case", *Econometrica*, 69(4): 1077-1098.

[4] Bhattacharya,S. and J. Ritter, 1983, "Innovation and Communication: Signalling with Partial Disclosure", *Review of Economic Studies,* 50: 331-346.

[5] Chen, Y. and Z.J. Zhang, 2001, "Competitive Targeted Pricing with Strategic Customers", mimeo, Stern NYU and Columbia Business School.

[6] Culter, D. and J. Gruber, 1996, "Does Public Insurance Crows Out Private Insurance", *Quarterly Journal of Economics*, 111(2):391-430.

[7] Epstein, L. and M. Peters, 1999, "A Revelation Principle for Competing Mechanisms", *Journal of Economic Theory*, 88: 119-160.

[8] Gertner, R., R. Gibbons and D. Scharfstein, 1988, "Simultaneous Signalling to the Capital and Product Markets", *RAND Journal of Economics*, 19(2): 173-190.

[9] Hall, P.,1935, "On Representatives of Subsets", *Journal of the London Mathematical Society*, 10: 26-30.

[10] Laffont, J.-J. and D. Martimort, 2002, The Theory of Incentives: The Principal-Agent Model, Princeton University Press.

[11] Lizzeri, A., 1999, "Information Revelation and Certification Intermediaries", *RAND Journal of Economics*, 30: 214-231.

[12] Padilla, A.J. and M. Pagano, 1997, "Endogenous Communication among Lenders and Entrepreneurial Incentives", *Review of Financial Studies*, 10(1): 205-236.

[13] Marcet, A. and R. Marimon, "Recursive Contracts", mimeo, Universitat Pompeu Fabra, Barcelona.

[14] Pagano, M. and T. Jappelli, 1993, "Information Sharing in Credit Markets", *Journal of Finance*, 48: 1693-1718.

[15] Pavan, A. and G. Calzolari, 2006, "Sequential Contracting with Multiple Principals", mimeo, Northwestern University.

[16] Peyrache, E. and L. Quesada,2004, "Strategic Certification", mimeo HEC Paris and Wisconsin-Madison.

[17] Phelan, C. and R. Townsend, 1991, "Computing Multiperiod Information-Constrained Optima", *Review of Economic Studies,* 58(5): 853-881.

[18] Rubin, H. and O. Wester, 1958, "A Note on Convexity in Euclidean n-Space", *Proceedings of the American Mathematical Society*, 9: 522-523.

[19] Sharpe, S., 1990, "Asymmetric Information, Bank Lending and Implicit Contracts: A Stylized Model of Customer Relationships", *Journal of Finance*, 45(4): 1069-1087.

[20] Strausz, R., 2004, "Deterministic Mechanisms are Optimal in Standard Principal - Agent Models", mimeo, Free University Berlin

[21] Taylor, C., 2002, "Private Demands and Demands for Privacy: Dynamic Pricing and the Market for Customer Information", mimeo, Duke University.

[22] Villas-Boas, J.M., 1999, "Dynamic Competition with Customer Recognition", *RAND Journal of Economics*, 30(4): 604-631.

Appendix

Public Communication

Proof of Proposition 1

To prove proposition 1, we derive the first order conditions implied by incentive efficiency in the following lemma. Then we apply the result of the lemma to establish the main result of Proposition 1.

Lemma 3 *Let* $(x_{N-2}^{-}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ *be incentive efficient. Then there exists* $\mu_{N-1} = (\mu_{N-1,1}, ..., \mu_{N-1,T_{N-1}}) \in \mathbf{R}^{T_{N-1}}$ *such that*

$$\begin{aligned} & \sum_{\theta_t \in \Theta_{N-1}} p_{N,t}(m) \int_{\Delta_{N-1}} v_{i,t}(x_{N-2}^{-}, x_{N-1}(m), x_N(x_{N-1}, p_N)) d\beta_{N-1}(m) \\ = & \sum_{\theta_t \in \Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} p_{N-1,t}(m) \quad \overline{\sigma}_{N-1} - \textit{almost everywhere.} \end{aligned}$$

Proof of Lemma 3. Let $K = \{K_1, ..., K_k, ...\}$ be a $\sigma-$ partition of M_{N-1}. Then for any $\lambda = (\lambda_1, ..., \lambda_k, ...)$ such that

$$\lambda_k \geq 0 \text{ for all } K_k \in K, \quad \sum_k \lambda_k \sigma_{N-1,t}(K_k) = 1 \text{ for all } \theta_t \in \Theta_{N-1}. \tag{15}$$

we can define a new reporting strategy σ'_{N-1} to P_{N-1} by setting

$$\sigma'_{N-1,t}(H) \equiv \sum_k \lambda_k \sigma_{N-1,t}(H \cap K_k)$$

for all $\theta_t \in \Theta_{N-1}$ and $H \in \mathcal{M}_{N-1}$. Indeed, (15) implies that $\sigma'_{N-1,t} \in S_{N-1}$ for all $\theta_t \in \Theta_{N-1}$.

Next we prove that incentive feasibility of $(x_{N-2}^{-}, \sigma_{N-1}, p_N, \beta_{N-1} | M_{N-1})$ implies that $(x_{N-2}^{-}, \sigma'_{N-1}, p_N, \beta_{N-1} | M_{N-1})$ is incentive feasible. Optimality of σ_{N-1} implies that

$$\begin{aligned} & \int_{\Delta_{N-1}} u_{N-1,t}(x_{N-2}^{-}, x_{N-1}(m), x_N(p_N(m), x_{N-1}(m))) d\beta_{N-1}(m) \qquad (16) \\ = & \; U_{N-1,t}(x_{N-2}^{-}, \sigma_{N-1}, p_N, \beta_{N-1} | M_{N-1}) \;\; \sigma_{N-1,t} - \text{almost everywhere.} \end{aligned}$$

Therefore,

$$\begin{aligned} & U_{N-1,t}(x^-_{N-2},\sigma'_{N-1},p_N,\beta_{N-1}\big|\, M_{N-1}) && (17)\\ = & \sum_k \lambda_k \int\limits_{K_k}\int\limits_{\Delta_{N-1}} u_{N-1,t}(x^-_{N-2},x_{N-1}(m),x_N(p_N(m),x_{N-1}(m)))d\beta_{N-1}(m)d\sigma_{N-1,t}(m) \\ = & U_{N-1,t}(x^-_{N-2},\sigma_{N-1},p_N,\beta_{N-1}\big|\, M_{N-1}) \end{aligned}$$

Hence, $\sigma'_{N-1,t}$ maximizes A's expected payoff and satisfies (3). To verify that $\sigma'_{N-1,t}$ is consistent with the Bayes rule, consider any $H \in \mathcal{M}_{N-1}$ and let $H_k \equiv H \cap K_k$. Then by condition (5)

$$\int\limits_{H_k} p_{N,t}(m)d\overline{\sigma}_{N-1} = p_{N-1,t}\sigma_{N-1,t}(H_k),$$

whenever $\overline{\sigma}_{N-1}(H_k) > 0$. Thus,

$$\int\limits_{H} p_{N,t}(m)d\overline{\sigma}'_{N-1} = \sum_k \int\limits_{H_k} p_{N,t}(m)\lambda_k d\overline{\sigma}_{N-1} = \sum_k p_{N-1,t}\lambda_k\sigma_{N-1,t}(H_k) = p_{N-1,t}\sigma'_{N-1,t}(H),$$

and the consistency with the Bayes rule is satisfied.

Given $(x^-_{N-2},\sigma'_{N-1},p_N,\beta_{N-1}\big|\, M_{N-1})$, P_{N-1} obtains the payoff

$$\sum_k \lambda_k \sum_{\Theta_{N-1}} p_{N-1,t} \int\limits_{K_k}\int\limits_{\Delta_{N-1}} v_{N-1,t}(x^-_{N-2},x_{N-1}(m),x_N(p_N(m),x_{N-1}(m)))d\beta_{N-1}(m)d\sigma_{N-1,t}(m).$$

Incentive efficiency of $(x^-_{N-2},\sigma_{N-1},p_N,\beta_{N-1}\big|\, M_{N-1})$ implies that $\lambda = (1,...,1,...)$ maximizes this payoff subject to (15). Therefore, there exists $\mu_{N-1} = (\mu_{N-1,1},...,\mu_{N-1,T_{N-1}}) \in \mathbf{R}^{T_{N-1}}$ such that λ' satisfies the first order condition:

$$\begin{aligned} & \sum_{\Theta_{N-1}} p_{N,t} \int\limits_{K_k}\int\limits_{\Delta(F_{N-1})} v_{N-1,t}(x^-_{N-2},x_{N-1}(m),x_N(p_N(m),x_{N-1}(m)))d\beta_{N-1}(m)d\sigma_{N-1,t}(m) \\ = & \sum_{\Theta_{N-1}} \mu_{N-1,t}\sigma_{N-1,t}(K_k) \end{aligned}$$

for all k. By the Bayes rule (5), this condition is identical to

$$\begin{aligned} & \sum_{\Theta_{N-1}} \int\limits_{K_k} p_{N,t}(m) \int\limits_{\Delta_{N-1}} v_{N-1,t}(x^-_{N-2},x_{N-1}(m),x_N(p_{N-1}(m),x_{N-1}(m)))d\beta_{N-1}(m)d\overline{\sigma}_{N-1}(m) \\ = & \sum_{\Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} \int\limits_{K_k} p_N(m)d\overline{\sigma}_{N-1}(m) && (18) \end{aligned}$$

for all $K_k \in K$. Since the above condition (18) must hold for any arbitrary $\sigma-$partition K on $\mathcal{M}_{N-1}$, we obtain the result of the Lemma. ■

Proof of the main result of Proposition 1. Note that by conditions (3) and (5)

$$[\int\limits_{\Delta_{N-1}} u_{N-1,t}(x^{-}_{N-2}, x_{N-1}(m), x_N(p_N(m), x_{N-1}(m)))d\beta_{N-1}(m) \tag{19}$$
$$-U_{N-1,t}(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1}|\, M_{N-1})]p_{N,t}(m) = 0 \quad \overline{\sigma}_{N-1} - \text{almost everywhere.}$$

Indeed, suppose there is an $H \in \mathcal{M}_{N-1}$ such that $\int_H p_{N,t}(m)d\overline{\sigma}_{N-1} > 0$ and

$$\int\limits_{\Delta_{N-1}} u_{N-1,t}(x^{-}_{N-2}, x_{N-1}(m), x_N(p_N(m), x_{N-1}(m)))d\beta_{N-1}(m) \neq U_{N-1,t}(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1}|\, M_N)$$

for all $m \in H$. Then (5) requires that $\sigma_{N-1,t}(H) > 0$. But since (3) implies (16), then $\sigma_{1,t}(H) = 0$, a contradiction.

By Lemma 3 and (19) the support of $\overline{\sigma}_{N-1}$ contains a set of messages $\overline{M} \in \mathcal{M}_{N-1}$ with $\overline{\sigma}_{N-1}(\overline{M}) = 1$ such that the following two properties are satisfied: First, there is $\mu_{N-1} \in \mathbf{R}^{T_{N-1}}$ such that

$$\sum_{\Theta_{N-1}} p_{N-1,t}(m) \int\limits_{\Delta_{N-1}} v_{N-1,t}(x^{-}_{N-2}, x_{N-1}(m), x_N(x_{N-1}(m), p_N(m)))d\beta_{N-1}(m)$$
$$= \sum_{\Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} p_{N,t}(m) \quad \text{for all } m \in \overline{M}.$$

Second,

$$[\int\limits_{\Delta_{N-1}} u_{N-1,t}(x^{-}_{N-2}, x_{N-1}(m), x_N(p_N(m), x_{N-1}(m)))d\beta_{N-1}(m) \tag{20}$$
$$- U_{N-1,t}(x^{-}_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1}|\, M_{N-1})]p_{N,t}(m) = 0 \text{ for all } m \in \overline{M}.$$

Since $\overline{\sigma}_{N-1}(\overline{M}) = 1$ implies $\sigma_{N-1,t}(\overline{M}) = 1$, it follows form the Bayes rule that

$$\int\limits_{\overline{M}} p_N(m)d\overline{\sigma}_{N-1} = p_{N-1}, \tag{21}$$

where $p_N(m) = (p_{N,t}(m))_{\theta_t \in \Theta_{N-1}}$.

Define $\overline{P} = \{p(m) \mid m \in \overline{M}\}$ and let $co(\overline{P})$ denote the convex hull of $\overline{P}$. By a theorem of Rubin and Wester (1958), (21) implies that $p_{N-2} \in co(\overline{P})$. Since $co(\overline{P})$ lies in the hyperplane $\{p \in \mathbf{R}^{T_{N-1}} \mid \Sigma_i p_i = 1\}$, it may be represented as a set in $\mathbf{R}^{T_{N-1}-1}$. Therefore, by Caratheodory's theorem, p_{N-2} can be written as a convex combination of $|M'| \leq T_{N-1}$ linearly independent vectors $p(m_1), ..., p(m_{|M'|})$ in $\overline{P}$. Thus there exists $\alpha = (\alpha_1, ..., \alpha_{|M'|})$ such that $\alpha_h \geq 0$, $\Sigma_h \alpha_h = 1$ and

$$\sum_h \alpha_h p_N(m_h) = p_{N-1}. \tag{22}$$

Consider a message set $M' = \{m_1, ..., m_{|M'|}\}$ associated with vectors $p(m_1), ..., p(m_{|M'|})$ and define a new reporting strategy for the agent by setting

$$\sigma'_{N-1,t}(H) = \sum_{m_h \in H} \frac{\alpha_t}{p_{N-1,t}} p_{N,t}(m_h). \tag{23}$$

By (22), $\sigma'_{N-1} \in S_{N-1}$. The vectors $\sigma'_{N-1}(m_h)$, $h = 1, ..., |M'|$ are linearly independent because the vectors $p(m_h)$, $h = 1, ..., |M'|$, are linearly independent.

The next step is to show that $(x^-_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1} | M'_{N-1})$ is incentive feasible. First, β_N remains an optimal strategy of P_N because $(x^-_{N-2}, \sigma'_{N-1}, p_{N-1}, \beta_{N-1} | M'_{N-1})$ and

$(x^-_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ differ only in the reporting strategy of the agent to P_{N-1}. Second, σ'_{N-1} is an optimal reporting strategy for the agent. Note that $\sigma'_{N-1,t}(M') = 1$ and $\sigma'_{N-1,t}(m_h) > 0$ only if $p_{N,t}(m_h) > 0$. Since $M' \subset \overline{M}$, together with (20) these conditions imply

$$\begin{aligned} & \sum_h \int_{\Delta_{N-1}} \sigma'_{N-1,t}(m_h) u_{N-1,t}(x^-_{N-2}, x_{N-1}(m_h), x_N(p_N(m_h), x_{N-1}(m_h))) d\beta_{N-1}(m_h) \\ = \; & U_{N-1,t}(x^-_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1}) \\ \geq \; & \int_M \int_{\Delta_{N-1}} u_{N-1,t}(x^-_{N-2}, x_{N-1}(m_h), x_N(p_N(m_h), x_{N-1}(m_h))) d\beta_{N-1}(m) d\sigma''_{N-1,t}(m) \end{aligned}$$

for all $\sigma''_{N-1,t} \in S_{N-1}$.

Third, $(x^-_{N-2}, \sigma'_{N-1}, p_{N-1}, \beta_{N-1} | M_{N-1})$ satisfies the Bayes rule because $\overline{\sigma}_{N-1}(M') = 1$ and

$$\begin{aligned} p_{N,t}(m_h) \sum_j p_{N-1,j} \sigma'_{N-1,j}(m_h) &= \sigma'_{N-1,t}(m_h) \frac{p_{N-1,t}}{\alpha_h} \sum_j \alpha_h p_{N,j}(m_h) \\ &= p_{N-1,t} \sigma'_{N-1,t}(m_h) \text{ for all } m_h \in M'. \end{aligned}$$

Finally, $(x^{-}_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ and $(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ are payoff equivalent. P_N obtains the same payoff because $(x^{-}_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ and

$(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ differ only in the agent's strategy to P_{N-1}. Suppose that $V_{N-1}(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1}) \neq V_{N-1}(x^{-}_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$. Incentive efficiency of $(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ implies

$$V_{N-1}(x^{-}_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} \middle| M_{N-1}) > V_{N-1}(x^{-}_{N-2}, \sigma'_{N-1}, p_{N-1}, \beta_{N-1} \middle| M_{N-1}).$$

Let us denote $\int\limits_{I_{N-1}} \equiv \sum\limits_{\Theta_{N-1}} \int\limits_{M_{N-1}} \int\limits_{\Delta_{N-1}}$. Therefore, by Lemma 3 and $\overline{\sigma}'(M') = \overline{\sigma}'(\overline{M}) = 1$, we obtain

$$\begin{aligned}
& \int\limits_{I_{N-1}} p_{N,t}(m) V_{N-1,t}(x^{-}_{N-2}(m), x_{N-1}(m), x_N(x_{N-1}(m), p_N(m))) d\beta_{N-1} d\overline{\sigma}_{N-1} \\
= & \sum_{\Theta_{N-1}} \int\limits_{M_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} p_{N,t}(m) d\overline{\sigma}_{N-1}(m) \qquad (24) \\
> & \sum_{\Theta_{N-1}} \int\limits_{M_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} p_{N,t}(m) d\overline{\sigma}'_{N-1}(m) \\
= & \int\limits_{I_{N-1}} p_{N,t}(m) V_{N-1,t}(x^{-}_{N-2}(m), x_{N-1}(m), x_N(x_{N-1}(m), p_N(m))) d\beta_{N-1}(m) d\overline{\sigma}(m).
\end{aligned}$$

By the Bayes rule (5),

$$\int\limits_{M_{N-1}} p_{N,t}(m) d\overline{\sigma}_{N-1}(m) = p_{N-1,t} = \int\limits_{M_{N-1}} p_{N,t}(m) d\overline{\sigma}'_{N-1}(m).$$

Thus the inequality (24) cannot hold. A contradiction. Hence $(x^{-}_{N-2}, \sigma'_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ and $(x^{-}_{N-2}, \sigma_{N-1}, p_N, \beta_{N-1} \middle| M_{N-1})$ are payoff equivalent. ■

Proof of Proposition 2. Because we can simply delete from M_{N-1} any $H \in \mathcal{M}_{N-1}$ such that $\overline{\sigma}_{N-1}(H) = 0$, Proposition 1 guarantees that there exists a mechanism $\Gamma'_{N-1} = (M'_{N-1}, \beta'_{N-1})$ with $\left|M'_{N-1}\right| \leq T_{N-1}$ and an incentive feasible $(x^{-}_{N-2}, \sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} \middle| M'_{N-1})$ which is payoff-equivalent to $(x^{-}_{N-2}, \sigma_{N-1}, p_{N-1}, \beta_{N-1} \middle| M_{N-1})$. Let $\Omega_{M'_{N-1}} = [\sigma'_{N-1}(m_h)]_{m_h \in M'_{N-1}}$ denote the $T_{N-1} \times \left|M'_{N-1}\right|$ matrix with column vectors $\sigma'_{N-1}(m_h)$, $h = 1, ..., \left|M'_{N-1}\right|$. Note that by Proposition 1 the column vectors of $\Omega_{M'}$ are linearly independent.

Define the correspondence $D : M' \rightrightarrows \Theta_{N-1}$ by $D(m_h) = \{\theta_t \mid \sigma'_{N-1,t}(m_h) > 0\}$. It follows that $D(H) = \bigcup_H \{\theta_t \mid \sigma'_{N-1,t}(m_h) > 0, m_h \in H\}$. We claim that $|D(H)| \geq |H|$

for all $H \subseteq M'_{N-1}$. Indeed, fix $H \subseteq M'_{N-1}$ and consider the $T_{N-1} \times |H|$ matrix $\Omega_H = [\sigma'_{N-1}(m_h)]_{m_h \in M'_{N-1}}$. Since the matrix $\Omega_{M'_{N-1}}$ consists of linearly independent column vectors, this also holds for Ω_H. Thus, rank$(\Omega_H) = |H|$. Note further that the matrix Ω_H has only $|D(H)|$ non-null row vectors. This implies that rank$(\Omega_H) \leq |D(H)|$. Hence it follows that $|D(H)| \geq |H|$. By the Marriage theorem, there exists a mapping $d : M'_{N-1} \to \Theta_{N-1}$ with $d(m_h) \in D(m_h)$ and the property that $d(m_h) = d(m_k)$ implies $m_h = m_k$.

We now use $d(\cdot)$ to construct a mapping $c : \Theta_{N-1} \to M'_{N-1}$ in the following way. Since the mapping $d(\cdot)$ is invertible we can set $c(d(m_h)) = m_h$ for each $m_h \in M'_{N-1}$. As $d(m_h) \in D(m_h)$, we have $\sigma'_{N-1,t}(c(\theta_t)) > 0$ for all $\theta_t \in \Theta^0_{N-1} \equiv \{d(m_h) \mid m_h \in M'_{N-1}\}$. To each $\theta_t \notin \Theta^0_{N-1}$ we can assign an arbitrary $c(\theta_t) \in M'_{N-1}$ such that $\sigma'_{N-1}(c(\theta_t)) > 0$. Such $c(\theta_t)$ exists because $\Sigma_h \sigma'_{N-1}(m_h) = 1$. Thus the mapping $c(\cdot)$ satisfies $\sigma_{N-1,t}(c(\theta_t)) > 0$ for all $\theta_t \in \Theta$. Moreover, as $\bigcup_{\Theta^0_{N-1}} c(\theta_t) = M'_{N-1}$ we have that

$$S(m_h) \equiv \{\theta_t \mid m_h = c(\theta_t)\} \neq \emptyset \text{ for all } m_h \in M'_{N-1}.$$

Now we replace the mechanism (M'_{N-1}, β'_{N-1}) and $(x^-_{N-2}, \sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} \big| M'_{N-1})$ by a direct mechanism $(\Theta_{N-1}, \beta^d_{N-1})$ and a $(x^-_{N-2}, \sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} \big| \Theta_{N-1})$ that is defined in the following way:

$$\sigma^d_{N-1,t}(\theta_j) = \frac{\sigma'_{N-1,t}(c(\theta_j))}{|S(c(\theta_j))|}, \quad p^d(\theta_j) = p'(c(\theta_j)), \quad \beta^d_{N-1}(\theta_j) = \beta'_{N-1}(c(\theta_j)). \tag{25}$$

Note that $\sigma^d_{N-1,t}(\theta_j) > 0$ for all $\theta_t \in \Theta_{N-1}$. Thus, to complete the proof it is sufficient to show that $(x^-_{N-2}, \sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} \big| \Theta_{N-1})$ is incentive feasible and payoff equivalent to $(x^-_{N-2}, \sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} \big| M'_{N-1})$.

By (25), $\sigma^d_{N-1,t}(\theta_j) = \frac{\sigma'_{N-1,t}(m_h)}{|S(m_h)|}$ for all $\theta_j \in S(m_h)$. Therefore,

$$\sum_{\theta_j \in \Theta_{N-1}} \sigma^d_{N-1,t}(\theta_j) = \sum_h \sum_{\theta_j \in S(m_h)} \frac{\sigma'_{N-1,t}(m_h)}{|S(m_h)|} = \sum_h \sigma'_{N-1,t}(m_h) = 1,$$

so that $\sigma^d_{N-1,t}$ defines a probability distribution on Θ_{N-1}. Since $\beta^d_{N-1,j} = \beta'_{N-1}(c(\theta_j))$, any allocation that an agent induces by some message $\theta_j \in \Theta_{N-1}$ under the mechanism $\Gamma^d_{N-1} = (\Theta, \beta^d_{N-1})$ it can also induce by the message $c(\theta_j) \in M'_{N-1}$ under mechanism (M'_{N-1}, β'_{N-1}). Conversely, as for each $m_h \in M_{N-1}$ there is a $\theta_t \in \Theta_{N-1}$ such that $m_h = c(\theta_t)$. Anything that A can induce under (M'_{N-1}, β'_{N-1}) it can also induce under (Θ, β^d_{N-1}). Therefore, $U_{N-1,t}(\sigma^d_{N-1}, \beta^d_{N-1}) = U_{N-1,t}(\sigma'_{N-1}, \beta'_{N-1})$ for all $\theta_t \in \Theta_{N-1}$.

Moreover, $\sigma^d_{N-1,t}(\theta_j) > 0$ if an only if $\sigma'_{N-1,t}(c(\theta_j)) > 0$. Thus, σ^d_{N-1} satisfies condition (3).

The principal's belief p^d_{N-1} is consistent with the Bayes rule (5) because

$$p^d_{N-1,t}(\theta_j) = \frac{p_{N-2,t}\sigma^d_{N-1,t}(\theta_j)}{\Sigma_k p_{N-2,k}\sigma^d_{N-1,k}(\theta_j)} = \frac{p_{N-2,t}\sigma'_{N-1,t}(c(\theta_j))}{\Sigma_k p_{N-2,k}\sigma'_{N-1,k}(c(\theta_j))} = p'_t(c(\theta_j)).$$

Thus, under the direct mechanism the optimality of the reporting strategy of A and the Bayes rule are satisfied, so it is incentive feasible.

P_N and A do not deviate from the original strategy (β^d_N, σ^d_N). First, $p^d_{N-1}(\theta_j) = p'_{N-1}(c(\theta_j))$ and $\beta^d_{N-1}(\theta_j) = \beta'_{N-1}(c(\theta_j))$. Second, under $(\sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} | \, M'_{N-1})$ the θ_t-agent induces a decision $\beta'_{N-1}(m_h)$ with probability $\sigma'_{N-1,t}(m_h)$. Under $(\sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \, \Theta_{N-1})$ it induces the same decision with the same probability, as $\Sigma_{\theta_j \in S(m_h)}\sigma^d_{N-1,t}(\theta_j) = \sigma'_{N-1,t}(m_h)$.

Since β^d_{N-1} is an optimal mechanism of P_{N-1}, and replacing the original equilibrium profile $(\beta_{N-1}, \sigma_{N-1})$ does not change the optimal choice of P_N and A at stage N, we conclude that $(\sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \, \Theta_{N-1})$ and $(\sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} | \, M'_{N-1})$ are payoff equivalent, and

$V_{N-1}(\sigma^d_{N-1}, p^d_{N-1}, \beta^d_{N-1} | \, \Theta_{N-1}) = V_{N-1}(\sigma'_{N-1}, p'_{N-1}, \beta'_{N-1} | \, M'_{N-1})$. ■

Proof of Proposition 3. In Proposition 2 we establish the result for Θ_{N-1}. Consider a type $\theta_j \in \Theta \backslash \Theta_{N-1}$. From the support of the original strategy $\sigma_{N-1,j}$ select some message $m \in M_{N-1}$. Define the direct mechanism of P_{N-1} as $\beta'_{N-1}(\theta_j) = \beta_{N-1}(m)$ and a reporting strategy of the agent

$$\sigma_{N-1,t}(\widetilde{\theta}) = \begin{cases} 1, \text{ if } \widetilde{\theta} = \theta_t, \\ 0, \text{ otherwise.} \end{cases}$$

Let the posterior belief be the same as the belief for the message m, $p_{N-1}(\theta_t) = p_{N-1}(m)$. The reporting strategy of the agent is optimal because by reporting its type it induces the same decision as by sending the message m. The posterior belief upon observing message θ_j is the same as upon observing message m. So we conclude that this extension preserve the distribution over allocations of the original profile $(\beta_{N-1}, \sigma_{N-1})$ also for out-of-equilibrium types. Consequently, when P_{N-1} offers a direct mechanism, P^-_{N-2} and A at preceding stages $1, ..., N-2$ do not deviate from the original equilibrium profile $(\beta_{N-2}, \sigma_{N-2})$. ■

Private Communication

The Revelation Principle

Proof of Theorem 2. Consider an equilibrium profile (β, σ) of $\Gamma = (\beta, M)$. For some P_i, $i \geq 2$ consider a direct mechanism

$$\beta_i^d(x_i, \theta_t) = \int_{M_i} \beta_i(x_i \mid m, x_{i-1}^-) d\sigma_{i,t}(m)$$

and a reporting strategy of A

$$\sigma_{i,t}^d(\widetilde{\theta}) = \begin{cases} 1 \text{ if } \widetilde{\theta} = \theta_t, \\ 0 \text{ otherwise,} \end{cases}$$

where $\sigma_{i,t}^d(\widetilde{\theta})$ is the probability to report type $\widetilde{\theta}$ when the true type is θ_t.

Step 1. We first verify that $(\sigma_i^d, p_i, \beta_i^d, x_{i-1}^- \big| \Theta)$ is incentive feasible. Indeed, σ_i^d is an optimal reporting strategy for the agent. Suppose the reverse. Then there exists a reporting strategy $\widetilde{\sigma}_i : \Theta \to \Theta$, with $\widetilde{\sigma}_i(\widetilde{\theta} \mid \theta) > 0$ at least for some $\widetilde{\theta} \neq \theta$, such that A obtains higher expected payoff by manipulating its report according to $\widetilde{\sigma}_i$:

$$\begin{aligned} &\sum_{\widetilde{\theta} \in \Theta} \widetilde{\sigma}_i(\widetilde{\theta} \big| \theta) \int_{F_i} u_i(x_i, x_{i-1}^-, \theta) d\beta_i^d(x_i \mid \widetilde{\theta}, x_{i-1}^-) \\ &\qquad > \int_{F_i} u_i(x_i, x_{i-1}^-, \theta) d\beta_i^d(x_i \mid \theta, x_{i-1}^-). \end{aligned}$$

This condition is equivalent to

$$\begin{aligned} &\sum_{\widetilde{\theta} \in \Theta} \widetilde{\sigma}_i(\widetilde{\theta} \big| \theta) \int_{M_i} \int_{F_i} u_i(x_i, x_{i-1}^-, \theta) d\beta_i(x_i \mid m_i, x_{i-1}^-) d\sigma(m_i \mid \widetilde{\theta}, x_{i-1}^-) \\ = &\int_{M_i} \int_{F_i(x_{i-1}^-)} u_i(x_i, x_{i-1}^-, \theta) d\beta_i(x_i \mid m_i, x_{i-1}^-) d\left[\sum_{\widetilde{\theta} \in \Theta} \widetilde{\sigma}_i(\widetilde{\theta} \mid \theta) \sigma(m_i \mid \widetilde{\theta}, x_{i-1}^-)\right] \\ &\qquad > \int_{M_i} \int_{F_i} u_i(x_i, x_{i-1}^-, \theta) d\beta_i(x_i \mid m_i, x_{i-1}^-) d\sigma(m_i \mid \theta, x_{i-1}^-), \end{aligned}$$

which contradicts optimality of σ_i. Thus, σ_i^d is optimal.

Furthermore, the profile $(\sigma_i^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta)$ induces the same posterior beliefs as the original profile

$$\begin{aligned} p_{i,t}^d(x_i^-) &= \frac{p_{i-1,t}\beta_i^d(x_i, \theta_t)}{\sum\limits_{j:p_{i-1,j}>0} p_{i-1,j}\beta_i^d(x_i, \theta_j)} \\ &= \frac{p_{i-1,t}\int_{M_i} \beta_i(x_i \mid m, x_{i-1}^-) d\sigma_{i,t}(m)}{\sum\limits_{j:p_{i-1,j}>0} p_{i-1,j}\int_{M_i} \beta_i(x_i \mid m, x_{i-1}^-) d\sigma_{i,j}(m)} = p_{i,t}(x_i^-), \end{aligned}$$

so the Bayes rule is satisfied. Therefore, $(\sigma_i^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta)$ is incentive feasible.

Step 2. The profile $(\sigma_i^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta)$ is incentive efficient. Suppose the reverse, so that there exists another direct mechanism $\widetilde{\beta}_i^d$ such that P_i obtains a strictly higher payoff under $\widetilde{\beta}_i^d$ than under β_i^d and A is indifferent between $\widetilde{\beta}_i^d$ and β_i^d, that is

$$\sum_t p_{i-1,t}[V_{i,t}(\sigma_{i,t}^d, \widetilde{p}_i, \widetilde{\beta}_i^d, x_{i-1}^- \Big| \, \Theta) - V_{i,t}(\sigma_{i,t}^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta)] > 0 \tag{26}$$
$$\text{and } U_{i,t}(\sigma_{i,t}^d, \widetilde{p}_i, \widetilde{\beta}_i^d, x_{i-1}^- \Big| \, \Theta) = U_{i,t}(\sigma_{i,t}^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta).$$

Note that $(\sigma_i^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta)$ is payoff equivalent to the original profile $(\sigma_i, p_i, \beta_i, x_{i-1}^- | \, M_i)$. The reason is that it induces the same probability measure that type θ_t is assigned an allocation x_i:

$$\beta_{i,t}^d(x_i) = \int_{M_i} \beta_i(x_i \mid m, x_{i-1}^-) d\sigma_{i,t}(m) = \beta_{i,t}(x_i),$$

and it results in the same posterior beliefs p_i. Then, the first expression in (26) can be written as

$$\begin{aligned} \sum_t p_{i-1,t} V_{i,t}(\sigma_{i,t}^d, p_i, \beta_i^d, x_{i-1}^- | \, \Theta) &= \sum_t p_{i-1,t} V_{i,t}(\sigma_{i,t}, p_i, \beta_i, x_{i-1}^- | \, M_i) \\ &< \sum_t p_{i-1,t}[V_{i,t}(\sigma_{i,t}^d, \widetilde{p}_i, \widetilde{\beta}_i^d, x_{i-1}^- \Big| \, \Theta), \end{aligned}$$

and it contradicts that (β, σ) is an equilibrium profile. Thus, the direct mechanism (β_i^d, σ_i^d) of P_i is incentive efficient.

In the next two steps we show that the original profile $(\beta_{-i}, \sigma_{-i})$ remains optimal for P_{-i} and A.

Step 3. Since β_i^d induces the same probability distribution over allocations $F_i(x_{i-1}^-)$ and the same posterior beliefs p_i, in the continuation game starting at $i+1$ it remains optimal for P_{i+1}^+ and A to follow $(\beta_{i+1}^+, \sigma_{i+1}^+)$.

Step 4. To sustain the original equilibrium $(\beta_{i-1}^{-}, \sigma_{i-1}^{-})$, it may be necessary to preserve the out-of-equilibrium messages in β_i. Note that in the direct mechanism β_i^d these messages are replicated through the allocations assigned to the out-of-equilibrium types with $p_{i-1,t} = 0$. As the direct mechanism β_i^d induces the same probability distribution over the allocations for all types, including the out-of-equilibrium ones, then the optimal contract of any P_k, $k < j$ when anticipating β_i^d is the same as when anticipating β_i. Similarly, the reporting strategy of A to P_k when anticipating a direct mechanism of P_i is the same as when anticipating the original mechanism β_i. Therefore, $(\sigma_{i-1}^{-}, \beta_{i-1}^{-})$ remains optimal for P_{i-1}^{-} and A.

Step 5. Iterating the argument for all $i = 1, ..., N$, we conclude that a profile (β^d, σ^d) of direct mechanisms and incentive compatible communication strategies is incentive efficient.

Step 6. The profile of direct mechanisms (β^d, σ^d) induces the sequence of beliefs and the probability distribution over allocations $x \in X$ which are equivalent to the that of the original profile (β, σ). Therefore, (β, σ) and (β^d, σ^d) are payoff equivalent. ■

Proof of Lemma 1. Given the reporting strategy σ_N and the mechanism β_N of the original equilibrium profile, let us define a direct mechanism and a reporting strategy of the agent as

$$\begin{aligned} \beta_N^d(\theta_t) &= \int_{M_N} \beta_N(m) d\sigma_{N,t}(m), \\ \sigma_{N,t}^d(\widetilde{\theta}) &= \begin{cases} 1 \text{ if } \widetilde{\theta} = \theta_t, \\ 0 \text{ otherwise.} \end{cases} \end{aligned}$$

For all types $\theta_t \in \Theta_N$, by the standard argument of the Revelation Principle, the profile (β_N, σ_N) is incentive feasible and payoff equivalent to (β_N, σ_N). To verify that P_{N-1}^{-} and A do not deviate form $(\beta_{N-1}^{-}, \sigma_{N-1}^{-})$, note that the out-of-equilibrium mechanisms that may be necessary to sustain (β, σ) are replicated in the direct mechanism (β_N^d, σ_N^d) through the allocations designed for out-of-equilibrium types $\theta_t \in \Theta \backslash \Theta_N$. ■

Structure of the optimal mechanism

In the following lemma we derive the first order conditions implied by incentive efficiency. Then we apply the result of the lemma to establish that an original mechanism can be replaced by a payoff equivalent mechanism that employs at most Θ_{N-1} allocations.

Lemma 4 (First order condition) *Let* $(\beta^{-}_{N-2}, p_N, \beta_{N-1}, \beta_N | F_{N-1})$ *be incentive efficient. Then there exists* $\mu_{N-1} = (\mu_{N-1,1}, ..., \mu_{N-1,T_{N-1}}) \in \mathbf{R}^{T_{N-1}}$ *such that*

$$\begin{aligned} & \sum_{\Theta_{N-1}} p_{N,t}(x_{N-1}) \int_{\Delta_{N-1}} v_{N-1,t}(x_{N-1}, x_N(x_{N-1}, p_N)) d\beta_{N-1}(x_{N-1}) \\ = & \sum_{\Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} p_{N,t}(m) \quad \overline{\beta}_{N-1} - \text{almost everywhere.} \end{aligned}$$

Proof. Let $K = \{K_1, ..., K_k, ...\}$ be a $\sigma-$ partition of F_{N-1}. Then for any $\lambda = (\lambda_1, ..., \lambda_k, ...)$ such that

$$\lambda_k \geq 0 \text{ for all } K_k \in K, \ \sum_k \lambda_k \beta_{N-1,t}(K_k) = 1 \text{ for all } \theta_t \in \Theta_{N-1}. \tag{27}$$

define a new mechanism β'_{N-1} by setting

$$\beta'_{N-1,t}(H) \equiv \sum_k \lambda_k \beta_{N-1,t}(H \cap K_k)$$

for all $\theta_t \in \Theta_{N-1}$ and $H \in X_{N-1}$. Indeed, (27) implies that $\beta'_{N-1,t} \in \Delta_{N-1}$ for all $\theta_t \in \Theta_{N-1}$.

Incentive feasibility of $(\beta^{-}_{N-2}, p_N, \beta_{N-1}, \beta_N | F_{N-1})$ implies that $(\beta^{-}_{N-2}, p_N, \beta'_{N-1}, \beta_N | F_{N-1})$ is incentive feasible. Indeed, optimality of β_{N-1} implies that

$$\begin{aligned} & \int_{\Delta_{N-1}} u_{N-1,t}(x^{-}_{N-2}, x_{N-1}, x_N(p_N, x_{N-1})) d\beta_{N-1}(x_{N-1}) \\ = & \ U_{N-1,t}(\beta^{-}_{N-2}, p_N, \beta_{N-1}, \beta_N | F_{N-1}) \quad \beta_{N-1,t} - \text{almost everywhere.} \end{aligned} \tag{28}$$

Consequently,

$$\begin{aligned} & U_{N-1,t}(\beta^{-}_{N-2}, p_N, \beta'_{N-1}, \beta_N | F_{N-1}) \\ = & \sum_k \lambda_k \int_{K_k} \int_{\Delta_{N-1}} u_{N-1,t}(x^{-}_{N-2}, x_{N-1}, x_N(p_N, x_{N-1})) d\beta_{N-1}(x_{N-1}) \\ = & \ U_{N-1,t}(\beta^{-}_{N-2}, p_N, \beta_{N-1}, \beta_N | F_{N-1}) \end{aligned} \tag{29}$$

Hence, $\beta'_{N-1,t}$ satisfies the incentive compatibility constraint of the agent. To verify that $\beta'_{N-1,t}$ is consistent with the Bayes rule, consider any $H \in F_{N-1}$ and let $H_k \equiv H \cap K_k$. Then by condition (4)

$$\int_{H_k} p_{N,t}(x) d\overline{\beta}_{N-1} = p_{N-1,t} \beta_{N-1,t}(H_k) \text{ for all } H_k \in \Delta_{N-1} \text{ with } \overline{\beta}_{N-1}(H_k) > 0.$$

Thus,

$$\int_H p_{N,t}(x)d\overline{\beta}'_{N-1} = \sum_k \int_{H_k} p_{N,t}(x)\lambda_k d\overline{\beta}_{N-1} = \sum_k p_{N-1,t}\lambda_k \beta_{N-1,t}(H_k) = p_{N-1,t}\beta'_{N-1,t}(H),$$

and the consistency with the Bayes rule is satisfied.

Given $(\beta^-_{N-2}, p_N, \beta'_{N-1}, \beta_N | F_{N-1})$, P_N obtains the payoff

$$\sum_k \lambda_k \sum_{\Theta_{N-1}} p_{N-1,t} \int_{K_k} v_{N-1,t}(x^-_{N-2}, x, x_N(p_N, x))d\beta_{N-1}(x).$$

Incentive efficiency of $(\beta^-_{N-2}, p_{N-1}, \beta_{N-1}, \beta_N | X_{N-1})$ implies that $\lambda = (1, ..., 1, ...)$ maximizes this payoff subject to (27). Therefore, there exists $\mu_{N-1} = (\mu_{N-1,1}, ..., \mu_{N-1,T_{N-1}}) \in \mathbf{R}^{T_{N-1}}$ such that λ' satisfies the first order condition:

$$\begin{aligned} & \sum_{\Theta_{N-1}} p_{N-1,t} \int_{K_k} v_{N-1,t}(x^-_{N-2}, x, x_N(p_N, x))d\beta_{N-1}(x) \\ = & \sum_{t:\theta_t \in \Theta_{N-1}} \mu_{N-1,t}\beta_{N-1,t}(K_k), \ \forall k \end{aligned}$$

Bayes rule (4) implies that this condition is equivalent to

$$\begin{aligned} & \sum_{\Theta_{N-1}} \int_{K_k} p_{N-1,t}(x_{N-1}) v_{N-1,t}(x^-_{N-2}, x_{N-1}, x_N(p_N, x_{N-1}))d\overline{\beta}_{N-1}(x_{N-1}) \\ = & \sum_{t:\theta_t \in \Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} \int_{K_k} p_N(x_{N-1})d\overline{\beta}_{N-1}(x_{N-1}) \end{aligned} \tag{30}$$

for all $K_k \in K$. Since the above condition holds for any arbitrary $\sigma-$partition K on F_{N-1}, we obtain the result of the Lemma. ■

Proposition 7 (Finite support of the mechanism) *Let $(p_{N-1}, \beta_{N-1}, \beta_N | F_{N-1})$ be incentive efficient. Then there exists an incentive feasible $(p_{N-1}, \beta'_{N-1}, \beta_N | F'_{N-1})$ and a finite set $F'_{N-1} \subset F_{N-1}$ with $|F'_{N-1}| \leq T_{N-1}$ and $\overline{\beta}'_{N-1}(F'_{N-1}) = 1$ such that*

$(p_{N-1}, \beta_{N-1}, \beta_N | F_{N-1})$ and $(p_{N-1}, \beta'_{N-1}, \beta_N | F'_{N-1})$ are payoff-equivalent. Moreover, the vectors $\beta'_{N-1}(x_h) = (\beta'_{N-1}(x_h), ..., \beta'_{N-1}(x_h))$, $x_h \in F'_{N-1}$, $h = 1, ..., |F'_{N-1}|$ are linearly independent.

Proof. Optimality of β_{N-1} and the Bayes rule imply that

$$[\int_{\Delta_{N-1}} u_{N-1,t}(x_{N-2}^{-}, x_{N-1}, x_N(p_N, x_{N-1}))d\beta_{N-1}(x_{N-1}) \tag{31}$$
$$-U_{N-1,t}(\beta_{N-2}^{-}, p_N, \beta_{N-1}, \beta_N \big| F_{N-1})]p_{N,t}(x_{N-1}) = 0 \quad \overline{\beta}_{N-1} - \text{almost everywhere.}$$

Indeed, suppose there is an $H \in F_{N-1}$ such that $\int_H p_{N,t} d\overline{\beta}_{N-1} > 0$ and

$$\int_{\Delta_{N-1}} u_{N-1,t}(x_{N-2}^{-}, x_{N-1}, x_N(p_N, x_{N-1}))d\beta_{N-1}(x_{N-1}) \neq U_{N-1,t}(\beta_{N-2}^{-}, p_N, \beta_{N-1}, \beta_N \big| F_{N-1})$$

for all $x \in H$. Then (4) requires that $\beta_{N-1,t}(H) > 0$. But optimality of β_{N-1} implies condition (28). Consequently, $\beta_{N-1,t}(H) = 0$. A contradiction.

Lemma 4 and condition (31) imply that support of $\overline{\beta}_{N-1}$ contains a set of allocations $\overline{X}_{N-1} \in X_{N-1}$ with $\overline{\beta}_{N-1}(\overline{X}_{N-1}) = 1$ such that the following two conditions are satisfied: First, there exists $\mu_{N-1} \in \mathbf{R}^{T_{N-1}}$ such that

$$\sum_{\Theta_{N-1}} \int_{K_k} p_{N,t}(x_{N-1}) v_{N-1,t}(x_{N-1}, x_N(p_N, x_{N-1}))d\overline{\beta}_{N-1}(x_{N-1})$$
$$= \sum_{\Theta_{N-1}} \frac{\mu_{N-1,t}}{p_{N-1,t}} \int_{K_k} p_N(x_{N-1})d\overline{\beta}_{N-1}(x_{N-1}) \text{ for } \forall x_{N-1} \in \overline{X}_{N-1}.$$

Second,

$$[\int_{\Delta_{N-1}} u_{N-1,t}(x_{N-1}, x_N(p_N, x_{N-1}))d\beta_{N-1}(x_{N-1}) \tag{32}$$
$$-U_{N-1,t}(\beta_{N-2}^{-}, p_N, \beta_{N-1}, \beta_N \big| F_{N-1})]p_{N-1,t}(x_{N-1}) = 0 \text{ for } \forall x_{N-1} \in \overline{X}_{N-1}.$$

Since $\overline{\beta}_{N-1}(\overline{X}_{N-1}) = 1$ implies $\beta_{N-1,t}(\overline{X}_{N-1}) = 1$, the Bayes rule implies that

$$\int_{\overline{X}_{N-1}} p_N(x_{N-1})d\overline{\beta}_{N-1} = p_{N-1}, \tag{33}$$

where $p_{N-1}(x_{N-1}) = (p_{N-1,t}(x_{N-1}))_{t=1}^{T_{N-1}}$.

Define $\overline{P} = \{p(x_{N-1}) \mid x_{N-1} \in \overline{X}_{N-1}\}$ and let $co(\overline{P})$ denote the convex hull of $\overline{P}$. By a theorem of Rubin and Wester (1958), it follows from (33) that $p_{N-1} \in co(\overline{P})$. Since $co(\overline{P})$ lies in the hyperplane $\{p \in \mathbf{R}^{T_{N-1}} \mid \Sigma_i p_i = 1\}$, it may be represented as a set in $\mathbf{R}^{T_{N-1}-1}$.

Therefore, by Caratheodory's theorem, p_{N-1} can be written as a convex combination of $|F'| \leq |\Theta_{N-1}|$ linearly independent vectors $p(x_1), ..., p(x_{|F'_{N-1}|})$ in $\overline{P}$. Thus there exists $\alpha = (\alpha_1, ..., \alpha_{|X'|})$ such that $\alpha_h \geq 0$, $\Sigma_h \alpha_h = 1$ and

$$\sum_h \alpha_h p_N(x_h) = p_{N-1}. \tag{34}$$

Consider a set of allocations $F'_{N-1} = \{x_1, ..., x_{|F'_{N-1}|}\}$ associated with vectors $p(x_1), ..., p(x_{|F'_{N-1}|})$ and define a new mechanism by setting

$$\beta'_{N-1,t}(H) = \sum_{x_h \in H} \frac{\alpha_t}{p_{N-1,t}} p_{N,t}(x_h). \tag{35}$$

By (34), $\beta'_{N-1} \in \Delta_{N-1}$. The vectors $\beta'_{N-1}(x_h)$, $h = 1, ..., |X'|$ are linearly independent because the vectors $p(x_h)$, $h = 1, ..., |X'|$, are linearly independent.

$(p_{N-1}, \beta_{N-1}, \beta_N | \, F'_{N-1})$ is incentive feasible. First, β_N remains an optimal strategy of P_N. Second, β'_{N-1} is optimal for P_{N-1}. Note that $\beta'_{N-1,t}(F'_{N-1}) = 1$, and $\beta'_{N-1,t}(x_h) > 0$ only if $p_{N,t}(x_h) > 0$. $F'_{N-1} \subset \overline{F}_{N-1}$ and (32) imply

$$\begin{aligned}
& \sum_h u_{N-1,t}(x^-_{N-2}, x_{N-1}, x_N(p_N, x_{N-1}))\beta'_{N-1}(x_h) \\
= \;& U_{N-1,t}(\beta^-_{N-2}, p_N, \beta_{N-1}, \beta_N | \, F_{N-1}) \\
\geq \;& \int_{\Delta_{N-1}} u_{N-1,t}(x^-_{N-2}, x_{N-1}, x_N(p_N, x_{N-1})) d\beta''_{N-1}(x) \\
\text{for all } \beta''_{N-1,t} \in \;& \Delta_{N-1}.
\end{aligned}$$

So, β'_{N-1} satisfies agent's incentive compatibility constraint. Third, $(\beta^-_{N-2}, p_N, \beta'_{N-1}, \beta_N | \, F_{N-1})$ satisfies the Bayes rule because $\overline{\beta}_{N-1}(F'_{N-1}) = 1$ and

$$\begin{aligned}
p_{N,t}(x_h) \sum_j p_{N-1,j} \beta'_{N-1,j}(x_h) &= \beta'_{N-1,t}(x_h) \frac{p_{N-1,t}}{\alpha_h} \sum_j \alpha_h p_{N,j}(x_h) \\
&= p_{N-1,t} \beta'_{N-1,t}(x_h) \text{ for all } x_h \in X'.
\end{aligned}$$

$(p_{N-1}, \beta_{N-1}, \beta_N | \, F'_{N-1})$ and $(p_{N-1}, \beta_{N-1}, \beta_N | \, F'_{N-1})$ are payoff equivalent. For each given outcome of β'_{N-1} P_N obtains the same payoff as under the original mechanism β'_{N-1}. Suppose that $U_{N-1}(p_{N-1}, \beta_{N-1}, \beta_N | \, F_{N-1}) \neq U_{N-1}(p_{N-1}, \beta'_{N-1}, \beta_N | \, F_{N-1})$. Because $(p_{N-1}, \beta_{N-1}, \beta_N | \, F_{N-1})$ is incentive efficient, it must be that $U_{N-1}(p_{N-1}, \beta_{N-1}, \beta_N | \, X_{N-1}) >$

$U_{N-1}(p_{N-1}, \beta'_{N-1}, \beta_N \big| X_{N-1})$. By Lemma 4 and condition $\overline{\beta}'(F'_{N-1}) = \overline{\beta}'(\overline{F}_{N-1}) = 1$, we obtain

$$\begin{aligned} & \sum_{\Theta_{N-1}} \int_{F_{N-1}} p_{N-1,t}(x) U_{N-1,t}(x_{N-1}, x_N(x_{N-1}, p_{N-1})) d\overline{\beta}_{N-1}(x) \\ = & \sum_{\Theta_{N-1}} \int_{F_{N-1}} \frac{\mu_{N-1,t}}{p_{N-2,t}} p_{N-1,t}(x) d\overline{\beta}_{N-1}(x) \\ > & \sum_{\Theta_{N-1}} \int_{F_{N-1}} \frac{\mu_{N-1,t}}{p_{N-2,t}} p_{N-1,t}(x) d\overline{\beta}'_{N-1}(x) \\ = & \sum_{\Theta_{N-1}} \int_{F_{N-1}} p_{N-1,t}(x) U_{N-1,t}(x_{N-1}, x_N(x_{N-1}, p_{N-1})) d\overline{\beta}'_{N-1}(x), \end{aligned} \tag{36}$$

Bayes rule (5) implies

$$\int_{X_{N-1}} p_{N-1,t}(x) d\overline{\beta}_{N-1}(x) = p_{N-2,t} = \int_{X_{N-1}} p_{N-1,t}(x) d\overline{\beta}'_{N-1}(x),$$

thus the inequality (36) cannot hold. A contradiction. Hence $(p_{N-1}, \beta_{N-1}, \beta_N \big| F_{N-1})$ and $(p_{N-1}, \beta'_{N-1}, \beta_N \big| F_{N-1})$ are payoff equivalent. Therefore, for any incentive efficient mechanism with an arbitrary support there exists a payoff equivalent mechanism that employs at most $|\Theta_{N-1}|$ allocations. ■