Mechanism Design with Bilateral Contracting[1]

Vianney Dequiedt[2] and David Martimort[3]

Revised version: September 25, 2008

Abstract: Suppose a principal cannot commit to a centralized grand-mechanism with all his privately informed agents but can only sign public bilateral contracts with each of them. The principal can manipulate what he learns by contracting with an agent when dealing with others. Introducing this possibility for manipulations may simplify significantly optimal mechanisms. It restores both the continuity of the principal's and the agents' payoffs and that of the optimal mechanism with respect to the information structure. Still, correlation remains useful to better extract the agents' information rent. In private values contexts, a *Revelation Principle with bilateral contracting* identifies the set of implementable allocations by means of simple *non-manipulability constraints.* Equipped with this tool, we characterize optimal non-manipulable mechanisms in various environments.

Keywords: Mechanism Design, Bilateral Contracting, Correlated Information, Robust Mechanism Design.

JEL Classification : D82.

[1]We thank for comments and discussions seminar participants at University of British Columbia, Lyon, Universidad Carlos III Madrid, Bocconi Milan, Northwestern, NICO-Northwestern, Paris-PSE, Paris-Institut Poincaré, Paris Séminaire Roy, Oxford Said Business School, Erasmus-Rotterdam, Toulouse, Marseille LAGV 2006, Lisbon ASSET 2006 Conferences and the Colin Clarke Lecture in Brisbane ESAM 2007. We also benefitted from helpful comments and discussions by Olivier Compte, Philippe Jehiel, Peter Klibanoff, Frédéric Koessler, Thomas Palfrey, Mike Peters, Salvatore Piccolo, Andy Postlewaite, Jérôme Pouyet, Jean-Charles Rochet and Jean Tirole. The usual disclaimer applies.

[2]CERDI, Université d'Auvergne, France. e-mail: vianney.dequiedt@u-clermont1.fr

[3]Toulouse School of Economics (IDEI-GREMAQ) and EHESS, France. e-mail: martimor@cict.fr

1 Introduction

Over the last thirty years, mechanism design has been the most powerful tool to understand how complex organizations and institutions are shaped. By means of the Revelation Principle,[1] this theory characterizes the set of implementable allocations in contexts where information is decentralized and privately known by agents of the organization. Once this first step of the analysis is performed, and a particular optimization criterion is specified at the outset, one can derive an optimal incentive feasible allocation and look for particular institutions (mechanisms) that may implement this outcome.

In the canonical framework for Bayesian collective choices,[2] the principal has the commitment ability to bring all agents to the "contracting table" and sign with them a grand-mechanism under the aegis of a single mediator (third-party, "machine" or Court of Law) who first collects messages from privately informed agents and second recommends them to play actions as requested by this mechanism. This paper modifies the mechanism design paradigm to take into account the principal's limited ability to rely on such centralized grand-mechanism. A mechanism is now viewed as a set of separate bilateral contracts linking the principal with each of his agents, each of those contracts being ruled by a separate mediator. What the principal learns when contracting with an agent can be manipulated by the principal himself if he finds it useful in his relationships with others. Examples of such bilateral contracting abound.

Internal organization of the firm: Consider the contracts linking the firm's management (the principal) with its workers (the agents). Each worker contracts separately with the firm, but what he communicates to the management on his own performances is by and large not observed by others. Nevertheless, the performances of a given worker can be used strategically by the management as a subjective evaluation of the latter's peers to determine their compensations.

Vertical contracting: A manufacturer deals with separate retailers who compete by selling differentiated products. Although wholesale contracts with those retailers might be publicly observable, how much intermediate good is traded with each retailer depends on his market demand. Retailers have private information both on demand and on how much they respectively trade with the manufacturer. The manufacturer may act opportunistically vis-à-vis each retailer and decide how much to sell to each retailer as a function of what he has privately learned from the others.

On-line auctions: A platform owner certainly does not bring all bidders to the "auction table." Each bidder never observes others' real bids but gets only information on the

[1]Gibbard (1973) and Green and Laffont (1977) among others.
[2]Myerson (1982 and 1991, Chapter 6.4) for instance.

winning bid that the auction designer publicly releases. Bidders must trust the platform to believe that such information is not manipulated.

Intermediated markets: Intermediaries play a significant role on most markets; financial markets being a leading example. On those markets, there is no mediator collecting aggregate supply and demand from both sides. Instead, intermediaries do so and enter into separate contractual relationships with sellers and buyers. Intermediaries might use their position at the nexus of contrats with each side of the market to act opportunistically.

Weak regulation: Regulatory and procurement environments often lack transparency (this is typically the case for developing countries). Regulators do not bring again all firms to the "contracting table" in a transparent way. Opacity opens possibilities for manipulations and those manipulations may lead to distorted regulatory decisions in case of favoritism towards some firms rather than others.

In those examples, the set of incentive feasible mechanisms must account for the fact that the principal might manipulate what is learned from his relationship with a given agent when contracting with others. Taking into account such manipulations may simplify significantly the characterization of incentive feasible allocations. It shall also allow us to reach more palatable conclusions on the design of contracts relative to those obtained when assuming that only a centralized grand-mechanism can be enforced.

Our main results are as follows.

Characterizing non-manipulable allocations: In a private values setting (i.e., when the agents' private information does not enter directly into the principal's objectives), a *Revelation Principle with bilateral contracting* characterizes the set of incentive feasible allocations under our bilateral contracting game. For a given implementation concept characterizing the agents' behavior (Bayesian-Nash or dominant strategy) there is no loss of generality in restricting the analysis to *non-manipulable* direct revelation mechanisms.

Non-manipulability constraints affect contract design. To see how, consider an organization with one principal and two agents, each running a different project on the principal's behalf. Agents have private information on their costs. Assume also that there is no productive externality between projects (technologies are separable) but informational externalities do exist (costs are correlated). Whereas private information is costless for the principal if he can design a grand-mechanism in such contexts,[3] this is no longer the case when only bilateral contracts are feasible. Such contracts are highly sensitive to the principal's manipulations: The principal can always claim that the performances of agent A_1 that he has privately observed conflicts with those of agent A_2 and punish the latter accordingly. To avoid such manipulations, the compensation of a

[3]Crémer and McLean (1985, 1988), McAfee and Reny (1992), Riordan and Sappington (1988), Johnson, Pratt and Zeckhauser (1990), d'Aspremont, Crémer and Gerard-Varet (1990), Matsushima (1991).

given agent must be less sensitive to what the principal has learned from others. With separable projects, non-manipulability is obtained with simple *sell-out contracts* that give to the principal a payoff independent of the agent's output.

Optimal mechanisms and rent/efficiency trade-off: Insisting on non-manipulability restores a genuine trade-off between rent extraction and efficiency even when the agents' types are correlated. Although the scope for yardstick competition is now more limited than with a centralized grand-mechanism, correlated information is still useful when writing bilateral contracts. Correlation makes it easier to extract the agents' information rents without nullifying those rents.

With separable projects, the optimal mechanism trades off the marginal efficiency of the agents' productions with virtual marginal costs that generalize those found in independent types environments. Allowing for more general production externalities between the agents' activities, we characterize non-manipulable contracts and show how they generalize the "sell-out" contracts found with separable projects. The optimal non-manipulable mechanism solves a system of partial differential equations that generalizes the standard second-best optimality conditions found in models with independent types and replicates those conditions in the limit of no correlation.

Continuity of payoffs and mechanisms: When a grand-mechanism can be used, privately informed agents get no rent if their types are correlated whereas they do so if types are independent. This lack of continuity of the optimal mechanism with respect to the information structure is a weakness in view of the so-called *"Wilson Doctrine"* which points out that mechanisms should be robust to small perturbations of the modelling. Taking into account non-manipulability constraints restores such continuity. Not only the principal's and the agents' payoffs vary continuously with the correlation but also the optimal mechanism keeps the same structure. To illustrate, sell-out contracts are optimal for separable projects both at zero and at a positive correlation.

Simple bilateral contracting: A *simple bilateral contract* uses only the corresponding agent's information and not what the principal might learn from others. Such mechanisms are non-manipulable. In Bayesian environments with separable projects, such simple bilateral contracts are dominated by non-manipulable bilateral mechanisms which use that information. By contrast, if dominant strategy and ex post participation constraints are imposed or if collusion between agents matters, simple bilateral contracts are optimal.

Section 2 presents our general model. Section 3 develops a simple example highlighting the fact that the principal's manipulations might constrain significantly mechanisms. Section 4 proves the Revelation Principle with bilateral contracting. Equipped with this tool, we characterize optimal mechanisms for separable projects (Section 5), and general production externalities (Section 6.1). A particular attention is given to production in teams (Section 6.2.1) and auctions (Section 6.2.2). For these two cases, the optimal non-manipulable mechanisms are derived in discrete type models. Section 7 analyzes various extensions allowing for dominant strategy (Section 7.1), collusion between agents (Section 7.2) and secret contracts (Section 7.3). Section 8 reviews the relevant literature. Section 9 proposes alleys for further research. All proofs are in an Appendix.

2 The Model

• **Preferences and Information:** We consider an organization with a principal (P) and n agents (A_i for $i = 1, ..., n$). Agent A_i produces a good in quantity q_i on the principal's behalf. Let $q = (q_1, ..., q_n)$ (resp. $t = (t_1, ..., t_n)$) denote the vector of goods (resp. transfers) which belongs to a set $\mathcal{Q} = \Pi_{i=1}^n \mathcal{Q}_i$ where $\mathcal{Q}_i \subset \mathbb{R}_+$ is compact and convex (resp. $\mathcal{T} = \Pi_{i=1}^n \mathcal{T}_i \subset \mathbb{R}^n$). By a standard convention, A_{-i} denotes the set of all agents except A_i and similar notations are used below for other variables.

The principal and his agents have quasi-linear utility functions defined respectively as:

$$V(q,t) = \tilde{S}(q) - \sum_{i=1}^{n} t_i \quad \text{and} \quad U_i(q,t) = t_i - \theta_i q_i.$$

The efficiency parameter θ_i is A_i's private information. It belongs to the set $\Theta = [\underline{\theta}, \bar{\theta}]$.[4] A vector of types is denoted $\theta = (\theta_1, ..., \theta_n) \in \Theta^n$. Types are jointly drawn from the common knowledge non-negative, bounded and atomless density function $\tilde{f}(\theta)$ whose support is Θ^n. Assuming, for simplicity, symmetric distributions,[5] we will denote marginal density, corresponding cumulative distribution and conditional density respectively as:

$$f(\theta_i) = \int_{\Theta^{n-1}} \tilde{f}(\theta_i, \theta_{-i}) d\theta_{-i}, \quad F(\theta_i) = \int_{\underline{\theta}}^{\theta_i} f(\theta_i) d\theta_i \quad \text{and} \quad \tilde{f}(\theta_{-i}|\theta_i) = \frac{\tilde{f}(\theta_i, \theta_{-i})}{f(\theta_i)}.$$

The principal's surplus function $\tilde{S}(\cdot)$ is increasing in each of its arguments q_i and concave in q. This formulation encompasses three cases of interest which will receive more attention in the sequel, specifically in organizations involving only two agents:

• *Separable projects:* $\tilde{S}(\cdot)$ is separable in both q_1 and q_2 and thus can be written as $\tilde{S}(q_1, q_2) = S(q_1) + S(q_2)$ for some function $S(\cdot)$ that is increasing and concave with the Inada condition $S'(0) = +\infty$, $S(0) = 0$ and $S'(+\infty) = 0$.

• *Perfect substitutability:* $\tilde{S}(\cdot)$ depends on the total production $q_1 + q_2$ only: $\tilde{S}(q_1, q_2) = S(q_1 + q_2)$ for some increasing and concave $S(\cdot)$ which still satisfies the above conditions.

[4] Sections 6.2.1 and 6.2.2 deal with the case of a discrete distribution.

[5] All our results could be straightforwardly adapted to asymmetric distributions.

• *Perfect complementarity:* $\tilde{S}(\cdot)$ can then be written as $\tilde{S}(q_1, q_2) = S(\min(q_1, q_2))$ where $S(\cdot)$ satisfies again the above conditions.

With separable projects, the only externality between agents is informational and comes from the possible correlation of their costs. Perfect substitutability arises instead in the context of auctions. Perfect complementarity occurs in team productions.[6]

Importantly, we consider a *private values* environment, i.e., the agents' private information does not enter directly into the principal's objective. The consequences of this assumption on some of our results will be discussed later.

• **Mechanisms:** The principal is unable to commit to a grand-mechanism bringing all agents to the contracting table and signs instead separate bilateral contracts with each of them. A bilateral contract with a given agent can, in full generality, use the principal's report on any information that he may get by contracting with others. Manipulations by the principal can arise because what a given agent communicates to the principal is not observed by others who will only learn about that information from the principal himself.

In our environment, a mechanism is a pair $(g(\cdot), \mathcal{M})$ where $g(\cdot)$ is an outcome function and $\mathcal{M} = \Pi_{i=1}^{n} \mathcal{M}_i$ the product space of the respective communication spaces $\mathcal{M}_i$ available to agent A_i to communicate with the principal. To capture the fact that the principal enters into a separate contractual relationship (sometimes referred to as a *sub-mechanism* in the sequel) with each of his agents, the outcome function $g(\cdot)$ is itself decomposed into a vector of n outcome functions $g(\cdot) = (g_1(\cdot), ..., g_n(\cdot))$. Each sub-mechanism $g_i(\cdot)$ maps $\mathcal{M} = \mathcal{M}_i \times \mathcal{M}_{-i}$ into the set $\Delta(\mathcal{Q}_i \times \mathcal{T}_i)$ of (possibly random) allocations for agent A_i. When playing the sub-mechanism $(g_i(\cdot), \mathcal{M})$, A_i communicates some message m_i to a mediator $\mathfrak{M}_i$. Such communication is observed by P. Then, the principal makes also a report $\hat{m}_{-i}$ to $\mathfrak{M}_i$ on whatever information he may have learned in observing the reports made by agents A_{-i} in the other sub-mechanisms $(g_{-i}(\cdot), \mathcal{M})$. Finally, the requested transfer $t_i(m_i, \hat{m}_{-i})$ and output $q_i(m_i, \hat{m}_{-i})$ for agent A_i are implemented.[7] Because of production or informational externalities, A_i's allocation should depend in full generality on the report $\hat{m}_{-i}$ made by the principal on the messages m_{-i} sent by other agents A_{-i} and observed by that principal.[8]

Remark 1: Standard mechanism design assumes that a unique mediator $\mathfrak{M}$ keeps one party's message secret from the other when running a centralized grand-mechanism. In sharp contrast, we suppose that agent A_i only observes the messages m_i he sends to the

[6]By simply changing variables, perfect complementarity is relevant for public good problems.

[7]When allocations are random, $q_i(m_i, \hat{m}_{-i})$ and $t_i(m_i, \hat{m}_{-i})$ should be accordingly viewed as distributions of outputs and transfers. In this case and with obvious notations, payoffs should be understood as expectations over those distributions.

[8]Because of bilateral contracting, the principal may a priori send different messages concerning what he learned from a third agent towards two different sub-mechanisms.

mediator $\mathfrak{M}_i$ ruling the sub-mechanism $g_i(\cdot)$ whereas P observes the whole array of messages $m = (m_1, ..., m_n)$ before communicating back to mediators in each sub-mechanism. This assumption is justified whenever mediators are not machine but may have their own financial objectives and may collude with the principal to share information they have gathered from the agents. Alternatively, this amounts to assuming that the only possible mediator available is the principal himself.[9] Under both interpretations, mediators make whatever information they learned from each agent available to the principal.[10] Finally, we also assume that agent A_i and the mediator $\mathfrak{M}_i$ (when distinct from the principal) do not observe the report m_{-i} made by A_{-i} into $g_{-i}(\cdot)$ and thus cannot compare this announcement with whatever P claims having learned from A_{-i}. That there is no single benevolent mediator having access to all the agents' reports leaves to the sole principal a strategic role at a nexus of all communication channels. Otherwise such a mediator could restore the possibilities of writing a grand-mechanism.[11] The constraint on observability under bilateral contracting could also be alleviated if each agent A_i could observe ex post other agents' trades (q_{-i}, t_{-i}) and infer from this (at least partly) whether the claim $\hat{m}_{-i}$ that the principal has made fits with the reports m_{-i} made by A_{-i}. To avoid such inferences we assume that those trades (q_{-i}, t_{-i}) remain nonobservable by A_i.[12] ■

Remark 2: A *simple bilateral contract* corresponds to an outcome function $g_i^*(\cdot)$ which only maps agent A_i's communication space $\mathcal{M}_i$ into $\Delta(\mathcal{Q}_i \times \mathcal{T}_i)$. With such contract, there is no scope for using what the principal has learned from observing A_{-i}'s messages to improve A_i's allocation. Those contracts are clearly non-manipulable. ■

Remark 3: For minimal departure from standard mechanism design, we assume that the mechanism $(g(\cdot), \mathcal{M})$ is publicly observable by all agents. Assuming private submechanisms, i.e., that A_i does not observe the submechanisms $(g_{-i}(\cdot), \mathcal{M})$, introduces another

[9] That mediators are not necessarily available is standard in the literature on limited commitment in dynamic contractual relationships. See for instance Laffont and Tirole (1993, Chapters 9 and 10).

[10] Suppose alternatively that mediators design private communication channels with each agent and keep their reports secret. Then there would be no scope for the principal communicating back in each sub-mechanism because he would not have observed the agents' reports in other sub-mechanisms. Only simple bilateral contracts (see Remark 2 below for their definition) are then available. Clearly, those mechanisms are most of the time suboptimal (see the formal analysis below) since they do not allow the principal to enjoy most of the benefits of having informational or production externalities between agents. In other words, if the principal could choose ex ante between having separate mediators running sub-mechanisms entertaining private communication only or having sub-mechanisms making agents' reports available to him, he would choose the latter mode of bilateral contracting.

[11] This "incomplete contracting" assumption is standard in the literature on vertical contracting and common agency (Martimort, 2007).

[12] One could think of less extreme situations where each agent may get a signal correlated with what the others are privately reporting to the principal. Of course, if this signal is public and verifiable, contingent mechanisms could be written to help circumvent the privacy problem. However, if this signal is only privately observed and can be manipulated, such contingent mechanisms lose again their force. Section 6.2.1 goes also in the direction of analyzing what happens when agents have access to some ex post information on other's reports.

dimension of private information in our model: the principal being now privately informed on contractual deals made with them. We investigate this issue in Section 7.3 below. ■

Remark 4: Because, we want to focus on the most extreme case of limited commitment where the principal can only make output recommendations once he has learned all the agents' reports, we assume that bilateral contracting with each of them is simultaneous. The principal is not endowed with the commitment ability to first contract and trade with some agents after having learned information from others. ■

• **Timing:** The contracting game unfolds as follows. First, agents privately learn their respective efficiency parameters. Second, the principal offers a mechanism $(g(\cdot), \mathcal{M})$ to the agents. Third, agents simultaneously accept or refuse their respective sub-mechanisms $(g_i(\cdot), \mathcal{M})$. If agent A_i refuses, he gets a payoff normalized to zero. Fourth, agents simultaneously send messages m_i in $g_i(\cdot)$. Fifth, the principal reports in his contractual relationship with A_i a message $\hat{m}_{-i}$ on what he has learned from contracting with A_{-i}. Finally, agent A_i's outputs and transfers are implemented according to the messages $(m_i, \hat{m}_{-i})$ and the outcome function $g_i(\cdot)$.

The equilibrium concept is perfect Bayesian equilibrium (thereafter PBE).[13]

• **Benchmark:** With correlated types, the first-best outcome can be either obtained (with discrete types) or arbitrarily approached (with a continuum of types) when a grand-mechanism is possible. In sharp contrast with what intuition commends, there is no trade-off between efficiency and rent extraction in such correlated environments. In the case of separable projects, the (symmetric) first-best output requested from each agent trades off the marginal benefit of production against its marginal cost:

$$S'(q^{FB}(\theta_i)) = \theta_i, \quad i = 1, ..., n. \tag{1}$$

Instead with independent types, agents obtain costly information rents and a genuine trade-off between efficiency and rent extraction arises. The marginal benefit of production must balance its *virtual marginal cost.* With separable projects again, the (symmetric) second-best output is given by the so-called *Baron-Myerson* outcome[14] for each agent:

$$S'(q^{BM}(\theta_i)) = \theta_i + \frac{F(\theta_i)}{f(\theta_i)}, \quad i = 1, ..., n.^{15} \tag{2}$$

Remark 5: This Baron-Myerson outcome is also obtained when the principal uses a simple bilateral contract with each agent even when types are correlated. The discrepancy

[13] Except in Section 7.1 where dominant strategy implementation characterizes the agents' behavior.

[14] Baron and Myerson (1982).

[15] Provided that the *Monotone Hazard Rate Property* holds, namely $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) > 0 \quad \forall \theta \in \Theta$, $q^{BM}(\theta_i)$ is indeed the optimal output. Otherwise, bunching may arise (Guesnerie and Laffont 1984, and Laffont and Martimort 2002, Chapter 3).

between (1) and (2) measures then the efficiency loss incurred when a grand-mechanism is replaced by n simple bilateral contracts. ■

3 A Simple Example

To get some preliminary insights on our general analysis, let us consider a simple example[16] where the principal's ability to manipulate information significantly undermines optimal contracting. A buyer (the principal) wants to procure one unit of a good from a single seller (the agent). The principal's valuation for this unit is S. The seller's cost θ takes values in $\Theta = \{\underline{\theta}, \bar{\theta}\}$ (where $\Delta\theta = \bar{\theta} - \underline{\theta} > 0$) with respective probabilities ν and $1-\nu$. Trade is efficient with both types of seller under complete information whereas it is no longer always efficient under asymmetric information when

$$\bar{\theta} + \frac{\nu}{1-\nu}\Delta\theta > S > \bar{\theta}. \tag{3}$$

The buyer makes an optimal take-it-or-leave-it offer to the seller at a price $\underline{\theta}$. Only an efficient seller accepts such offer.

Suppose now that the buyer learns a signal $\sigma \in \{\underline{\theta}, \bar{\theta}\}$ on the agent's type ex post, i.e., once the agent has already reported his cost parameter. This signal is positively correlated with the agent's type:

$$\text{proba}\{\sigma = \underline{\theta}|\underline{\theta}\} = \text{proba}\{\sigma = \bar{\theta}|\bar{\theta}\} = \rho > \frac{1}{2} > \text{proba}\{\sigma = \underline{\theta}|\bar{\theta}\} = \text{proba}\{\sigma = \bar{\theta}|\underline{\theta}\} = 1-\rho.$$

Let assume that σ is publicly verifiable. The price $t(\theta,\sigma)$ paid by the buyer for one unit of the good depends in full generality both of the seller's report on his cost and on the realized value of the signal. Looking for prices that implement the first-best trades, incentive compatibility for both types of seller requires now respectively:

$$\begin{aligned}
\rho t(\underline{\theta},\underline{\theta}) + (1-\rho)t(\underline{\theta},\bar{\theta}) &\geq \rho t(\bar{\theta},\underline{\theta}) + (1-\rho)t(\bar{\theta},\bar{\theta}),\\
(1-\rho)t(\bar{\theta},\underline{\theta}) + \rho t(\bar{\theta},\bar{\theta}) &\geq (1-\rho)t(\underline{\theta},\underline{\theta}) + \rho t(\underline{\theta},\bar{\theta}).
\end{aligned}$$

Normalizing at zero the seller's outside opportunities, the respective participation constraints of both types (assuming again efficient trades) can be written as:

$$\begin{aligned}
\rho t(\underline{\theta},\underline{\theta}) + (1-\rho)t(\underline{\theta},\bar{\theta}) - \underline{\theta} &\geq 0,\\
(1-\rho)t(\bar{\theta},\underline{\theta}) + \rho t(\bar{\theta},\bar{\theta}) - \bar{\theta} &\geq 0.
\end{aligned}$$

[16]This example is in the spirit of the principal/agent model in Riordan and Sappington (1988) where a signal correlated with the agent's type is produced exogenously and not by another agent's report as in our general analysis below.

The buyer can extract the seller's surplus and implement the first-best trades by properly designing transfers. Among many other possibilities given by Farkas' Lemma, the following prices suffice:[17]

$$t(\underline{\theta},\underline{\theta}) = \frac{\rho}{2\rho-1}\underline{\theta} > 0, t(\underline{\theta},\bar{\theta}) = -\frac{1-\rho}{2\rho-1}\underline{\theta} < 0, t(\bar{\theta},\underline{\theta}) = -\frac{1-\rho}{2\rho-1}\bar{\theta} < 0, t(\bar{\theta},\bar{\theta}) = \frac{\rho}{2\rho-1}\bar{\theta} > 0.$$

This mechanism punishes the seller whenever his report conflicts with the public signal. Otherwise the seller is rewarded and paid more than his marginal cost.

Consider now the case where the principal privately observes σ. The mechanism above becomes manipulable. Once the seller has already reported his type, the buyer may want to claim that he receives conflicting evidence on the agent's report to pocket the corresponding punishment instead of giving a reward. To avoid those manipulations and implement the efficient allocation, prices must be independent of the realized signal:

$$t(\theta,\underline{\theta}) = t(\theta,\bar{\theta}) \qquad \forall \theta \in \{\underline{\theta},\bar{\theta}\}.$$

This *non-manipulability constraint* brings us back to the traditional screening model without ex post information. Given (3), the principal is better off if trade only occurs with an efficient seller.

This simple example illustrates the consequences of having the principal manipulate information which, if otherwise publicly verifiable, would be used for screening purposes. In the sequel, informative signals are no longer exogenously produced but are learned from contracting with other agents who, in equilibrium, report truthfully their types. Second, non-manipulability constraints are derived rather than assumed. Moreover, and in contrast with the above example where output was fixed (one unit of the good had to be produced irrespectively of the observed/reported signal σ), non-manipulability may require distortions on both outputs and transfers.

4 Revelation Principle and Non-Manipulability in Private Values Environments

We provide now a full characterization of the set of allocations that can be achieved as PBEs of the overall contracting game where the principal offers any possible mechanism $(g(\cdot),\mathcal{M})$ under bilateral contracting in this private values context.

For any agents' reporting strategy $m^*(\cdot) = (m_1^*(\cdot),...,m_n^*(\cdot))$, let $\sup m_i^*(\cdot)$ denote the support of $m_i^*(\cdot)$, i.e., the set of messages m_i that are sent with positive probability by

[17] Those prices are obtained when all incentive and participation constraints are binding just to illustrate our purpose.

A_i given $m_i^*(\cdot)$. For a given mechanism with bilateral contracting $(g(\cdot), \mathcal{M})$ accepted by all types of agents, the continuation equilibria induced by such mechanism at the communication stage are described in the next lemma:

Lemma 1 *For any arbitrary mechanism $(g(\cdot), \mathcal{M})$, a continuation equilibrium is a pair $\{m^*(\cdot), \hat{m}^*(\cdot)\}$ such that:*

- *The agents' strategy vector $m^*(\theta) = (m_1^*(\theta_1), ..., m_n^*(\theta_n))$ from Θ^n into $\mathcal{M} = \Pi_{i=1}^n \mathcal{M}_i$ forms a Bayesian equilibrium given the principal's optimal manipulation $\hat{m}^*(\cdot)$*

$$m_i^*(\theta_i) \in \arg\max_{m_i \in \mathcal{M}_i} E_{\theta_{-i}} \left(t_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i}^*(\theta_{-i}))) - \theta_i q_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i}^*(\theta_{-i})))|\theta_i \right); \tag{4}$$

- *The principal's optimal manipulation $\hat{m}^*(\cdot) = (\hat{m}_{-1}^*(\cdot), ..., \hat{m}_{-n}^*(\cdot))$ from $\mathcal{M}$ on $\Pi_{i=1}^n \mathcal{M}_{-i}$ satisfies $\forall m = (m_1, ..., m_n) \in \mathcal{M}$*

$$(\hat{m}_{-1}^*(m), ..., \hat{m}_{-n}^*(m))$$
$$\in \arg\max_{(\hat{m}_{-1}, ..., \hat{m}_{-n}) \in \Pi_{i=1}^n \mathcal{M}_{-i}} \tilde{S}(q_1(m_1, \hat{m}_{-1}), ..., q_n(m_n, \hat{m}_{-n})) - \sum_{i=1}^n t_i(m_i, \hat{m}_{-i}). \tag{5}$$

Given a mechanism $(g(\cdot), \mathcal{M})$, a continuation equilibrium induces an allocation $a(\theta) = g(m^*(\theta), \hat{m}^*(m^*(\theta)))$ which maps Θ^n into $\Delta(\mathcal{Q} \times \mathcal{T})$. In this private values context, updated beliefs held by the principal following the agents' reports $m^*(\theta)$ do not influence his optimal manipulation. This can be seen more precisely on equation (5) which is written ex post, i.e., for each realization of the agents' reports.[18]

The following definitions are useful:

Definition 1 *A mechanism $(g(\cdot), \mathcal{M})$ is non-manipulable if and only if $\hat{m}_{-i}^*(m) = m_{-i}$, for all $m \in \sup m^*(\cdot)$ and i at a continuation equilibrium.*[19]

Definition 2 *A direct mechanism $(\bar{g}(\cdot), \Theta^n)$ is truthful if and only if $m^*(\theta) = \theta$, for all $\theta \in \Theta$ at a continuation equilibrium.*

Proposition 1 ***The Revelation Principle with Bilateral Contracting.*** *In a private values context, any allocation $a(\cdot)$ achieved at a continuation equilibrium of any arbitrary mechanism $(g(\cdot), \mathcal{M})$ with bilateral contracting can also be implemented through a truthful and non-manipulable direct mechanism $(\bar{g}(\cdot), \Theta^n)$.*

[18] That aspect of the modeling simplifies significantly the analysis by avoiding any signalling issue when agents communicate their types.

[19] Note that our concept of non-manipulability is weak and that we do not impose the more stringent requirement that the mechanism is non-manipulable at all continuation equilibria.

With such direct revelation mechanisms, the agents' Bayesian incentive compatibility constraints are written as usual:

$$\underset{\theta_{-i}}{E}\left(t_i(\theta_i,\theta_{-i})-\theta_i q_i(\theta_i,\theta_{-i})|\theta_i\right) \geq \underset{\theta_{-i}}{E}\left(t_i(\hat{\theta}_i,\theta_{-i})-\theta_i q_i(\hat{\theta}_i,\theta_{-i})|\theta_i\right) \quad \forall(\theta_i,\hat{\theta}_i)\in\Theta^2. \quad (6)$$

The following *non-manipulability* constraints stipulate that the principal does not misrepresent what he has learned from others' reports in his relationship with any agent:

$$\tilde{S}(q(\theta))-\sum_{i=1}^{n} t_i(\theta) \geq \tilde{S}(q_1(\theta_1,\hat{\theta}_{-1}),...,q_n(\theta_n,\hat{\theta}_{-n}))-\sum_{i=1}^{n} t_i(\theta_i,\hat{\theta}_{-i}), \quad \forall(\theta,\hat{\theta}_{-1},...,\hat{\theta}_{-n}). \quad (7)$$

Remark 6: Taxation Principle. We could have started with nonlinear prices as primitives of our analysis, i.e., sub-mechanisms $g_i(\cdot)$ mapping Θ into $\mathfrak{T}_i=\{T_i(\cdot):\mathcal{Q}_i\to\mathcal{T}_i\}$. With this approach, everything happens as if agent A_i picks first one such nonlinear price within the family $\{T_i(\cdot,\hat{\theta}_i)\}_{\hat{\theta}_i\in\Theta}$, and the principal optimally chooses afterwards the particular output q_i and the corresponding transfers $T_i(q_i,\hat{\theta}_i)$ that this agent receives as a function of the selected nonlinear prices. In other words, the constraints imposed by the manipulability of the mechanisms are akin to assuming that the principal can commit to offer menus of nonlinear prices $\{T_i(\cdot,\hat{\theta}_i)\}_{\hat{\theta}_i\in\Theta}$ to his agents in the first place *but* cannot commit to a particular output schedule $\{q(\hat{\theta})\}_{\hat{\theta}\in\Theta^n}$ beforehand.[20]

Starting from a non-manipulable direct revelation mechanism, we may define the nonlinear price $T_i(q_i,\theta_i)$ as $T_i(q_i,\theta_i)=t_i(\theta_i,\theta_{-i})$ for $q_i=q_i(\theta_i,\theta_{-i})$ and the definition is unambiguous since any θ_{-i} such that $q_i(\theta_i,\theta_{-i})=q_i$ corresponds to the same transfer $t_i(\theta_i,\theta_{-i})=t_i$ otherwise, from (10), manipulations would arise. Written in terms of those nonlinear prices, the non-manipulability constraints (7) become:

$$q(\theta)\in\arg\max_q \tilde{S}(q)-\sum_{i=1}^{n} T_i(q_i,\theta_i). \quad (8)$$

■

In the sequel, we analyze the impact of the non-manipulability constraints (7) on optimal mechanisms in different contexts involving two agents. Those constraints become:

$$\tilde{S}(q(\theta))-\sum_{i=1}^{2} t_i(\theta) \geq \tilde{S}(q_1(\theta_1,\hat{\theta}_2),q_2(\hat{\theta}_1,\theta_2))-\sum_{i=1}^{2} t_i(\theta_i,\hat{\theta}_{-i}), \quad \forall(\theta,\hat{\theta}_1,\hat{\theta}_2). \quad (9)$$

[20] This assumption on no commitment to any output schedule also rules out de facto any sequential timing of the game where the principal would commit to such output schedules for a subset of agents after having learned the messages of the remaining subset only.

5 Separable Projects

Let us start with the simplest setting with only two agents working each on a different project without any production externality between those projects. The principal's gross surplus function is separable and writes as $\tilde{S}(q_1, q_2) = \sum_{i=1}^{2} S(q_i)$. This case provides a useful benchmark to understand how non-manipulability constraints affect contract design when only informational externality between agents matter.

The non-manipulability constraints (9) imply the existence of functions $H_i(\theta_i)$ such that:

$$S(q_i(\theta_i, \theta_{-i})) - t_i(\theta_i, \theta_{-i}) = H_i(\theta_i) \qquad (10)$$

Equation (10) shows that each agent is made residual claimant for the part of the principal's objective function which is directly related to his own output.

The direct mechanism above can be transformed into a nonlinear price $T_i(q_i, \theta_i) = t_i(\theta_i, \theta_{-i})$ for $q_i(\theta_i, \theta_{-i}) = q_i$. Such nonlinear price corresponds to a *sell-out contract*:

$$T_i(q_i, \theta_i) = S(q_i) - H_i(\theta_i). \qquad (11)$$

With such scheme, agent A_i pays upfront a fixed-fee $H_i(\theta_i)$ to produce on the principal's behalf. Then, the principal once informed on all agents' reports choose an output and agent A_i enjoys all the benefit $S(q_i)$ on the project he is running. The principal's payoff in his relationship with A_i is $H_i(\theta_i)$ which does not depend on the agent's output. These fixed-fees are chosen so that the mechanism is incentive compatible and all types, even the least efficient one, participate.[21]

Let us denote by $U_i(\theta_i)$ the information rent of an agent A_i with type θ_i:

$$U_i(\theta_i) = \underset{\theta_{-i}}{E} \left(S(q_i(\theta_i, \theta_{-i})) - \theta_i q_i(\theta_i, \theta_{-i}) | \theta_i\right) - H_i(\theta_i). \qquad (12)$$

Individual rationality implies:

$$U_i(\theta_i) \geq 0 \qquad \forall i, \quad \forall \theta_i \in \Theta. \qquad (13)$$

Bayesian incentive compatibility requires:

$$U_i(\theta_i) = \arg\max_{\hat{\theta}_i \in \Theta_i} \underset{\theta_{-i}}{E} \left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i}) | \theta_i\right) - H_i(\hat{\theta}_i) \quad \forall i, \ \forall \theta_i \in \Theta. \qquad (14)$$

What is remarkable here is the similarity of this formula with the Bayesian incentive constraint that would be obtained had types been independently distributed. In that case, the agent's expected payment is independent of his true type and can also be separated

[21] Shutting down the least efficient types is never optimal given the Inada condition $S'(0) = +\infty$.

in the expression of the incentive constraint exactly as the function $H_i(\cdot)$ in (14). This renders the analysis of the set of non-manipulable incentive compatible allocations close to standard mechanism design with independent types.

Assume for simplicity that $q_i(\cdot)$ is differentiable.[22] Simple revealed preferences arguments show that $H_i(\cdot)$ is also differentiable. The local first-order condition for Bayesian incentive compatibility becomes:

$$\dot{H}_i(\theta_i) = \underset{\theta_{-i}}{E}\left((S'(q_i(\theta_i,\theta_{-i})) - \theta_i)\frac{\partial q_i}{\partial \theta_i}(\theta_i,\theta_{-i})|\theta_i\right) \quad \forall i,\ \forall \theta_i \in \Theta; \tag{15}$$

Consider thus any output schedule $q_i(\cdot)$ which is monotonically decreasing in θ_i and which lies below the first-best. From (15), $H_i(\cdot)$ is necessarily also decreasing in θ_i: Less efficient types produce less and pay lower up-front payments. The incentive constraint (15) captures the trade-off faced by an agent with type θ_i. By overreporting, this agent pays a lower up-front payment but he also produces less and enjoys a lower expected surplus. Incentive compatibility is achieved when those two effects just compensate each other.

To highlight the trade-off between efficiency and rent extraction, it is useful to rewrite incentive compatibility in terms of the agents' information rent. Equation (15) becomes:

$$\dot{U}_i(\theta_i) = -\underset{\theta_{-i}}{E}(q_i(\theta_i,\theta_{-i})|\theta_i) + \underset{\theta_{-i}}{E}\left((S(q_i(\theta_i,\theta_{-i})) - \theta_i q_i(\theta_i,\theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right). \tag{16}$$

To better understand the right-hand side of (16), consider an agent with type θ_i mimicking a less efficient type $\theta_i + d\theta_i$. By doing so, the θ_i agent produces the same amount than the $\theta_i + d\theta_i$ one but at a lower marginal cost. This gives to type θ_i a first source of information rent which is worth the first term on this right-hand side. By mimicking the $\theta_i + d\theta_i$ type, a θ_i agent A_i affects also how the principal interprets the other agent's report to adjust A_i's own production. The corresponding marginal rent is the second term on the right-hand side of (16). It may in fact be either positive or negative. Some intuition is provided below after having derived the optimal mechanism.

Finally, the local second-order condition for incentive compatibility can be written as:

$$-\underset{\theta_{-i}}{E}\left(\frac{\partial q_i}{\partial \theta_i}(\theta_i,\theta_{-i})|\theta_i\right) + \underset{\theta_{-i}}{E}\left((S'(q_i(\theta_i,\theta_{-i})) - \theta_i)\frac{\partial q_i}{\partial \theta_i}(\theta_i,\theta_{-i})\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) \geq 0$$
$$\forall i = 1,2,\ \forall \theta_i \in \Theta. \tag{17}$$

[22]Because conditional expectations depend on A_i's type, one cannot derive from revealed preferences arguments that either $q_i(\cdot)$ or $E(q_i(\cdot)|\theta_i)$ is itself monotonically decreasing in θ_i. However, it is possible to use the envelope theorem in integral form (Milgrom and Segal, 2002), to characterize the rent obtained by the agents without assuming differentiability of q_i. Therefore the assumption of differentiability is imposed here only to gain better intuition. See the proof of Proposition 2 for details.

A regular incentive problem is such that the agent's first-order condition (15) is both necessary and sufficient for the optimality of a truthful strategy and the right-hand side of (16) is negative so that countervailing incentives do not arise.

The optimal non-manipulable allocation $\{(q_i(\theta), U_i(\theta_i))_{i=1,2}\}$ solves:

$$(\mathcal{P}): \quad \max_{\{(q_i(\theta),U_i(\theta_i))_{i=1,2}\}} \quad \underset{\theta}{E}\left(\sum_{i=1}^{2} S(q_i(\theta)) - \theta_i q_i(\theta) - U_i(\theta_i)\right)$$

subject to constraints (13) to (17).[23]

To get sharp predictions on the solution, we need to generalize to environments with correlated information the well-known assumption of monotonicity of the virtual cost:

Assumption 1 $\varphi(\theta_i, \theta_{-i}) = \theta_i + \frac{\frac{F(\theta_i)}{f(\theta_i)}}{1+\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\frac{F(\theta_i)}{f(\theta_i)}}$ *is strictly increasing in* θ_i *and decreasing in* θ_{-i}.

The monotonicity of the generalized virtual cost ensures that optimal outputs are non-increasing with own types, a condition which is neither sufficient nor necessary for implementability as it can be seen from (17) but which remains a useful ingredient for it. Note that Assumption 1 implies also the Monotone Likelihood Ratio Property (MLRP) $\frac{\partial}{\partial\theta_{-i}}\left(\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right) \geq 0$ for all $\theta \in \Theta^2$.

Also, we assume that there is a lower bound on the possible level of correlation expressed by the following condition:

Assumption 2

$$\left|\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right| \leq \min\left\{\frac{f(\theta_i)}{F(\theta_i)}, \frac{q^{BM}(\theta_i)}{S(q^{FB}(\theta_i)) - \theta_i q^{FB}(\theta_i)}\right\} \text{ for all } \theta \in \Theta^2,$$

and

$$\max_{(\theta_i,\theta_{-i})\in\Theta^2}\left|\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)\right| \leq \min_{\theta_i\in\Theta} f(\theta_i)\frac{(\min_{(\theta_i,\theta_{-i})\in\Theta^2}\tilde{f}(\theta_{-i}|\theta_i))^2}{2\max_{(\theta_i,\theta_{-i})\in\Theta^2}\tilde{f}(\theta_{-i}|\theta_i)}.$$

[23] With correlated types, the local second-order conditions (17) are not sufficient to guarantee global incentive compatibility even if the agents' utility function satisfies a Spence-Mirrlees condition. However, this is the case if the correlation is small enough as requested by Assumption 2 below. See the proof of Proposition 2 in the Appendix for details.

Assumption 2 ensures that the incentive problem is regular as defined above.[24]

Proposition 2 *Assume that Assumptions 1 and 2 both hold and projects are separable (i.e.,* $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} = 0$*). The agents' incentive problems are regular and the optimal non-manipulable Bayesian mechanism entails:*

• *A downward output distortion* $q^{SB}(\theta_i, \theta_{-i})$ *which satisfies the following "modified Baron-Myerson" formula*

$$S'(q^{SB}(\theta_i, \theta_{-i})) = \varphi(\theta_i, \theta_{-i}), \tag{18}$$

with "no distortion at the top" $q^{SB}(\underline{\theta}, \theta_{-i}) = q^{FB}(\underline{\theta}, \theta_{-i}), \quad \forall \theta_{-i} \in \Theta$ *and the following monotonicity conditions*

$$\frac{\partial q^{SB}}{\partial \theta_{-i}}(\theta) \geq 0 \quad \text{and} \quad \frac{\partial q^{SB}}{\partial \theta_i}(\theta) < 0;$$

• *Agents always get a positive rent except for the least efficient ones*

$$U_i^{SB}(\theta_i) \geq 0 \quad (\text{with } = 0 \text{ at } \theta_i = \bar{\theta}).$$

As already stressed, Bayesian incentive constraints with non-manipulability look very similar to what they are with independent types. This suggests that the trade-off between efficiency and rent extraction that occurs under independent types carries over here also even with correlation. This intuition is confirmed by equation (18) which highlights the output distortion capturing this trade-off.

[24] As an example, consider the bivariate normal distribution truncated on $[\theta_0 - \lambda\sigma^2, \theta_0 + \lambda\sigma^2]^2$ with density

$$\tilde{f}(\theta_1, \theta_2) = \frac{C(\rho, \lambda\sigma^2)}{2\pi\sigma^2(1-\rho^2)^{\frac{1}{2}}} exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(\theta_1-\theta_0)^2}{\sigma^2} + \frac{(\theta_2-\theta_0)^2}{\sigma^2} - 2\frac{\rho}{\sigma^2}(\theta_1-\theta_0)(\theta_2-\theta_0)\right)\right].$$

The case $\rho = 0$ corresponds to independent types. For ρ small enough, we have up to terms of order at least ρ^2 : $C(\rho, \lambda\sigma^2) = (\Phi(\lambda) - \Phi(-\lambda))^{-2} + o(\rho^2)$ where $\Phi(x)$ is the cumulative of the standard normal distribution. Using this property, we derive successively:

$$\tilde{f}(\theta_1, \theta_2) = \frac{1}{2\pi\sigma^2 (\Phi(\lambda) - \Phi(-\lambda))^2} exp\left[-\frac{(\theta_1-\theta_0)^2}{2\sigma^2} - \frac{(\theta_2-\theta_0)^2}{2\sigma^2}\right]\left(1 + \frac{\rho}{\sigma^2}(\theta_1-\theta_0)(\theta_2-\theta_0)\right) + o(\rho^2)$$

and

$$\tilde{f}(\theta_1) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}} (\Phi(\lambda) - \Phi(-\lambda))} exp\left(-\frac{(\theta_1-\theta_0)^2}{2\sigma^2}\right) + o(\rho^2),$$

i.e., each cost is approximatively distributed according to a truncated normal distribution. Finally, the likelihood ratio

$$\frac{\tilde{f}_{\theta_1}(\theta_2|\theta_1)}{\tilde{f}(\theta_2|\theta_1)} = \frac{\rho}{\sigma^2}(\theta_2 - \theta_0) + o(\rho^2),$$

satisfies MLRP and conditions in Assumption 2 are verified when ρ is small enough.

With independent types, the right-hand sides of (2) and (18) are the same. The principal finds useless the report of an agent to better design the other agent's incentives. He must give up some information rent to induce information revelation anyway. Outputs are distorted downward to reduce those rents and the standard Baron-Myerson distortions follow. The optimal mechanism with separable projects and independent types can be implemented with simple bilateral contracts which are *de facto* non-manipulable by the principal. Non-manipulability constraints have no bite in this case.

When types are instead correlated, a similar logic to that of Section 3 applies here with an added twist. Indeed, in our earlier example, non-manipulability puts only a restriction on transfers. More generally, non-manipulability imposes only that the principal's payoff remains constant over all possible transfer-output pairs offered to an agent. This still allows the principal to link agent A_i's payment to what he learns from agent A_{-i}'s report as long as A_i's output varies accordingly. By doing so, the principal may still be able to use the benefits of correlated information. Simple bilateral contracts are not optimal.

To understand the nature of the output distortions and the role of the correlation, it is useful to compare the solution found in (18) with the standard Baron-Myerson formula (2) which corresponds also to the optimal mechanism had the principal offered (non-manipulable) simple bilateral contracts to his agents. Using (16), we observe that the second term on the right-hand side is null for a simple bilateral contract implementing the Baron-Myerson outcome $q^{BM}(\theta_i)$ since $\underset{\theta_{-i}}{E}\left(\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) = 0$.

By having A_i's output depend on θ_{-i}, one departs from the Baron-Myerson outcome, and the principal can use A_{-i}'s report to reduce A_i's information rent. Suppose indeed that the principal starts from the simple bilateral Baron-Myerson contract with A_i but slightly modifies it to improve rent extraction once he has learned A_{-i}'s type. By using A_{-i}'s report the principal should infer how likely it is that A_i lies on his type.

¿From (MLRP) there exists $\theta^*_{-i}(\theta_i)$ such that $\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \geq 0$ if and only if $\theta_{-i} \geq \theta^*_{-i}(\theta_i)$. Hence, the principal's best estimate of A_i's type is θ_i if he learns $\theta_{-i} = \theta^*_{-i}(\theta_i)$ from A_{-i}. Everything happens as if A_{-i}'s report did not bring more information on A_i's type. The only principal's concern remains reducing the first-term on the right-hand side of (16) and the optimal output corresponds to the Baron-Myerson outcome. Think now of an observation $\theta_{-i} > \theta^*_{-i}(\theta_i)$. Such signal let the principal think that the agent has not exaggerated his cost parameter and there is less need for distorting output. The distortion with respect to the first-best outcome is less than Baron-Myerson. Instead, a signal $\theta_{-i} < \theta^*_{-i}(\theta_i)$ confirms the agent's report if he exaggerates his type. This requires increasing the distortion beyond the Baron-Myerson solution.

Corollary 1 *: Under the assumptions of Proposition 2, the following property holds:*

$$q^{SB}(\theta_i, \theta_{-i}) \geq q^{BM}(\theta_i) \quad \Leftrightarrow \quad \theta_{-i} \geq \theta^*_{-i}(\theta_i) \quad \forall \theta_i \in \Theta.$$

6 General Environments

6.1 Characterizing Non-Manipulability

With separable projects, non-manipulability constraints are also separable and it was straightforward to derive the form of non-manipulable schemes. With production externalities, things are more complex. We now propose an approach that enables us to derive second-best distortions in those more general environments.[25]

Suppose that the principal wants to implement the vector of outputs $q(\theta) = (q_1(\theta), q_2(\theta))$ in a non-manipulable way. Assume that the following properties hold for such outputs:

Assumption 3 $q(\theta) = (q_1(\theta), q_2(\theta))$ *is continuously differentiable and satisfies:*

$$\frac{\partial^2 \tilde{S}}{\partial q_i \partial q_{-i}}(q(\theta)) \frac{\partial q_i}{\partial \theta_{-i}}(\theta) \frac{\partial q_{-i}}{\partial \theta_{-i}}(\theta) \geq 0 \qquad \textit{for } i = 1, 2, \forall \theta \in \Theta^2. \tag{19}$$

For substitutes, Assumption 3 is satisfied when a given agent's output decreases with his own cost and increases with that of his peer. For complements, the output of an agent should decrease with both marginal costs.

As it will appear in Lemma 2 below, (19) is indeed a second-order condition ensuring that the principal's best strategy is telling the truth on whatever he has learned from the other agent. This condition is similar to those found in standard screening problems.[26]

Assumption 4 $q(\theta) = (q_1(\theta), q_2(\theta))$ *satisfies:*

$$\left| \frac{\partial q_{-i}}{\partial \theta_{-i}}(\theta) \right| \geq \left| \frac{\partial q_i}{\partial \theta_{-i}}(\theta) \right| \; \forall \theta \in \Theta^2. \tag{20}$$

[25] For simplicity, we still focus on the case of two agents.

[26] In such problems, a single crossing assumption on the agent's utility function is enough to derive the almost everywhere differentiability of the screening variable. Here instead, when dealing with the non-manipulability of his report $\hat{\theta}_{-i}$ vis-à-vis agent A_i computing the cross-derivative of the principal's objective can only be done once it is assumed that the screening variable $q_{-i}(\theta)$ is continuously differentiable with respect to θ_{-i}. This leads us to restrict a priori to differentiable schedules instead of deriving this property from revealed preferences arguments or from the envelope theorem as in the case of separable projects. A similar trick is used in all the applied literature on common agency (see Martimort (1992) and Stole (1991)).

Assumption 4 simply means that the own-impact of an agent's cost parameter on his output is greater than its impact on the other agent's output.

Next lemma provides a local characterization of non-manipulable allocations with continuously differentiable schedules.

Lemma 2 *Assume that $q(\theta)$ satisfies Assumptions 3 and 4. The following necessary first-order conditions for the non-manipulability constraints (7) are also locally sufficient:*

$$\frac{\partial \tilde{S}}{\partial q_i}(q(\theta))\frac{\partial q_i}{\partial \theta_{-i}}(\theta) = \frac{\partial t_i}{\partial \theta_{-i}}(\theta) \ \forall \theta \in \Theta^2. \tag{21}$$

Turning now to the issue of global optimality for the principal of non-manipulating what he has learned, we have:

Lemma 3 *Assume that $q(\theta)$ satisfies Assumptions 3 and 4. A sufficient condition for global optimality of the principal's non-manipulating strategy is:*

$$\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q_1, q_2) = \lambda \in \mathbb{R} \quad \forall (q_1, q_2) \in \mathcal{Q}^2. \tag{22}$$

Integrating (21) immediately yields the following expressions of the transfers:

$$t_i(\theta) = \int_{\underline{\theta}}^{\theta_{-i}} \frac{\partial \tilde{S}}{\partial q_i}(q(\theta_i, x))\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, x)dx - H_i(\theta_i) \text{ for } i = 1, 2 \tag{23}$$

where $H_i(\theta_i)$ is some arbitrary function. For a given output schedule $q(\theta)$ satisfying the conditions of Lemma 2, non-manipulable transfers are thus determined up to some functions $H_i(\cdot)$. The transfers obtained in (23) generalize the sell-out contracts obtained with separable activities to the case of production externalities.

To understand the new distortions involved with a production externality, it is useful thinking of the case of a small production externality (i.e., $\left|\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q_1, q_2)\right| = |\lambda|$ small enough). The principal can still offer sell-out contracts $t_i(\theta) = \tilde{S}(q_i(\theta), 0) - H_i(\theta_i)$ with little modifications of the information rents left to the agents and little changes in allocative efficiency compared to the case without externality. However, these sell-out schemes are now manipulable. To see how, define the principal's payoff when informed on $\theta = (\theta_i, \theta_{-i})$ and choosing a manipulation $\hat{\theta} = (\hat{\theta}_i, \hat{\theta}_{-i})$ as:

$$V(\hat{\theta}, \theta) = \tilde{S}(q_i(\theta_i, \hat{\theta}_{-i}), q_{-i}(\hat{\theta}_i, \theta_{-i})) - \sum_{i=1}^{2} t_i(\theta_i, \hat{\theta}_{-i}).$$

When Assumption 3 holds, we get:

$$\frac{\partial V}{\partial \hat{\theta}_{-i}}(\hat{\theta}, \theta)|_{\hat{\theta}=\theta} = \left(\frac{\partial \tilde{S}}{\partial q_i}(q_i(\theta), q_{-i}(\theta)) - \frac{\partial \tilde{S}}{\partial q_i}(q_i(\theta), 0)\right)\frac{\partial q_i}{\partial \theta_{-i}}(\theta) = \lambda q_{-i}(\theta)\frac{\partial q_i}{\partial \theta_{-i}}(\theta) < 0 \text{ when } \lambda \neq 0.^{27} \tag{24}$$

To limit the scope for manipulating $\hat{\theta}_{-i}$, (24) shows that the principal has at his disposal roughly two strategies. The first one consists in reducing A_{-i}'s output. At the extreme, this would mean committing himself to always deal only with A_i, a non-manipulable but also highly inefficient contract in case agents exert complementary activities. The second strategy consists in making A_i's output q_i less sensitive to θ_{-i} like in a simple bilateral contract. Which strategy is preferred depends on which types realize and the nature of the externality. We now turn to necessary conditions that must be satisfied by the optimal mechanism in such settings with production externality before giving some hints on the nature of those distortions.

First, observe that condition (23) allows us to express the agent's incentive compatibility constraint as:

$$U_i(\theta_i) = \arg\max_{\hat{\theta}_i \in \Theta} \underset{\theta_{-i}}{E} \left(\int_{\underline{\theta}}^{\theta_{-i}} \frac{\partial \tilde{S}}{\partial q_i}(q(\hat{\theta}_i, x)) \frac{\partial q_i}{\partial \theta_{-i}}(\hat{\theta}_i, x) dx - \theta_i q_i(\hat{\theta}_i, \theta_{-i}) | \theta_i \right) - H_i(\hat{\theta}_i). \quad (25)$$

From this we derive the optimal mechanism that the principal would like to implement in this bilateral contracting environment.

Proposition 3 *Assume that (22) is satisfied, and that (19) and (20) hold for that solution and the agents' incentive problems are regular, the optimal non-manipulable output $q^{SB}(\theta)$ solves the system of partial derivative equations*

$$\text{For } i = 1, 2,\ \tilde{f}(\theta) \left(\left(1 + \frac{F(\theta_i)}{f(\theta_i)} \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \right) \left(\frac{\partial \tilde{S}}{\partial q_i}(q^{SB}(\theta)) - \theta_i \right) - \frac{F(\theta_i)}{f(\theta_i)} \right)$$

$$= \lambda \left(F(\theta_{-i}) \left(\int_{\underline{\theta}}^{\theta_i} \tilde{f}_{\theta_{-i}}(x|\theta_{-i}) dx \right) \frac{\partial q_{-i}^{SB}}{\partial \theta_i}(\theta) - F(\theta_i) \left(\int_{\underline{\theta}}^{\theta_{-i}} \tilde{f}_{\theta_i}(x|\theta_i) dx \right) \frac{\partial q_{-i}^{SB}}{\partial \theta_{-i}}(\theta) \right) \quad (26)$$

with the boundary conditions

$$\frac{\partial \tilde{S}}{\partial q_i}(q^{SB}(\underline{\theta}, \theta_{-i})) = \underline{\theta} \text{ and } \frac{\partial \tilde{S}}{\partial q_{-i}}(q^{SB}(\underline{\theta}, \theta_{-i})) = \varphi(\theta_{-i}, \underline{\theta})\ i = 1, 2. \quad (27)$$

The hyperbolic system of first-order partial derivative equations (26) generalizes the Baron-Myerson formula to the case of production externalities. Finding its solutions satisfying the boundary conditions (27) which determine outputs at $\theta_1 = \underline{\theta}$ and $\theta_2 = \underline{\theta}$

[27] This is so since, with substitutes $\frac{\partial q_i}{\partial \theta_{-i}}(\theta) > 0$ but $\frac{\partial \tilde{S}}{\partial q_{-i}}(q_i(\theta), q_{-i}(\theta)) - \frac{\partial \tilde{S}}{\partial q_{-i}}(q_i(\theta), 0) < 0$ when $q_{-i}(\theta) > 0$ whereas it is the reverse for complements.

requires numerical methods. In the Appendix, we nevertheless propose a method to approximate such solution near the boundary defined by (27) when $\lambda \neq 0$.[28] The idea is to find approximations of the characteristic curves associated to the system (26) close to the boundary $(\underline{\theta}, \theta_{-i})$ to solve explicitly the system at least locally. We then check ex post that Assumptions 3 and 4 both hold for the solution when λ is small enough.

Of course, the system (26) with the boundary conditions (27) helps to recover the solutions we already found for separable projects. Beyond that case, non-manipulability constraints force now the principal to take into account any impact of his output choice for agent A_i on the transfer he gives to A_{-i} and this introduces the new terms on the right-hand side of (26). However, as in the case of separable projects, non-manipulability does not matter for independent types. The optimal solution is then the second-best outcome not taking into account the possibility of manipulations. It corresponds to outputs distortions given by the familiar Baron-Myerson conditions:

$$\frac{\partial \tilde{S}}{\partial q_i}(q^{SB}(\theta)) = \theta_i + \frac{F(\theta_i)}{f(\theta_i)}.$$

In order to give more insights on the nature of these distortions, assume that $\tilde{S}(\cdot)$ is quadratic and writes as $\tilde{S}(q) = \mu(q_1 + q_2) - \frac{1}{2}(q_1^2 + q_2^2) + \lambda q_1 q_2$, where $|\lambda| < 1$ to ensure strict concavity of $\tilde{S}(\cdot)$. Denote $l = \frac{\tilde{f}_{\theta_i}(\underline{\theta}|\underline{\theta})}{\tilde{f}(\underline{\theta}|\underline{\theta})}$ the likelihood ratio at $(\underline{\theta}, \underline{\theta})$. This can be viewed as an index of the correlation across types. Assuming strict (MLRP), we have $l < 0$. Any real analytic solution to (26)-(27) close to $(\underline{\theta}, \underline{\theta})$ (which lies on the boundary surfaces defined in (27)) can be approximated locally as follows.

Corollary 2 *Assume that $\tilde{S}(\cdot)$ is quadratic as above and $f(\cdot)$ is real analytic with $-\frac{f'(\underline{\theta})}{2f(\underline{\theta})} = m$. Locally around $(\underline{\theta}, \underline{\theta})$, any symmetric real analytic solution to (26)-(27) admits the following approximation:*

$$q_i^{SB}(\theta_i, \theta_{-i}) - q^{FB}(\underline{\theta}, \underline{\theta}) = -\frac{1}{1-\lambda^2}\left((\theta_i - \underline{\theta}) + (l-m)(\theta_i - \underline{\theta})^2\right)$$

$$+\frac{\lambda}{1-\lambda^2}\left(-2(\theta_{-i} - \underline{\theta}) + (l-m)(\theta_{-i} - \underline{\theta})^2\right) - \frac{2\lambda l}{1-\lambda^2}(\theta_i - \underline{\theta})(\theta_{-i} - \underline{\theta}) + o(||\theta - \underline{\theta}||^2) \quad (28)$$

where $lim_{||\theta-\underline{\theta}|| \to 0} \frac{o(||\theta-\underline{\theta}||^2)}{||\theta-\underline{\theta}||^2} = 0.$

From (28), Assumptions 3 and 4 hold for the optimal output $q^{SB}(\theta)$ at least locally around $(\underline{\theta}, \underline{\theta})$. To understand the nature of the output distortions away from the first-best,

[28] Even when $\tilde{S}(\cdot)$ and $\tilde{f}(\cdot)$ are both real analytic, the Cauchy-Kowalevski Theorem (see John (1982) for instance) cannot be directly used to ensure that a solution to (26) exists which is real analytic close to the boundary defined by (27) since indeed the coefficients of $\frac{\partial q_{-i}^{SB}}{\partial \theta_i}(\theta)$ and $\frac{\partial q_{-i}^{SB}}{\partial \theta_{-i}}(\theta)$ into (26) are both zero simultaneously when the right-hand side of (26) vanishes.

it is necessary to decompose it into three elements. First, there is the *generalized virtual cost effect* that comes on the first bracketed term on the right-hand side of (28). This term survives when there is no production externality and is only due, as in Section 5, to the fact that costs are replaced by virtual generalized costs in evaluating the rent/efficiency trade-off under non-manipulability. Within that bracket, the negative term $-\frac{(l-m)}{1-\lambda^2}(\theta_i-\underline{\theta})^2$ captures how correlation affects optimal outputs. Because learning from agent A_{-i} that his type is close to $\underline{\theta}$ can only be bad news when A_i reports himself a type θ_i s above $\underline{\theta}$, this first effect leads to an exacerbated downward distortion of $q_i^{SB}(\theta_i, \theta_{-i})$. This term is reduced in absolute value when correlation diminishes. The next bracketed term captures the impact of the production externality that would arise if the right-hand side of (26) was set at zero. This *indirect effect of production externality* comes from the fact that, as the *generalized virtual cost effect* distorts the output of a given agent, substitutability or complementarity imply further distortion of the other agent's output. For instance, with substitutes the downward distortion of A_{-i}'s output due to the generalized virtual cost effect leads to raise A_i's output and that all the more that the correlation increases. Finally, the last term on the right-hand side of (28) captures the impact of the production externality on the principal's incentives to manipulate: *a direct effect of production externality.* It represents the extra distortions needed to move sell-out contracts towards being non-manipulable. This last term increases output distortions around $(\underline{\theta}, \underline{\theta})$ for substitutes. Indeed, manipulations vis-à-vis agent A_i are better fought by making A_i's output less sensitive to A_{-i}'s cost. Distortions are instead reduced with complements also for the same reason. Manipulations are then better fought by making outputs less sensitive to the other agent's cost. Note finally that, as the correlation increases (in the sense of having $|l|$ bigger), the direct effect of production externality on output distortions is magnified.

The models of Section 6.2 below further built intuition on distortions induced by non-manipulability constraints.

6.2 Further Results with Discrete Types

The complexity of finding solutions to (26)-(27) suggests to investigate now the nature of optimal non-manipulable mechanisms in a discrete types environment where full-fledged solutions could be found. We shall now assume that each agent's type belongs to $\Theta = \{\underline{\theta}, \bar{\theta}\}$ (denote $\Delta\theta = \bar{\theta} - \underline{\theta}$). The common knowledge distribution of types is still symmetric for simplicity and we let probabilities for the different type realizations be defined as

$$\tilde{p}(\underline{\theta}, \underline{\theta}) = \nu^2 + \alpha, \tilde{p}(\bar{\theta}, \underline{\theta}) = \tilde{p}(\underline{\theta}, \bar{\theta}) = \nu(1-\nu) - \alpha \text{ and } \tilde{p}(\bar{\theta}, \bar{\theta}) = (1-\nu)^2 + \alpha.$$

With that distribution, the marginal distribution of any type is $p(\underline{\theta}) = \nu$, $p(\bar{\theta}) = 1-\nu$ and the correlation coefficient is $\tilde{p}(\underline{\theta}, \underline{\theta})\tilde{p}(\bar{\theta}, \bar{\theta}) - \tilde{p}(\bar{\theta}, \underline{\theta})\tilde{p}(\underline{\theta}, \bar{\theta}) = \alpha$ that we take in $[0, \nu(1-\nu)]$ so that correlation is positive.

If manipulations were not a concern, it would be straightforward to implement the first-best with correlated types and the intuition built in Section 3 suggests that incentives to manipulate matter when the principal may report to either agent that the other has made a report which conflicts with his own. To analyze those incentives, we study two polar cases of interest: perfect complements and perfect substitutes.

6.2.1 Team Production

Consider a team where agents exert efforts q_1 and q_2 which are perfect complements in the production process. Denote by $q = \min(q_1, q_2)$ this output and by $S(q)$ the principal's surplus ($S'(0) = +\infty$, $S' > 0$, $S'' < 0$ with $S(0) = 0$).[29] Any symmetric mechanism is characterized by an output schedule with three possible elements $\{q(\bar{\theta}, \bar{\theta}), q(\bar{\theta}, \underline{\theta}) = q(\underline{\theta}, \bar{\theta}), q(\underline{\theta}, \underline{\theta})\}$ and a four-uple of transfers $\{t(\bar{\theta}, \bar{\theta}), t(\bar{\theta}, \underline{\theta}), t(\underline{\theta}, \bar{\theta}), t(\underline{\theta}, \underline{\theta})\}$.

Given the constraints on output observability, some manipulations are not possible: For instance, pretending that $(\hat{\theta}_1, \hat{\theta}_2) = (\underline{\theta}, \bar{\theta})$ when $(\theta_1, \theta_2) = (\underline{\theta}, \underline{\theta})$ is not feasible given that such report would require implementing $q(\underline{\theta}, \bar{\theta})$ vis-à-vis A_1 and a different output $q(\underline{\theta}, \underline{\theta})$ vis-à-vis A_2. The only two relevant non-manipulability constraints (7) are:

$$S(q(\underline{\theta}, \underline{\theta})) - 2t(\underline{\theta}, \underline{\theta}) \geq S(q(\underline{\theta}, \bar{\theta})) - 2t(\underline{\theta}, \bar{\theta}) \tag{29}$$

and

$$S(q(\bar{\theta}, \bar{\theta})) - 2t(\bar{\theta}, \bar{\theta}) \geq S(q(\bar{\theta}, \underline{\theta})) - 2t(\bar{\theta}, \underline{\theta}). \tag{30}$$

Constraint (29) comes from the fact that the principal can always report to an efficient agent that the other is not even when both are. Constraint (30) captures the fact that the principal can always report to an inefficient agent that the other is not so again even when both are. Since the principal can only lie to both agents at the same time, it is worth noticing that those constraints correspond to global deviations.

Proposition 4 *For α small enough, the optimal symmetric non-manipulable mechanism in a team production context is such that (30) is binding. Optimal outputs are respectively given by:*

$$S'(q^{SB}(\underline{\theta}, \underline{\theta})) = 2\underline{\theta},$$

$$S'(q^{SB}(\underline{\theta}, \bar{\theta})) = \underline{\theta} + \bar{\theta} + \frac{\nu}{1-\nu}\left(\frac{1 + \alpha\frac{3-2\nu}{\nu^2}}{1 - \alpha\frac{1-2\nu}{\nu(1-\nu)}}\right)\Delta\theta, S'(q^{SB}(\bar{\theta}, \bar{\theta})) = 2\bar{\theta} + \frac{2\nu}{1-\nu}\left(\frac{1 - \frac{\alpha}{\nu(1-\nu)}}{1 + \alpha\frac{2\nu}{(1-\nu)^3}}\right)\Delta\theta.$$

[29] The Inada condition again ensures that it is worth always contracting with both agents so that the issue of "shutting-down" the worst types again does not arise. Now the ability of the principal to manipulate reports towards both agents is constrained by the fact that, assuming there is no waste of their individual inputs, both agents observe the same final output on which contracts can be written.

Only an efficient agent gets a strictly positive information rent:

$$U^{SB}(\underline{\theta}) > U^{SB}(\bar{\theta}) = 0.$$

Observe that the optimal outputs above again converges towards the “Baron-Myerson” outcome in the limit of zero correlation. This outcome corresponds to the simple rule consisting in equalizing the marginal efficiency of production to the sum of the agents’ virtual costs.[30] Starting from this benchmark, the principal would like to use types correlation to reduce the rent of an efficient agent. This can be be done by reducing the payment $t(\bar{\theta}, \underline{\theta})$ that this agent could get by lying on his type when facing an efficient agent since such event is rather unlikely with a positive correlation. At the same time, satisfying the participation constraint of an inefficient agent requires also to raise $t(\bar{\theta}, \bar{\theta})$ and altogether those changes in payments makes it attractive to manipulate reports so that (30) necessarily binds. Relaxing this constraint requires to move further apart $q^{SB}(\bar{\theta}, \bar{\theta})$ and $q^{SB}(\underline{\theta}, \bar{\theta})$ bringing the former closer to the first-best and the latter closer to zero.

Turning now to the form of optimal transfers, (30) binding and the possibility of taking also (29) binding means that changes in an agent’s payment are half the incremental surplus triggered by a change in the over agent’s report: an easy generalization of the sell-out contract.

6.2.2 Unit Auction

The principal’s surplus from consuming 1 unit of the good is $S > \bar{\theta}$ so that, under complete information, trade would always be efficient. In this auction environment, each agent may produce that unit on the principal’s behalf. Any symmetric mechanism is characterized by a four-uple of non-negative probabilities for producing the requested unit $\{q(\bar{\theta}, \bar{\theta}), q(\bar{\theta}, \underline{\theta}), q(\underline{\theta}, \bar{\theta}), q(\underline{\theta}, \underline{\theta})\}$ and a four-uple of corresponding transfers $\{t(\bar{\theta}, \bar{\theta}), t(\bar{\theta}, \underline{\theta}), t(\underline{\theta}, \bar{\theta}), t(\underline{\theta}, \underline{\theta})\}$.

Manipulations are now constrained by the fact that the principal cannot lie so that he requests the good with probability more than one. In other words, manipulations are feasible when $q_1(\theta_1, \hat{\theta}_2) + q_2(\hat{\theta}_1, \theta_2) \leq 1$ for all $(\hat{\theta}_1, \hat{\theta}_2)$.

Intuition built by looking at the form taken by the efficient allocation suggests a solution such that the good is produced when both agents tie and report being efficient, i.e., $q(\underline{\theta}, \underline{\theta}) = \frac{1}{2}$ and that the good is allocated more often to the efficient agent when the other agent reports being inefficient, i.e., $q(\underline{\theta}, \underline{\theta}) \leq q(\underline{\theta}, \bar{\theta})$. With those conditions,

[30] It entails thus a double distortion $\frac{2\nu}{1-\nu}\Delta\theta$ when those two types are inefficient whereas there is a simple distortion $\frac{\nu}{1-\nu}\Delta\theta$ when only one agent is inefficient.

the principal certainly cannot manipulate reports in state $(\underline{\theta}, \underline{\theta})$ since this would lead to propose to buy the good with probability more than one. Manipulations are a priori feasible for other type realizations. In states $(\underline{\theta}, \bar{\theta})$, non-manipulability constraints become:

$$S(q(\underline{\theta}, \bar{\theta}) + q(\bar{\theta}, \underline{\theta})) - t(\underline{\theta}, \bar{\theta}) - t(\bar{\theta}, \underline{\theta})$$
$$\geq \max\{S(q(\underline{\theta}, \underline{\theta}) + q(\bar{\theta}, \underline{\theta})) - t(\underline{\theta}, \underline{\theta}) - t(\bar{\theta}, \underline{\theta}); S(q(\underline{\theta}, \underline{\theta}) + q(\bar{\theta}, \bar{\theta})) - t(\underline{\theta}, \underline{\theta}) - t(\bar{\theta}, \bar{\theta})\} \quad (31)$$

provided the corresponding manipulations are feasible, i.e., $q(\underline{\theta}, \underline{\theta}) + q(\bar{\theta}, \underline{\theta}) \leq 1$ and $q(\underline{\theta}, \underline{\theta}) + q(\bar{\theta}, \bar{\theta}) \leq 1$. Instead, in state $(\bar{\theta}, \bar{\theta})$, non-manipulability constraints write as:

$$2Sq(\bar{\theta}, \bar{\theta}) - 2t(\bar{\theta}, \bar{\theta}) \geq \max\{S(q(\bar{\theta}, \underline{\theta}) + q(\bar{\theta}, \bar{\theta})) - t(\bar{\theta}, \underline{\theta}) - t(\bar{\theta}, \bar{\theta}); 2Sq(\bar{\theta}, \underline{\theta}) - 2t(\underline{\theta}, \bar{\theta})\}$$

whenever $q(\bar{\theta}, \underline{\theta}) + q(\bar{\theta}, \bar{\theta}) \leq 1$. This can be simplified as

$$Sq(\bar{\theta}, \bar{\theta}) - t(\bar{\theta}, \bar{\theta}) \geq Sq(\bar{\theta}, \underline{\theta}) - t(\bar{\theta}, \underline{\theta}). \quad (32)$$

In contrast with the case of perfect complements seen above, this constraint is now local. It entails only a possible deviation towards a single agent.

Proposition 5 *For α small enough, the optimal non-manipulable auction is such that (32) is binding. Optimal probabilities are respectively given by:*

$$q^{SB}(\underline{\theta}, \underline{\theta}) = \frac{1}{2}, q^{SB}(\underline{\theta}, \bar{\theta}) = 1, q^{SB}(\bar{\theta}, \underline{\theta}) = 0,$$

and

$$q^{SB}(\bar{\theta}, \bar{\theta})) = \begin{cases} \frac{1}{2} & \text{if } S \geq \bar{\theta} + \frac{\nu}{1-\nu}\left(\frac{1-\frac{\alpha}{\nu(1-\nu)}}{1+\alpha\frac{1+\nu}{\nu(1-\nu)^2}}\right)\Delta\theta \\ 0 & \text{otherwise.} \end{cases}$$

Again, the optimal non-manipulable auction above implements an allocation which converges also towards the allocation implemented with independent types in the limit of zero correlation even if in this latter case the non-manipulability constraints are not taken into account. For a positive correlation and when the principal's surplus is large enough, the optimal auction has non-zero payments in all states of nature. Agents are awarded production with probability one half in case of a tie even if both are inefficient. The threshold on surplus for such efficient trade decreases with the correlation, showing that more correlated environments facilitate efficient trade.

That (32) is binding means that the principal is indifferent between lying or not to an inefficient agent. Again the mechanism looks like a generalized sell-out.

7 Extensions

7.1 Dominant Strategy and Simple Bilateral Contracting

Let us come back to the case where $\Theta = [\underline{\theta}, \bar{\theta}]$. Section 5 showed that, in a Bayesian setting, any information learned by the principal when contracting with a given agent is used to regulate another agent when types are correlated. We now strengthen the implementation concept and require that agents play dominant strategies in the mechanism offered by the principal. We ask whether it makes optimal non-manipulable mechanisms look more like a set of simple bilateral contracts: an extreme case of non-manipulability.

Notice first that the notion of non-manipulability is independent of the implementation concept used to describe the agents' behavior. Our framework can be easily adapted to dominant strategy implementation. For any arbitrary mechanism $(g(\cdot), \mathcal{M})$, a dominant strategy continuation equilibrium is characterized through the following Lemma.[31]

Lemma 4 *A continuation dominant strategy equilibrium induces a pair* $\{m^*(\cdot), \hat{m}^*(\cdot)\}$ *such that:*

- $m^*(\theta) = (m_1^*(\theta_1), ..., m_n^*(\theta_n))$ *from* Θ^n *into* $\mathcal{M} = \Pi_{i=1}^n \mathcal{M}_i$ *forms a dominant strategy equilibrium given the principal's optimal manipulation* $\hat{m}^*(\cdot)$

$$m_i^*(\theta_i) \in \arg\max_{m_i \in \mathcal{M}_i} t_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i})) - \theta_i q_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i})), \forall m_{-i} \in \mathcal{M}_{-i}; \quad (33)$$

- *The principal's optimal manipulation* $\hat{m}^*(\cdot) = (\hat{m}_{-1}^*(\cdot), ..., \hat{m}_{-n}^*(\cdot))$ *from* $\mathcal{M}$ *on* $\Pi_{i=1}^n \mathcal{M}_{-i}$ *satisfies (5).*

We immediately adapt our previous findings to get:

Proposition 6 ***The Revelation Principle for Dominant Strategy Implementation with Bilateral Contracting.*** *In a private values context, any allocation* $a(\cdot)$ *achieved at a dominant strategy equilibrium of any arbitrary mechanism* $(g(\cdot), \mathcal{M})$ *with bilateral contracting can alternatively be implemented as a truthful and non-manipulable dominant strategy equilibrium of a direct mechanism* $(\bar{g}(\cdot), \Theta^n)$.

Non-manipulability is independent of the implementation concept and still obtained with sell-out contracts if agents work on separable projects:

$$t_i(\theta_i, \theta_{-i}) = S(q_i(\theta_i, \theta_{-i})) - H_i(\theta_i).$$

[31] The proof mimics that of Lemma 1 and is thus omitted.

Denoting $u_i(\theta_i, \theta_{-i}) = t_i(\theta_i, \theta_{-i}) - \theta_i q_i(\theta_i, \theta_{-i})$ the ex post rent of an agent with type θ_i when the other agent reports θ_{-i}, dominant strategy incentive compatibility implies that $q_i(\theta_i, \theta_{-i})$ is weakly decreasing in θ_i, for all θ_{-i}, and

$$u_i(\theta_i, \theta_{-i}) = u_i(\bar{\theta}, \theta_{-i}) + \int_{\theta_i}^{\bar{\theta}} q_i(u, \theta_{-i}) du. \tag{34}$$

We also strengthen the participation condition and impose ex post participation constraints which hold irrespectively of the agents' beliefs on each other types:

$$u_i(\theta_i, \theta_{-i}) \geq 0, \quad \forall (\theta_i, \theta_{-i}) \in \Theta^2.$$

Proposition 7 *Assume projects are separable (i.e., $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} = 0$), dominant strategy implementation and ex post participation. The optimal non-manipulable mechanism can be achieved with a pair of simple bilateral contracts $(t_i^{BM}(\theta_i), q_i^{BM}(\theta_i))$ implementing the Baron-Myerson output for each agent:*

$$t_i^{BM}(\theta_i) = \theta_i q_i^{BM}(\theta_i) + \int_{\theta_i}^{\bar{\theta}} q_i^{BM}(u) du.$$

With dominant strategy and non-manipulability, informational externalities can no longer be exploited and the principal cannot do better than offering simple bilateral contracts. Therefore, the Baron-Myerson outcome is optimal even with correlated types.

Remark 7: Simple bilateral contracts are suboptimal if we do not impose non-manipulability even under dominant strategy implementation and ex post participation. Insisting only on dominant strategy and ex post participation, the optimal quantities are given by Baron-Myerson formulae taking into account the fact that the principal uses the correlation of types to update his beliefs accordingly. We get:

$$S'(q_i(\theta_i, \theta_{-i})) = \theta_i + \frac{\tilde{F}(\theta_i \mid \theta_{-i})}{\tilde{f}(\theta_i \mid \theta_{-i})}.$$

The optimal mechanism without the non-manipulability constraint yields a strictly higher payoff than a pair of bilateral contracts when types are correlated. Non-manipulability and dominant strategy implementation are two different concepts with different implications. These restrictions justify simple bilateral contracts only when taken in tandem. ■

7.2 Horizontal Collusion and Simple Bilateral Contracting

We now investigate the possibility of collusion between the agents. We do not tackle the important issue (left for future research) of knowing whether bilateral contracting might facilitate or hinder collusion.[32] Nevertheless, it is important to understand what sort of constraints would be added by also taking into account the agents' collusive behavior if any. To simplify let us again focus on the case of separable projects and, following Laffont and Maskin (1980), suppose that, when colluding, agents learn each other's types.

Given the form that (symmetric) non-manipulable mechanisms take in this environment (see (11)), coalition incentive compatibility requires:

$$(\theta_1, \theta_2) \in \arg \max_{(\hat{\theta}_1, \hat{\theta}_2) \in \Theta^2} \sum_{i=1}^{2} S(q_i(\hat{\theta}_i, \hat{\theta}_{-i})) - \theta_i q_i(\hat{\theta}_i, \hat{\theta}_{-i}) - H(\hat{\theta}_i).$$

This yields the necessary first-order conditions:

$$-H'(\theta_k) + \sum_{i=1}^{2} \left(S'(q_i(\theta_k, \theta_{-k})) - \theta_i\right) \frac{\partial q_i}{\partial \theta_k}(\theta_k, \theta_{-k}) = 0 \quad \text{for } k = 1, 2. \tag{35}$$

These collusion-proofness conditions are helpful to show the following result.

Proposition 8 *Assume that agents work on separable projects and can collude:*

- *The optimal mechanism described in Proposition 2 is not collusion-proof.*
- *The only differentiable output schedules $q_i(\theta_i, \theta_{-i})$ such that*

$$q_i(\theta_i, \theta_{-i}) \leq q^{FB}(\theta_i) \text{ and } \frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) \geq 0$$

and which can be implemented by a collusion-proof and non-manipulable mechanism are such that $\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) = 0$. The optimal mechanism within this class is a pair of simple bilateral contracts implementing the Baron-Myerson outcome $q^{BM}(\theta_i)$.

The optimal mechanism characterized in Proposition 2 when agents do not collude makes the output of any given agent depend also on the report of the other. From the coalition's viewpoint, revealing truthfully his own type is not optimal however. Indeed, given that agent A_i produces below the first-best for that mechanism, the coalition would like that he overstates his report because revealing such information has a positive effect on A_{-i}'s payoff. This points at the difficulty in reconciling non-manipulability and collusion-proofness unless the principal gives up any attempt to make the contract of an agent depend on information he learns from the other. Under some weak conditions, the only possibility left is to offer simple bilateral contracts.

[32]This requires a theory showing how the transaction costs of collusive behavior between asymmetrically informed agents are affected by bilateral contracting. Such theory is beyond the scope of this paper.

7.3 Secret Contracts

Our analysis so far focused on the case where the mechanisms offered by the principal are public. This assumption allowed us to focus on the role of privacy in communication only. An extra degree of privacy arises when the principal offers secret mechanisms to each agent. In this case, not only the particular choice of A_{-i} within the menu he receives but the mere menus are not observed by A_i. We investigate now how our previous results should be modified when assuming secret bilateral contracts.

In the case of separable projects, avoiding the principal's manipulation on whatever messages he receives from the other agents still imposes that

$$S(q_i^S(\theta_i, \theta_{-i})) - t_i^S(\theta_i, \theta_{-i}) = H_i^S(\theta_i),$$

for some functions $H_i^S(\cdot)$ where $(t_i^S(\theta_i, \theta_{-i}), q_i^S(\theta_i, \theta_{-i}))$ is the direct mechanism[33] offered to A_i in the game with secret bilateral contracts. It should be noted at this stage that with separable projects, the offer $(t_{-i}^S(\cdot), q_{-i}^S(\cdot))$ made to agent A_{-i} does not influence how the principal manipulates the report he makes to agent A_i. As a consequence, whether the offers are public or secret does not change the incentives faced by agent A_i: the conjectures about the offer made to A_{-i} following any unexpected deviation that the principal may envision $\{t_i(\theta), q_i(\theta)\}_{\theta \in \Theta^2}$ do not intervene in the reasoning of agent A_i. Therefore, we can immediately replicate the analysis made in Section 5.

Proposition 9 *When projects are separable (i.e., $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} = 0$), the equilibrium outcomes of the game with public contracts and the game with secret contracts coincide.*

8 Relationships with the Literature

This paper is linked to several trends of the mechanism design literature reviewed below:

Partial Commitment: Our modelling of the principal's limited inability to commit to a grand-mechanism leads to a tractable characterization of non-manipulable mechanisms with bilateral contracting by means of a simple Revelation Principle. More generally, models with partial commitment require giving up such simple approach and might involve partially revealing strategies on the agents' side (Bester and Strausz 2000 and 2001, Krishna and Morgan 2006). This latter difficulty is avoided in our context because we focus on private values environments where the principal's utility function does not directly depend on the agents' types. Hence, agents do not manipulate their reports to the

[33] In this setting, the restriction to direct mechanisms is without loss of generality, see the proof of Proposition 9 in the Appendix.

principal to affect his beliefs about their types and influence his optimal manipulations. Whatever information is learned by the principal with an agent, non-manipulability requires that it is truthfully revealed to others. Non-manipulability constraints can thus be interpreted as incentive compatibility constraints with respect to the information that the principal endogenously learns from the agents. This is reminiscent of the *posterior implementability* concept developed by Green and Laffont (1987) in which agents' equilibrium strategies are best-responses to each other even after they learned information revealed by the play of the mechanism itself. However, non-manipulability concerns the principal.

Finally, Baliga, Corchon and Sjostrom (1997) investigate implementation when the mediator himself is a player and reacts to whatever information privately informed agents may report by choosing a decision. Formally, the mechanism design game is transformed into a signalling game. We are less extreme in modelling the principal's lack of commitment and still allow some commitment to bilateral contracts.

Mechanism design in correlated environments: The striking results on the irrelevance of private information in correlated information environments have already been attacked on various fronts. A first approach is to introduce exogenous limits or costs on feasible punishments by means of risk-aversion and wealth effects (Robert 1991, Eso 2004), limited liability (Demougin and Garvie 1991), ex post participation constraints (Demski and Sappington 1988, Dana 1993), or limited enforceability (Compte and Jehiel 2006). In our paper instead, the benefits of using correlated information is undermined by non-manipulability constraints on the principal's side.

A second approach argues that correlation may not be as generic as suggested by the earlier literature. Enriching the information structure may significantly simplify mechanisms. Neeman (2004) points out that the type of an agent should not simultaneously determine his beliefs on others and be payoff-relevant. Such extension of the type space might reinstall some sort of conditional independence and avoid full extraction.[34] Bergemann and Morris (2005) model higher order beliefs and show that robust implementation may amount to ex post implementation.[35] Chung and Ely (2007) show that a maxmin principal may want to rely on dominant strategy implementation. Although important, these approaches lead also to somewhat extreme results since Bayesian mechanisms end up being given up.[36] Our approach still relaxes the common knowledge requirements assumed in standard mechanism design but bilateral contracting between the principal/mediator

[34] Heifetz and Neeman (2006) exhibit conditions under which this conditional independence is generic.

[35] If the aim of the analysis is to model long-run institutions, it is not clear that agents remain in such high degree of ignorance on each other unless they are also boundedly rational and cannot learn about others' types distributions from observing past performances.

[36] This might appear as too extreme in view of the recent (mostly) negative results pushed forward by the ex post implementation literature in interdependent values environments (Dasgupta and Maskin 2000, Perry and Reny 2002 and Jehiel and al. 2006)

and his agents does so in a simple and tractable way.[37] As a result, Bayesian implementation keeps much of its force. Resolution techniques to derive optimal mechanisms are also similar to those already well-known for independent types.

A last approach to avoid full surplus extraction in correlated environments consists in adding collusive behavior. Laffont and Martimort (2000) show that mechanisms extracting entirely all the agents' rent are not robust to horizontal collusion between the agents.[38] Key to this horizontal collusion possibility is the fact that agents can coordinate their strategies in any grand-mechanism offered by the designer. This coordination is certainly harder when no such grand-mechanism is available and agents contract separately with their principal. Our focus on bilateral contracting points at another polar case which leaves less scope for such horizontal collusion but introduces the possibility of manipulations by the principal himself.[39]

Subjective evaluations: There is a literature on the design of incentive contracts between a principal and his agents in moral hazard contexts where the principal's evaluation of the agents' performance is subjective, i.e., private information of the principal himself. McLeod (2003) and Fuchs (2007) show the difficulty in inducing simultaneously a principal to report truthfully whatever evaluations he has on an agent and to induce the latter's effort which requires having the principal paying sometimes more for the agent's output and sometimes less. One recommended solution is to restore those incentives by having the principal "burning money." Another solution, suggested by Rahman and Obara (2007), is to use correlated strategy as the implementation concept. Then, a mediator can make private recommendations on the agent's effort and on which subjective evaluation strategy should be followed by the principal. In our model, the principal's knowledge of the other agents' reports can be viewed as a private signal on an agent's type (exactly as in Section 3). The solutions above might have some appeal. However, both the option of "burning money" and that of appointing another mediator making secret recommendations rely on the ability of the principal not to collude with a third-party. If such collusion arises, those solutions lose their bites and we are back to our original analysis.

Contractual externalities with bilateral contracts: Our work is also related to the IO literature on bilateral contracting (for instance, Hart and Tirole 1990, O'Brien and Shaffer 1992, McAfee and Schwartz 1994, Segal 1999 and Segal and Whinston 2003).

[37]Readers accustomed with the moral hazard literature know that correlation between the agents' performances may be used to better design incentives without of course voiding the agency problem of its interest. Our results have the same flavor.

[38]Their model has only two agents. With more than two agents and in the absence of sub-coalitional behavior, Che and Kim (2006) showed that correlation can still be used to the principal's benefits.

[39]Gromb and Martimort (2007) proposed a specific model of expertise involving both moral hazard in information gathering and adverse selection and showed that private communication between the principal and each of his experts opens the possibility for some vertical collusion which is harmful for the organization.

Those papers analyze complete information environments with the assumption that a manufacturer (principal) can contract with his retailers (agents) only through simple bilateral contracts. Their focus is on the consequences of the principal's opportunistic behavior that arises when he strikes each of those bilateral deals independently and the retailers' payoffs depend on each others' contracting variables with the principal. Although, we share with this literature some concerns in studying the opportunistic behavior of a principal, this is in a different context. Our paper concentrates instead on informational externalities across agents. Since non-manipulability constraints depend only on the principal's payoff, introducing payoff externalities between agents would not change significantly our analysis. Moreover, for most of our analysis above, mechanisms are public and, the principal's opportunistic behavior comes from his possibilities to manipulate communication and not to sign independent secret deals.

Common agency: Our modelling of the principal's lack of commitment is reminiscent of the common agency literature.[40] This should come at no surprise. In our framework, the key issue is to prevent the principal's opportunistic behavior vis-à-vis each of his agents. Under common agency, the same kind of opportunistic behavior occurs, with the common agent reacting to his principals' offers. Beside the allocation of bargaining powers between parties, there is another difference between common agency and the environment described in this paper. The principal has more commitment power here since he can design this mechanism in a first stage and restrict the choice of the uninformed agents. Although a priori minor, this latter difference simplifies the analysis. This instilled minimal level of commitment allows us to maintain much of the optimization techniques available in standard mechanism design without falling into the difficulties faced when characterizing Nash equilibria in the context of multi-contracting mechanism design.[41]

9 Conclusion

This paper has relaxed the assumption of a centralized grand-contracting in mechanism design. Considering bilateral contracts paves the way to a theory which responds to some of the most often heard criticisms addressed to the mechanism design methodology. Even in correlated information environments, considering non-manipulable mechanisms restores a genuine trade-off between efficiency and rent extraction. This leads to a more standard

[40] Bernheim and Whinston (1986), Stole (1991), Martimort (1992, 2007), Mezzetti (1997), Martimort and Stole (2002, 2003, 2007), Peters (2001, 2003). Most often private information is modelled on the common agent's side in this literature, an exception being Martimort and Moreira (2007).

[41] The most noticeable difficulty being of course the multiplicity of equilibria. Martimort (2007) argues that one should look for minimal departures of the centralized mechanism design framework which go towards modelling multi-contracting settings to avoid this difficulty. The non-manipulability constraint modeled above can precisely be viewed as such a minimal departure.

second-best analysis. In several environments of interest (separable projects, auctions, team productions, more general production externalities), we analyzed this trade-off and characterized optimal non-manipulable mechanisms.

Each of these particular settings certainly deserves further studies either by specializing the information structure, by generalizing preferences or by focusing on organizational problems coming from the analysis of real world institutions in particular contexts (political economy, regulation, vertical restraints in a IO context, etc..).

Of particular importance may be the extension of our framework to the case of auctions with interdependent valuations and/or common values. Our approach for simplifying mechanisms could be an attractive alternative to the somewhat too demanding ex post implementation pushed forward by the recent vintage of the literature on that topic. More generally, the analysis of non-manipulable trading mechanisms in correlated environments deserves further analysis. Simple institutions like market mechanisms might perform quite well if one insists on non-manipulability.

Introducing a bias in the principal's preferences towards either agent could also raise interesting issues. First by making the principal's objective function somewhat congruent with that of one of the agents, one goes towards a simple modelling of vertical collusion and favoritism. Second, this congruence may introduce interesting aspects related to the common values element that arises in such environment and that have been set aside by our focus on a private values setting.

Also, it would be worth investigating further what is the scope for horizontal collusion between agents in the environments depicted in this paper. Considering collusion may justify the constraint on bilateral contracting in the first place. Indeed, bilateral contracting introduces private communication between agents which may make it difficult for agents to enforce any collusive agreement compared to the case of a grand-mechanism making all players' strategies public information. Such analysis could lead to an interesting trade-off between the cost of the principal's opportunism under bilateral contracting and the fact that collusion is facilitated with more centralized procedures.

In practice, the degree of transparency of communication in an organization may be intermediate between what we have assumed here with bilateral contracting and the more usual assumption of having a grand-mechanism bringing all agents to the "contracting table". Reputation-like arguments on the principal's side might help in circumventing non-manipulability constraints but the extent by which it is so remains to uncover.

All those are extensions that we plan to analyze in further research.

References

Aspremont, C. d', Crémer, J., and L.A. Gérard-Varet (1990): "Incentives and the Existence of Pareto-Optimal Revelation Mechanisms," *Journal of Economic Theory*, 51, 233-254.

Baliga, S., L. Corchon and T. Sjostrom (1997): "The Theory of Implementation when the Planner is a Player," *Journal of Economic Theory*, 77, 15-33.

Baron, D. and R. Myerson (1982): "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50, 911-930.

Bergemann, D. and S. Morris (2005): "Robust Mechanism Design," *Econometrica*, 73, 1771-1814.

Bernheim, D. and M. Whinston (1986): "Menu Auctions, Resource Allocations and Economic Influence," *Quarterly Journal of Economics*, 101, 1-31.

Bester, H. and R. Strausz (2000): "Imperfect Commitment and the Revelation Principle: The Multi-Agent Case," *Economics Letters* 69, 165-171.

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

Biglaiser, G. and C. Mezzetti (1993): "Principals Competing for an Agent in the Presence of Adverse Selection and Moral Hazard," *Journal of Economic Theory*, 61, 302-330.

Che, Y.-K. and J. Kim (2006): "Robustly Collusion-Proof Implementation," *Econometrica*, 74, 1063-1108.

Chung, K.S. and J. Ely (2007): "Foundations for Dominant Strategy Mechanisms," *Review of Economic Studies*, 74, 447-476.

Compte, O. and P. Jehiel (2006): "Veto Constraints in Mechanism Design: Inefficiency with Correlated Types," mimeo PSE CERAS.

Crémer, J., and R. McLean (1985): "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist when Demands are Interdependent," *Econometrica*, 53, 345-362.

Crémer, J., and R. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica* , 56, 1247-1257.

Dana, J. (1993): "The Organization and Scope of Agents: Regulating Multiproduct Industries," *Journal of Economic Theory*, 59, 288-310.

Dasgupta, P. and E. Maskin (2000): “Efficient Auctions,” *Quarterly Journal of Economics*, 115, 341-388.

Demougin, D. and D. Garvie (1991): “Contractual Design with Correlated Information under Limited Liability,” *RAND Journal of Economics*, 22, 477-489.

Demski, J. and D. Sappington (1989): “Incentive Schemes with Multiple Agents and Bankruptcy Constraints,” *Journal of Economic Theory*, 44, 156-167.

Esö, P. (2004): “An Optimal Auction with Correlated Values and Risk-Aversion,” *Journal of Economic Theory*, 125, 78-89.

Fuchs, W. (2007): “Contracting with Repeated Moral Hazard and Subjective Evaluations,” *American Economic Review*, 97, 1432-1448.

Gelfand, I. and S. Fomin, (2000), *Calculus of Variations,* Dover: Mineola.

Gibbard, A. (1973): “Manipulation for Voting Schemes,” *Econometrica*, 41, 617-631.

Green, J. and J.-J. Laffont (1977): “Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods,” *Econometrica*, 45, 427-438.

Green, J. and J.-J. Laffont (1987): “Posterior Implementability in a Two-Person Decision Problem,” *Econometrica*, 55, 69-94.

Gromb, D. and D. Martimort (2007): “Collusion and the Organization of Delegated Expertise,” forthcoming *Journal of Economic Theory.*

Guesnerie, R. and J.-J. Laffont (1984): “A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm,” *Journal of Public Economics*, 25, 329-369.

Hart, O. and J. Tirole (1990): “Vertical Integration and Market Foreclosure,” *Brooking Papers in Economic Activity*, 205-285.

Heifetz, A. and Z. Neeman (2006): “On the Generic (Im-)possibility of Surplus Full Extraction in Mechanism Design,” *Econometrica*, 74, 213-233.

Jehiel, P., M. Meyer-Ter-Vehn, B. Moldovanu, and W. Zame (2006): “The Limits of Ex-Post Implementation,” *Econometrica*, 74, 585-610.

John, F. (1982): *Partial Differential Equations*, Berlin: Springer.

Johnson, S., J. Pratt and R. Zeckhauser (1990): “Efficiency Despite Mutually Payoff-Relevant Information: The Finite Case,” *Econometrica*, 58, 873-900.

Krishna, V. and J. Morgan (2006): "Contracting for Information under Imperfect Commitment," *mimeo*, Penn State University.

Laffont, J.-J. and D. Martimort (1997): "Collusion under Asymmetric Information," *Econometrica*, 65, 875-912.

— (2000): "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309-342.

— (2002): *The Theory of Incentives: The Principal-Agent Model*, Princeton: Princeton University Press.

Laffont, J.-J. and E. Maskin (1980): "A Differential Approach to Dominant Strategy Mechanisms," *Econometrica*, 48, 1507-1520.

Laffont, J.-J. and J. Tirole (1993): *A Theory of Incentives in Procurement and Regulation*, Cambridge: MIT Press.

Lewis, T. and D. Sappington (1989): "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294-313.

Martimort, D. (1992): "Multi-Principaux avec Anti-Selection," *Annales d'Economie et de Statistiques*, 28, 1-38.

— (2007): "Multicontracting Mechanism Design," in *Advances in Economic Theory , Proceedings of the 2005 World Congress of the Econometric Society.*

Martimort, D. and H. Moreira (2007): "Common Agency with Informed Principals," mimeo IDEI Toulouse.

Martimort, D. and L. Stole (2002): "The Revelation and Delegation Principles in Common Agency Games," *Econometrica*, 70, 1659-1674.

— (2003): "Contractual Externalities and Common Agency Equilibria," *Advances in Theoretical Economics*, Vol. 3, Issue 1, Article 4. http://www.bepress.com/bejte.

— (2007): "Selecting Equilibria in Common Agency Games," mimeo IDEI and University of Chicago.

Maskin, E. and J. Tirole (1990): "The Principal-Agent Relationship with an Informed Principal, 1: Private Values," *Econometrica*, 58, 379-410.

Matsushima, H., (1991): "Incentive Compatible Mechanisms with Full Transferability," *Journal of Economic Theory*, 54, 198-203.

McAfee, P. and M. Schwartz (1994): "Opportunism in Multilateral Vertical Contracting: Non-Discrimination, Exclusivity and Uniformity," *American Economic Review*, 84, 210-230.

McAfee, P. and P. Reny (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395-421.

McLeod, B. (2003): "Optimal Contracting with Subjective Evaluations," *American Economic Review*, 93, 216-240.

Mezzetti, C. (1997): "Common Agency with Horizontally Differentiated Principals," *RAND Journal of Economics*, 28, 323-345.

Milgrom, P. and I. Segal, (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583-601.

Myerson, R., (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Models", *Journal of Mathematical Economics,* 10: 67-81.

— (1991): *Game Theory: Analysis of Conflict*, Harvard: Harvard University Press.

O'Brien, D. and G. Shaffer (1992): "Vertical Control with Bilateral Contracts," *RAND Journal of Economics,* 23, 299-308.

Perry, M. and P. Reny (2002): "An Ex Post Efficient Auction," *Econometrica*, 70, 1199-1212.

Peters, M. (2001): "Common Agency and the Revelation Principle," *Econometrica*, 69, 1349-1372.

— (2003): "Negotiation and Take-It-Or-Leave-It in Common Agency," *Journal of Economic Theory*, 111, 88-109.

Neeman, Z. (2004): "The Relevance of Private Information in Mechanism Design," *Journal of Economic Theory*, 117, 55-77.

Rahman, D. and I. Obara (2007): "Secret Contracts for Efficient Partnerships," mimeo University Minnesota.

Riordan, M. and D. Sappington (1988): "Optimal Contracts with Public Ex Post Information," *Journal of Economic Theory*, 45, 189-199.

Robert, J. (1991): "Continuity in Auction Design," *Journal of Economic Theory*, 55, 169-179.

Segal, I. (1999): "Contracting with Externalities," *Quarterly Journal of Economics*, 104, 337-388.

Segal, I. and M. Whinston (2003): "Robust Predictions for Bilateral Contracting with Externalities," *Econometrica*, 71, 757-792.

Stole, L. (1991): "Mechanism Design under Common Agency," mimeo University of Chicago.

Appendix

• **Proof of Lemma 1:** Take any arbitrary mechanism $(g(\cdot), \mathcal{M}) = ((g_1(\cdot), \mathcal{M}_1), ..., (g_n(\cdot), \mathcal{M}_n))$ for any arbitrary communication space $\mathcal{M} = \Pi_{i=1}^n \mathcal{M}_i$. Consider also a perfect Bayesian continuation equilibrium of the overall contractual game induced by $(g(\cdot), \mathcal{M})$. Such continuation PBE is a triplet $\{m^*(\cdot), \hat{m}^*(\cdot), d\mu(\theta|m)\}$ that satisfies:

• Agent A_i with type θ_i reports a private message $m_i^*(\theta_i)$ to the principal. The strategy $m^*(\theta) = (m_1^*(\theta_1), ..., m_n^*(\theta_n))$ forms a Bayesian-Nash equilibrium among the agents. The corresponding equilibrium conditions are stated in (4) .

• P updates his beliefs on the agents' types following Bayes' rule whenever possible, i.e, when $m \in \text{supp} \quad m^*(\cdot)$. Otherwise, beliefs are arbitrary. Let denote $d\mu(\theta|m)$ the updated beliefs following the observation of a vector of messages m.

• Given any such vector m (either on or out of the equilibrium path) and the corresponding posterior beliefs, the principal reports the messages $(\hat{m}_{-1}^*(m), ..., \hat{m}_{-n}^*(m))$ which maximizes his expected payoff, i.e.,

$$(\hat{m}_{-1}^*(m), ..., \hat{m}_{-n}^*(m))$$

$$\in \arg \max_{(\hat{m}_{-1},...,\hat{m}_{-n}) \in \Pi_{i=1}^n \mathcal{M}_{-i}} \int_{\Theta^n} \left\{ \tilde{S}(q_1(m_1, \hat{m}_{-1})), ..., q_n(m_n, \hat{m}_{-n})) - \sum_{i=1}^n t_i(m_i, \hat{m}_{-i}) \right\} d\mu(\theta|m). \tag{A.1}$$

Because we are in a private values context where the agents' types do not enter directly into the principal's utility function, expectations do not matter and (A.1) can be rewritten more simply as (5).

• **Proof of Proposition 1:** Consider the agents' Bayesian incentive compatibility conditions that must be satisfied by $m^*(\cdot)$. For A_i, we have for instance

$$m_i^*(\theta_i) \in \arg \max_{\tilde{m}_i \in \mathcal{M}_i} \underset{\theta_{-i}}{E} \left(t_i \left(\tilde{m}_i, \hat{m}_{-i}^*(\tilde{m}_i, m_{-i}^*(\theta_{-i})) \right) - \theta_i q_i \left(\tilde{m}_i, \hat{m}_{-i}^*(\tilde{m}_i, m_{-i}^*(\theta_{-i})) \right) |\theta_i \right).$$

The proof of a Revelation Principle will now proceed in two steps. In the first one, we replace the mechanism $(g(\cdot), \mathcal{M})$ by another mechanism $(\tilde{g}(\cdot), \mathcal{M})$ which is not manipulable by the principal. In the second step, we replace $(\tilde{g}(\cdot), \mathcal{M})$ by a direct, truthful and still non-manipulable mechanism $(\bar{g}(\cdot), \Theta^n)$.

Step 1: Consider the new mechanism $(\tilde{g}(\cdot), \mathcal{M})$ defined as:

$$\tilde{t}_i(m_i, m_{-i}) = t_i(m_i, \hat{m}_i^*(m_i, m_{-i})) \text{ and } \tilde{q}_i(m_i, m_{-i}) = q_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i})) \text{ for } i = 1, ..., n. \tag{A.2}$$

Lemma 5 *$(\tilde{g}(\cdot), \mathcal{M})$ is not manipulable by the principal, i.e., $\hat{m}_{-i}^*(m) = m \quad \forall m \in \mathcal{M}$ given that $\tilde{g}(\cdot)$ is offered.*

Proof: Fix any $m = (m_1, ..., m_n) \in \mathcal{M}$. By (5), we have:

$$\tilde{S}(q_1(m_1, \hat{m}^*_{-1}(m)), .., q_n(m_n, \hat{m}^*_{-n}(m))) - \sum_{i=1}^{n} t_i(m_i, \hat{m}^*_{-i}(m))$$

$$\geq \tilde{S}(q_1(m_1, \tilde{m}_{-1}), ..., q_n(m_n, \tilde{m}_{-n})) - \sum_{i=1}^{n} t_i(m_i, \tilde{m}_{-i}) \quad \forall (\tilde{m}_{-1}, ..., \tilde{m}_{-n}) \in \mathcal{M}^n_{-i}.$$

In particular, $\forall m' = (m_i, m'_{-i}) \in \mathcal{M}^n$ we get:

$$\tilde{S}(q_1(m_1, \hat{m}^*_{-1}(m)), .., q_n(m_n, \hat{m}^*_{-n}(m))) - \sum_{i=1}^{n} t_i(m_i, \hat{m}^*_{-i}(m))$$

$$\geq \tilde{S}(q_1(m_1, \hat{m}^*_{-1}(m_1, m'_{-1})), ..., q_n(m_n, \hat{m}^*_{-n}(m_n, m'_{-n}))) - \sum_{i=1}^{n} t_i(m_i, \hat{m}^*_{-i}(m_i, m'_{-i})). \quad \text{(A.3)}$$

Then, using the definition of $\tilde{g}(\cdot)$ given in (A.2), (A.3) ensures that $\forall (m'_{-1}, ..m'_{-n}) \in \mathcal{M}^n_{-i}$:

$$\tilde{S}(\tilde{q}(m)) - \sum_{i=1}^{n} \tilde{t}_i(m) \geq \tilde{S}(\tilde{q}_1(m_1, m'_{-1}), .., \tilde{q}_n(m_n, m'_n)) - \sum_{i=1}^{n} \tilde{t}_i(m_i, m'_{-i}). \quad \text{(A.4)}$$

Given that $\tilde{g}(\cdot)$ is played, the best manipulation made by the principal is $\hat{m}^*_{-i}(m) = m$ for all m. $\tilde{g}(\cdot)$ is not manipulable by the principal. ■

It is straightforward to check that the new mechanism $\tilde{g}(\cdot)$ still induces an equilibrium strategy vector $m^*(\theta) = (m^*_1(\theta_1), ..., m^*_n(\theta_n))$ for the agents. Indeed, $m^*(\cdot)$ satisfies by definition the following Bayesian-Nash constraints:

$$m^*_i(\theta_i) \in \arg\max_{m_i} \underset{\theta_{-i}}{E} \left(t_i(m_i, \hat{m}^*_{-i}(m_i, m^*_{-i}(\theta_{-i}))) - \theta_i q_i(m_i, \hat{m}^*_{-i}(m_i, m^*_{-i}(\theta_{-i}))) | \theta_i \right)$$

which can be rewritten as:

$$m^*_i(\theta_i) \in \arg\max_{m_i} \underset{\theta_{-i}}{E} \left(\tilde{t}_i(m_i, m^*_{-i}(\theta_{-i})) - \theta_i \tilde{q}_i(m_i, m^*_{-i}(\theta_{-i})) | \theta_i \right). \quad \text{(A.5)}$$

Hence, $m^*(\cdot)$ is a Bayesian-Nash equilibrium of the new mechanism $\tilde{g}(\cdot)$.

Step 2: Consider now the direct revelation mechanism $(\bar{g}(\cdot), \Theta^n)$ defined as:

$$\bar{t}_i(\theta) = \tilde{t}_i(m^*(\theta)) \text{ and } \bar{q}_i(\theta) = \tilde{q}_i(m^*(\theta)) \quad \text{for } i = 1, ..., n. \quad \text{(A.6)}$$

Lemma 6 $\bar{g}(\cdot)$ *is truthful in Bayesian incentive compatibility and not manipulable.*

Proof: First consider the non-manipulability of the mechanism $\bar{g}(\cdot)$. From (A.4), we get:

$$\tilde{S}\left(\bar{q}(\theta)\right) - \sum_{i=1}^{n} \bar{t}_i(\theta) \geq$$

$$\tilde{S}\left(\tilde{q}_1(m_1^*(\theta_1), m'_{-1}), .., \tilde{q}_n(m_n^*(\theta_n), m'_n)\right) - \sum_{i=1}^{n} \tilde{t}_i(m_i^*(\theta_i), m'_{-i}) \quad \forall m'_{-i} \in \mathcal{M}_{-i}. \tag{A.7}$$

Taking $m'_{-i} = m^*_{-i}(\theta'_{-i})$, (A.7) becomes

$$\tilde{S}\left(\bar{q}(\theta)\right) - \sum_{i=1}^{n} \bar{t}_i(\theta) \geq$$

$$\tilde{S}(\bar{q}_1(\theta_1, \theta'_{-1}), .., \bar{q}_n(\theta_n, \theta'_{-n}))) - \sum_{i=1}^{n} \bar{t}_i(\theta_i, \theta'_{-i}) \quad \forall(\theta'_{-1}, ..., \theta'_{-n}) \in \Theta^{n(n-1)}. \tag{A.8}$$

Hence, $\bar{g}(\cdot)$ is non-manipulable.

Turning to (A.5), it is immediate to check that the agents' Bayesian incentive constraints can be written as:

$$\theta_i \in \arg\max_{\hat{\theta}_i} \underset{\theta_{-i}}{E} \left(\bar{t}_i(\hat{\theta}_i, \theta_{-i}) - \theta_i \bar{q}_i(\hat{\theta}_i, \theta_{-i}) | \theta_i\right). \tag{A.9}$$

■

• **Proof of Proposition 2:** Let us define

$$\tilde{U}_i(\hat{\theta}_i, \theta_i) = \underset{\theta_{-i}}{E} \left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i}) | \theta_i\right) - H_i(\hat{\theta}_i).$$

$\tilde{U}_i(\hat{\theta}_i, \cdot)$ is differentiable for all $\hat{\theta}_i$. Without loss of generality we can restrict attention to quantity schedules that are bounded above by $\bar{q}$ large enough. Therefore there exists an integrable function $b(\theta_i)$ such that

$$\left|\frac{\partial \tilde{U}_i}{\partial \theta_i}(\hat{\theta}_i, \theta_i)\right| = \left|\underset{\theta_{-i}}{E} \left(q_i(\hat{\theta}_i, \theta_{-i}) - (S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\Big|\theta_i\right)\right| \leq b(\theta_i),$$

for all $\hat{\theta}_i$ and almost all θ_i. We can now apply Theorem 2 in Milgrom and Segal (2002, p. 586) to ensure that

$$U_i(\theta_i) = U_i(\bar{\theta}) + \int_{\theta_i}^{\bar{\theta}} \underset{\theta_{-i}}{E} \left(q_i(x, \theta_{-i}) - (S(q_i(x, \theta_{-i})) - x q_i(x, \theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|x)}{\tilde{f}(\theta_{-i}|x)}\Big|x\right) dx.$$

Therefore, we obtain:

$$\underset{\theta_i}{E}(U_i(\theta_i)) = U_i(\bar{\theta})$$

$$+ \int_{\underline{\theta}}^{\bar{\theta}} f(\theta_i) \left(\int_{\theta_i}^{\bar{\theta}} \underset{\theta_{-i}}{E} \left(q_i(x, \theta_{-i}) - (S(q_i(x, \theta_{-i})) - x q_i(x, \theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|x)}{\tilde{f}(\theta_{-i}|x)}\Big|x\right) dx\right) d\theta_i.$$

Integrating by parts yields

$$\underset{\theta_i}{E}(U_i(\theta_i)) = U_i(\bar{\theta}) + \underset{\theta}{E}\left(\frac{F(\theta_i)}{f(\theta_i)}\left(q_i(\theta) - (S(q_i(\theta)) - \theta_i q_i(\theta))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right)\right). \quad \text{(A.10)}$$

First, let us suppose that (13) is binding only at $\theta_i = \bar{\theta}$. Of course minimizing the agents' information rent requires to set $U_i(\bar{\theta}) = 0$ when the right-hand side in (16) is negative; something that will be checked later. Inserting (A.10) into the principal's objective function and optimizing pointwise yields (18).

Monotonicity conditions: Assumption 1 and strict concavity of $S(\cdot)$ immediately imply that $\frac{\partial q^{SB}}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) \geq 0$ and $\frac{\partial q^{SB}}{\partial \theta_i}(\theta_i, \theta_{-i}) < 0$.

Monotonicity of $U^{SB}(\theta_i)$: This monotonicity is ensured whenever the following sufficient condition holds

$$q^{SB}(\theta_i, \theta_{-i}) \geq \left(S(q^{SB}(\theta_i, \theta_{-i})) - \theta_i q^{SB}(\theta_i, \theta_{-i})\right)\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \quad \text{(A.11)}$$

since then integrating over θ_{-i} yields that the right-hand side of (16) is negative and thus $U_i(\theta_i)$ is non-increasing as supposed. Note that (A.11) holds when $\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \leq 0$. When instead $\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} > 0$, we have $q^{FB}(\theta_i) > q^{SB}(\theta_i, \theta_{-i}) > q^{BM}(\theta_i)$. Therefore, a sufficient condition for (A.11) is

$$\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \leq \frac{q^{BM}(\theta_i)}{S(q^{FB}(\theta_i)) - \theta_i q^{FB}(\theta_i)}$$

as requested in Assumption 2.

Second-order conditions: For $q^{SB}(\theta_i, \theta_{-i})$ the local second-order condition (17) becomes

$$\underset{\theta_{-i}}{E}\left(\frac{\frac{\partial q^{SB}}{\partial \theta_i}(\theta_i, \theta_{-i})}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\frac{F(\theta_i)}{f(\theta_i)}}\Big|\theta_i\right) \geq 0$$

which holds since $\frac{\partial q^{SB}}{\partial \theta_i}(\theta_i, \theta_{-i}) \leq 0$ from Assumption 1 and $1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\frac{F(\theta_i)}{f(\theta_i)} > 0$ from Assumption 2.

Global incentive compatibility: Observe that

$$\frac{\partial U^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_i) = \underset{\theta_{-i}}{E}\left((S'(q^{SB}(\hat{\theta}_i, \theta_{-i})) - \theta_i)\frac{\partial q^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i})|\theta_i\right) - \dot{H}^{SB}(\hat{\theta}_i).$$

Taking into account the first-order condition (15), we get:

$$\frac{\partial U^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_i)$$

$$= \underset{\theta_{-i}}{E}\left((S'(q^{SB}(\hat{\theta}_i, \theta_{-i})) - \theta_i)\frac{\partial q^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i})|\theta_i\right) - \underset{\theta_{-i}}{E}\left((S'(q^{SB}(\hat{\theta}_i, \theta_{-i})) - \hat{\theta}_i)\frac{\partial q^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i})|\hat{\theta}_i\right)$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial q^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i})\left(\left(\hat{\theta}_i - \theta_i + \frac{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}\right)\tilde{f}(\theta_{-i}|\theta_i) - \frac{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}{1 + \frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}}\tilde{f}(\theta_{-i}|\hat{\theta}_i)\right)d\theta_{-i}$$

$$= \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial q^{SB}}{\partial \hat{\theta}_i}(\hat{\theta}_i, \theta_{-i})\psi(\hat{\theta}_i, \theta_i, \theta_{-i})d\theta_{-i}$$

where

$$\psi(\hat{\theta}_i, \theta_i, \theta_{-i}) = \left(\hat{\theta}_i - \theta_i + \frac{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}\right)\tilde{f}(\theta_{-i}|\theta_i) - \frac{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}{1 + \frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}}\tilde{f}(\theta_{-i}|\hat{\theta}_i).$$

Observe first that $\psi(\theta_i, \theta_i, \theta_{-i}) = 0$. By the Intermediate Values Theorem, we have $\tilde{f}(\theta_{-i}|\hat{\theta}_i) - \tilde{f}(\theta_{-i}|\theta_i) = \tilde{f}_{\theta_i}(\theta_{-i}|\xi)(\hat{\theta}_i - \theta_i)$, for some $\xi \in]\theta_i, \hat{\theta}_i[$. Denote $M = \max_{(\theta_i, \theta_{-i}) \in \Theta^2} \tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)$, we have $\tilde{f}(\theta_{-i}|\hat{\theta}_i) - \tilde{f}(\theta_{-i}|\theta_i) \leq M(\hat{\theta}_i - \theta_i)$ for $\hat{\theta}_i \geq \theta_i$ and $\psi(\hat{\theta}_i, \theta_i, \theta_{-i}) \geq 0$ for $\hat{\theta}_i \geq \theta_i$ when

$$M \leq \tilde{f}(\theta_{-i}|\theta_i)\left(\frac{1 + \frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}}{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}\right). \tag{A.12}$$

Since $\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)} \geq -\frac{M}{\min_{(\hat{\theta}_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\hat{\theta}_i)}$, (A.12) holds when

$$M \leq \frac{f(\hat{\theta}_i)}{F(\hat{\theta}_i)}\frac{\tilde{f}(\theta_{-i}|\theta_i)}{1 + \frac{\tilde{f}(\theta_{-i}|\theta_i)}{\min_{(\hat{\theta}_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\hat{\theta}_i)}} \quad \forall \hat{\theta}_i \geq \theta_i, \forall \theta_{-i}.$$

A sufficient condition for this is

$$M \leq \min_{\theta_i \in \Theta} f(\theta_i)\frac{(\min_{(\theta_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\theta_i))^2}{2\max_{(\theta_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\theta_i)}$$

as requested by Assumption 2 since then

$$\min_{\theta_i \in \Theta} f(\theta_i)\frac{(\min_{(\theta_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\theta_i))^2}{2\max_{(\theta_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\theta_i)} \leq \min_{\theta_i \in \Theta} f(\theta_i)\frac{\tilde{f}(\theta_{-i}|\theta_i)}{1 + \frac{\tilde{f}(\theta_{-i}|\theta_i)}{\min_{(\hat{\theta}_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\hat{\theta}_i)}} \leq \frac{f(\hat{\theta}_i)}{F(\hat{\theta}_i)}\frac{\tilde{f}(\theta_{-i}|\theta_i)}{1 + \frac{\tilde{f}(\theta_{-i}|\theta_i)}{\min_{(\hat{\theta}_i, \theta_{-i}) \in \Theta^2} \tilde{f}(\theta_{-i}|\hat{\theta}_i)}}.$$

Turning now to the case $\hat{\theta}_i < \theta_i$. Note that we have then $\tilde{f}(\theta_{-i}|\hat{\theta}_i) - \tilde{f}(\theta_{-i}|\theta_i) \geq M(\hat{\theta}_i - \theta_i)$ for $\hat{\theta}_i \leq \theta_i$. Therefore, we get:

$$\psi(\hat{\theta}_i, \theta_i, \theta_{-i}) \leq (\hat{\theta}_i - \theta_i)\left(\tilde{f}(\theta_{-i}|\theta_i) + M\frac{\frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}}{1 + \frac{F(\hat{\theta}_i)}{f(\hat{\theta}_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\hat{\theta}_i)}{\tilde{f}(\theta_{-i}|\hat{\theta}_i)}}\right) \leq 0 \text{ for } \hat{\theta}_i < \theta_i$$

when Assumption 2 holds.

■

• **Proof of Lemma 2:** Starting from (7) and writing a first-order condition yields (21). To prove that those conditions are also locally sufficient, denote the principal's ex post profit as:

$$V(\hat{\theta},\theta) = \tilde{S}(q_1(\theta_1,\hat{\theta}_2), q_2(\hat{\theta}_1,\theta_2)) - t_1(\theta_1,\hat{\theta}_2) - t_2(\hat{\theta}_1,\theta_2).$$

We have:

$$\frac{\partial^2 V}{\partial \hat{\theta}_2^2}(\hat{\theta},\theta) = \frac{\partial^2 \tilde{S}}{\partial q_1^2}(q_1(\theta_1,\hat{\theta}_2), q_2(\hat{\theta}_1,\theta_2))\left(\frac{\partial q_1}{\partial \theta_2}(\theta_1,\hat{\theta}_2)\right)^2$$
$$+\frac{\partial \tilde{S}}{\partial q_1}(q_1(\theta_1,\hat{\theta}_2), q_2(\hat{\theta}_1,\theta_2))\frac{\partial^2 q_1}{\partial \theta_2^2}(\theta_1,\hat{\theta}_2) - \frac{\partial^2 t_1}{\partial \theta_2^2}(\theta_1,\hat{\theta}_2).$$

Taking into account (21) and differentiating with respect to θ_2 yields:

$$\frac{\partial^2 \tilde{S}}{\partial q_1^2}(q(\theta))\left(\frac{\partial q_1}{\partial \theta_2}\right)^2 + \frac{\partial \tilde{S}}{\partial q_1}(q(\theta))\frac{\partial^2 q_1}{\partial \theta_2^2}(\theta) - \frac{\partial^2 t_1}{\partial \theta_2^2}(\theta) = -\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q(\theta))\frac{\partial q_1}{\partial \theta_2}(\theta)\frac{\partial q_2}{\partial \theta_2}(\theta).$$

Finally, we get:

$$\frac{\partial^2 V}{\partial \hat{\theta}_2^2}(\theta,\theta) = -\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q(\theta))\frac{\partial q_1}{\partial \theta_2}(\theta)\frac{\partial q_2}{\partial \theta_2}(\theta) \leq 0$$

when Assumption 3 holds.

Direct computations yields also $\frac{\partial^2 V}{\partial \hat{\theta}_1 \partial \hat{\theta}_2}(\theta,\theta) = \frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q(\theta))\frac{\partial q_1}{\partial \theta_2}(\theta)\frac{\partial q_2}{\partial \theta_1}(\theta)$. Finally, we have:

$$\frac{\partial^2 V}{\partial \hat{\theta}_2^2}(\theta,\theta)\frac{\partial^2 V}{\partial \hat{\theta}_1^2}(\theta,\theta) - \left(\frac{\partial^2 V}{\partial \hat{\theta}_1 \partial \hat{\theta}_2}(\theta,\theta)\right)^2$$
$$= \left(\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2}(q(\theta))\right)^2 \frac{\partial q_1}{\partial \theta_2}(\theta)\frac{\partial q_2}{\partial \theta_1}(\theta)\left(\frac{\partial q_1}{\partial \theta_1}(\theta)\frac{\partial q_2}{\partial \theta_2}(\theta) - \frac{\partial q_1}{\partial \theta_2}(\theta)\frac{\partial q_2}{\partial \theta_1}(\theta)\right) \geq 0$$

which ensures concavity of the principal's problem at $\hat{\theta} = \theta$ when Assumption 4 holds. ■

• **Proof of Lemma 3:** To prove global optimality of a non-manipulable strategy, it turns out that an approach in terms of nonlinear prices helps. Define thus $T_i(q_i,\theta_i) = t_i(\theta_i,\theta_{-i})$ for θ_{-i} such that $q_i(\theta_i,\theta_{-i})$. From (7), this definition is without any ambiguity because all type θ_{-i} which corresponds to the same output $q_i(\theta_i,\theta_{-i})$ must also correspond to the same transfer $t_i(\theta_i,\theta_{-i})$ otherwise a valuable manipulation would be feasible. Assume now that $\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i,\theta_{-i}) \neq 0$ so that $q_i(\theta_i,\theta_{-i})$ is invertible with respect to θ_{-i}. Denote $\Theta_{-i}(q_i,\theta_i)$ the inverse function.

The non-manipulability constraints can be rewritten as:

$$q(\theta) = (q_1(\theta), q_2(\theta)) \in \arg\max_{q} \tilde{S}(q_1,q_2) - \sum_{i=1}^{2} T_i(q_i,\theta_i) = J(\theta,q). \quad \text{(A.13)}$$

It can be checked that:

$$\frac{\partial T_i}{\partial q_i}(q_i,\theta_i) = \frac{\partial \tilde{S}}{\partial q_i}(q_i, q_{-i}(\theta_i, \Theta_{-i}(q_i,\theta_i))).$$

So that the first-order conditions for (A.13) defines $q(\theta)$. The local analysis above also proves that second-order conditions are always locally satisfied.

We turn next to the global concavity of $J(\theta, q)$. Observe that:

$$\frac{\partial^2 T_i}{\partial q_i^2}(q_i,\theta_i) = \frac{\partial^2 \tilde{S}^2}{\partial q_i^2}(q_i, q_{-i}(\theta_i, \Theta_{-i}(q_i,\theta_i)))+\frac{\partial^2 \tilde{S}^2}{\partial q_i \partial q_{-i}}(q_i, q_{-i}(\theta_i, \Theta_{-i}(q_i,\theta_i)))\frac{\frac{\partial q_{-i}}{\partial \theta_{-i}}(\theta_i, \Theta_{-i}(q_i,\theta_i))}{\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \Theta_{-i}(q_i,\theta_i))}.$$

Assume now that $\frac{\partial^2 \tilde{S}^2}{\partial q_i \partial q_{-i}}(q_i, q_{-i}) = \lambda$, then $\frac{\partial^3 \tilde{S}^2}{\partial q_i^2 \partial q_{-i}}(q_i, q_{-i}) = 0$ so that $\frac{\partial^2 \tilde{S}^2}{\partial q_i^2}(q_i, q_{-i}(\theta_i, \Theta_{-i}(q_i,\theta_i))) = \frac{\partial^2 \tilde{S}^2}{\partial q_i^2}(q_i, q_{-i})$ for any q_{-i}. From Assumption 3 and $\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) \neq 0$, we finally get:

$$\frac{\partial^2 J}{\partial q_i^2}(\theta, q) = -\frac{\partial^2 T_i}{\partial q_i^2}(q_i,\theta_i) + \frac{\partial^2 \tilde{S}^2}{\partial q_i^2}(q_i, q_{-i}) = \lambda \frac{\frac{\partial q_{-i}}{\partial \theta_{-i}}(\theta_i, \Theta_{-i}(q_i,\theta_i))}{\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \Theta_{-i}(q_i,\theta_i))} \leq 0.$$

Similarly, we have:

$$\left(\frac{\partial^2 J}{\partial q_1^2}\frac{\partial^2 J}{\partial q_2^2} - \left(\frac{\partial^2 J}{\partial q_1 \partial q_2}\right)^2\right)(\theta, q) = \lambda^2 \left(\frac{\frac{\partial q_2}{\partial \theta_2}(\theta_1, \Theta_2(q_1,\theta_1))\,\frac{\partial q_1}{\partial \theta_1}(\Theta_1(q_2,\theta_2),\theta_2)}{\frac{\partial q_1}{\partial \theta_2}(\theta_1, \Theta_2(q_1,\theta_1))\,\frac{\partial q_2}{\partial \theta_1}(\Theta_1(q_2,\theta_2),\theta_2)} - 1\right) \geq 0$$

when Assumption 4 holds. ■

• **Proof of Proposition 3:** Using the Envelope Theorem yields the expression of the derivative of agent's information rent for a given differentiable output vector $q(\theta)$:

$$\dot{U}_i(\theta_i) = -\underset{\theta_{-i}}{E}(q_i(\theta_i,\theta_{-i})|\theta_i)+\underset{\theta_{-i}}{E}\left(\left(\int_{\underline{\theta}}^{\theta_{-i}} \frac{\partial \tilde{S}}{\partial q_i}(q(\theta_i,x))\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i,x)dx - \theta_i q_i(\theta_i,\theta_{-i})\right)\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right). \tag{A.14}$$

Integrating by parts the second term of (A.14) and taking into account that $\int_{\underline{\theta}}^{\bar{\theta}} \tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)d\theta_{-i} = 0$ yields a new expression of $\dot{U}_i(\theta_i)$ as

$$\dot{U}_i(\theta_i) = -\underset{\theta_{-i}}{E}(q_i(\theta_i,\theta_{-i})|\theta_i)-\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\underline{\theta}}^{\theta_{-i}} \tilde{f}_{\theta_i}(x|\theta_i)dx\right)\left(\frac{\partial \tilde{S}}{\partial q_i}(q(\theta_i,\theta_{-i}))dx - \theta_i\right)\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i,\theta_{-i})d\theta_{-i}. \tag{A.15}$$

From (A.15), agent A_i's information rent is decreasing when Assumption 2 holds and thus (13) is binding at $\bar{\theta}$. This yields the following expression of A_i's expected rent:

$$\underset{\theta_i}{E}(U_i(\theta_i)) = \underset{\theta}{E}\left(\frac{F(\theta_i)}{f(\theta_i)}q_i(\theta)\right)$$

$$+\underset{\theta}{E}\left(\frac{F(\theta_i)}{\tilde{f}(\theta_i,\theta_{-i})}\left(\int_{\underline{\theta}}^{\theta_{-i}} \tilde{f}_{\theta_i}(x|\theta_i)dx\right)\left(\frac{\partial \tilde{S}}{\partial q_i}(q(\theta_i,\theta_{-i}))dx - \theta_i\right)\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i,\theta_{-i})\right).$$

Inserting these expected rents into the principal's objective function yields the following calculus of variations problem:

$$\max_{\{q(\cdot)\}} \int_{\Theta^2} \Phi\left(\theta, q(\theta), \nabla q(\theta)\right) d\theta$$

where admissible arcs $q(\cdot)$ are in $\mathcal{C}^1$, and

$$\Phi\left(\theta, q(\theta), \nabla q(\theta)\right) = \tilde{f}(\theta)\left(\tilde{S}(q(\theta)) - \sum_{i=1}^{2}\left(\theta_i + \frac{F(\theta_i)}{f(\theta_i)}\right) q_i(\theta)\right)$$

$$-\sum_{i=1}^{2} F(\theta_i)\left(\int_{\underline{\theta}}^{\theta_{-i}} \tilde{f}_{\theta_i}(x|\theta_i)dx\right)\left(\frac{\partial \tilde{S}}{\partial q_i}(q(\theta_i, \theta_{-i})) - \theta_i\right)\frac{\partial q_i}{\partial \theta_{-i}}(\theta).$$

Given that Assumption 2 holds, $\Phi(\theta, s, v)$ is concave in (s, v), the necessary conditions for optimality are also sufficient. The first necessary conditions are the Euler-Lagrange conditions[42] satisfied by $q^{SB}(\theta)$. They are obtained by looking at variations of the functional inside the square Θ^2 and can be written as:

$$\Phi_{q_i} = \sum_{k=1}^{2} \frac{\partial \Phi_{q_{i\theta_k}}}{\partial \theta_k} \quad \text{for } i = 1, 2. \tag{A.16}$$

Simplifying yields (26).

The second set of necessary conditions for optimality is obtained by looking at variations of the functional on the boundary Γ of Θ^2. They can be written as:

$$\int_{\Gamma} \psi(\theta)\overrightarrow{G}_i \cdot d\overrightarrow{n} = 0 \quad \text{for } i = 1, 2. \tag{A.17}$$

for any function $\psi(\theta) \in \mathcal{C}^1$ where $d\overrightarrow{n}$ is the normal outward to Γ and $\overrightarrow{G}_i = (\Phi_{q_{i\theta_1}}, \Phi_{q_{i\theta_2}})$. These conditions are obviously satisfied since terms of the form $F(\theta_i)\left(\int_{\underline{\theta}}^{\theta_{-i}} \tilde{f}_{\theta_i}(x|\theta_i)dx\right)$ are zero on the boundary. The boundary conditions (27) come from taking $\theta_i = \underline{\theta}$ into (26).

Using characteristics to approximate solutions close to the boundary surfaces: When $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} = \lambda > 0$, we can rewrite the system of partial differential equations (26) as:

$$a(\theta_1, \theta_2)\frac{\partial q_1^{SB}}{\partial \theta_1}(\theta) - a(\theta_2, \theta_1)\frac{\partial q_1^{SB}}{\partial \theta_2}(\theta) =$$

$$-\frac{\tilde{f}(\theta)}{\lambda}\left(\left(1 + \frac{F(\theta_2)}{f(\theta_2)}\frac{\tilde{f}_{\theta_2}(\theta_1|\theta_2)}{\tilde{f}(\theta_1|\theta_2)}\right)\left(\frac{\partial \tilde{S}}{\partial q_2}(q^{SB}(\theta)) - \theta_2\right) - \frac{F(\theta_2)}{f(\theta_2)}\right), \tag{A.18}$$

$$a(\theta_1, \theta_2)\frac{\partial q_2^{SB}}{\partial \theta_1}(\theta) - a(\theta_2, \theta_1)\frac{\partial q_2^{SB}}{\partial \theta_2}(\theta) =$$

[42] Gelfand and Fomin (2000, p. 153).

$$\frac{\tilde{f}(\theta)}{\lambda}\left(\left(1+\frac{F(\theta_1)}{f(\theta_1)}\frac{\tilde{f}_{\theta_1}(\theta_2|\theta_1)}{\tilde{f}(\theta_2|\theta_1)}\right)\left(\frac{\partial\tilde{S}}{\partial q_1}(q^{SB}(\theta))-\theta_1\right)-\frac{F(\theta_1)}{f(\theta_1)}\right) \quad (A.19)$$

where $a(\theta_1,\theta_2) = F(\theta_2)\left(\int_{\underline{\theta}}^{\theta_1}\tilde{f}_{\theta_2}(x|\theta_2)dx\right) \leq 0$.

Let introduce the variable $z \in \mathbb{R}$ to parameterize characteristic curves which are tangent at each point to the surfaces $q_i = q_i^{SB}(\theta)$ defined through the system (A.18)-(A.19). We set:

$$\frac{d\theta_1}{dz}(z) = a(\theta_1,\theta_2) \text{ and } \frac{d\theta_2}{dz}(z) = -a(\theta_2,\theta_1). \quad (A.20)$$

Let $Q(z) = (Q_1(z), Q_2(z)) = q^{SB}(\theta(z))$. Equations (A.18) and (A.19) define a system of differential equations such that for $i = 1, 2$:

$$\dot{Q}_i(z) = \frac{(-1)^i\tilde{f}(\theta(z))}{\lambda}\left(\left(1+\frac{F(\theta_{-i}(z))}{f(\theta_{-i}(z))}\frac{\tilde{f}_{\theta_{-i}}(\theta_i(z)|\theta_{-i}(z))}{\tilde{f}(\theta_i(z)|\theta_{-i}(z))}\right)\left(\frac{\partial\tilde{S}}{\partial q_{-i}}(Q(z)))-\theta_{-i}(z)\right)-\frac{F(\theta_{-i}(z))}{f(\theta_{-i}(z))}\right). \quad (A.21)$$

At this stage the difficulty in using the standard method of characteristics (as in John (1982) for instance) comes from the fact that the boundary conditions (27) correspond to characteristic curves. Nevertheless, with a little bit more work, we can prove existence (locally around the boundary) and provide an approximation for a solution to (26)-(27).

Let choose the initial values for $z = 0$ as

$$\theta_1(0) = \theta_2(0) = \theta_0 \in (\underline{\theta}, \bar{\theta}). \quad (A.22)$$

Since $a(\cdot)$ is continuous and satisfies a Lipschitz condition, the Uniqueness Theorem for ordinary differential equations ensures that the system (A.20) with these initial conditions has a unique solution. It can be easily checked that $\theta_1(z)$ (resp. $\theta_2(z)$) is strictly decreasing (resp. increasing). Moreover, in the (θ_1, θ_2) space the curve corresponding to the solution of (A.20)-(A.22) cannot reach the boundary $\theta_2 = \bar{\theta}$ for some z_0 such that $\theta_1(z_0) > \underline{\theta}$ because the unique solution to (A.20) such that, for some finite z_0 we have $\theta_2(z_0) = \bar{\theta}$ and $\theta_1(z_0) > \underline{\theta}$, is such that $\theta_2(z) = \bar{\theta}$ for all z by the Uniqueness Theorem for ordinary differential equations. This would contradict the initial conditions (A.22). Moreover, because $\theta_1(z)$ (resp. $\theta_2(z)$) is decreasing and thus bounded below by $\underline{\theta}$ (resp. increasing) (bounded above by $\bar{\theta}$), it has a limit when $z \to +\infty$ and this limit has to be $\underline{\theta}$ (resp. some θ_2^* such that $\theta_2^* < \bar{\theta}$). (Note that the limit is not reached. Indeed, by the Uniqueness Theorem, there exists a unique solution to (A.20) with the conditions $\theta_1(z_1) = \underline{\theta}$ and $\theta_2(z_1) = \theta_2^*$ for some $z_1 < +\infty$ and this limit is such that $\theta_1(z) = \underline{\theta}$ and $\theta_2(z) = \theta_2^*$ for all z contradicting (A.22).) Note for each initial condition θ_0, the corresponding value of θ_2^* as $\theta_2^*(\theta_0)$. This function is weakly increasing (otherwise, there would be a contradiction with the Uniqueness Theorem for differential equations), obviously continuous in θ_0 and such that first, since $\theta_2^*(\theta_0) \geq \theta_0$ we have $lim_{\theta_0\to\bar{\theta}}\theta_2^*(\theta) = \bar{\theta}$, and second $\theta_2^*(\underline{\theta}) = \underline{\theta}$. Hence, any $\theta_2^* \in \Theta$ is the limit of a schedule $\theta_2(z)$ for some initial condition θ_0.

To understand how the system $(\theta_1(z), \theta_2(z))$ behaves as $z \to +\infty$, observes that (A.20) can be approximated as:

$$\dot{\theta}_1(z) = F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)(\theta_1(z)-\underline{\theta}) \text{ and } \dot{\theta}_2(z) = -f(\underline{\theta})\left(\int_{\underline{\theta}}^{\theta_2^*} \tilde{f}_\theta(x|\underline{\theta})dx\right)(\theta_1(z)-\underline{\theta}). \quad \text{(A.23)}$$

Integrating yields:

$$\theta_1(z) = \underline{\theta}+\mu exp(F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)z) \text{ and } \theta_2(z) = \theta_2^*-\mu\frac{f(\underline{\theta})\left(\int_{\underline{\theta}}^{\theta_2^*} \tilde{f}_\theta(x|\underline{\theta})dx\right)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)}exp(F(\theta_2^*)f_\theta(\underline{\theta}|\theta_2^*)z) \quad \text{(A.24)}$$

for one constant $\mu \in \mathbb{R}$ which depends on the initial condition θ_0. Changing variables, we set $y = \mu exp(F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)z)$ so that $F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)dz = \frac{dy}{y}$. Slightly abusing notations, we get:

$$\theta_1(y) = \underline{\theta} + y, \text{ and } \theta_2(y) = \theta_2^* - \frac{f(\underline{\theta})\left(\int_{\underline{\theta}}^{\theta_2^*} \tilde{f}_\theta(x|\underline{\theta})dx\right)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)}y \quad \text{(A.25)}$$

and (A.21) becomes:

$$\dot{Q}_i(y) = \frac{(-1)^i\tilde{f}(\theta(y))}{yF(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)\lambda}\times$$

$$\left(\left(1+\frac{F(\theta_{-i}(y))}{f(\theta_{-i}(y))}\frac{\tilde{f}_{\theta_{-i}}(\theta_i(y)|\theta_{-i}(y))}{\tilde{f}(\theta_i(y)|\theta_{-i}(y))}\right)\left(\frac{\partial\tilde{S}}{\partial q_{-i}}(Q(y)))-\theta_{-i}(y)\right)-\frac{F(\theta_{-i}(y))}{f(\theta_{-i}(y))}\right) \quad \text{(A.26)}$$

with the initial data $Q(0) = q^{SB}(\underline{\theta}, \theta_2^*)$ we obtain thereby a solution $Q(y, \theta_2^*)$. These ordinary differential equations (A.26) have singularities at $y = 0$ since the numerator and denominator on the right-hand side of (A.26) are both equal to zero at that point.

However, using Lhospital's rule, the system (A.26) gives us the derivatives at 0, namely $(\dot{Q}_1(0,\theta_2^*), \dot{Q}_2(0,\theta_2^*))$, of that solutions as the solutions to (A.27)-(A.28) below:

$$\dot{Q}_1(0,\theta_2^*) = -\frac{\tilde{f}(\underline{\theta},\theta_2^*)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)\lambda}\left(\left(1+\frac{F(\theta_2^*)}{f(\theta_2^*)}\frac{\tilde{f}_{\theta_2}(\underline{\theta}|\theta_2^*)}{\tilde{f}(\underline{\theta}|\theta_2^*)}\right)\left(\lambda\dot{Q}_1(0,\theta_2^*)+S_{22}\dot{Q}_2(0,\theta_2^*)-\dot{\theta}_2(0)\right)\right.$$

$$\left.-\dot{\theta}_2(0)\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)|_{\theta_2^*}+\frac{d}{dy}\left(\frac{F(\theta_2(y))}{f(\theta_2(y))}\frac{\tilde{f}_{\theta_2}(\theta_1(y)|\theta_2(y))}{\tilde{f}(\theta_1(y)|\theta_2(y))}\right)|_{y=0}(\varphi(\underline{\theta},\theta_2^*)-\theta_2^*)\right) \quad \text{(A.27)}$$

$$\dot{Q}_2(0,\theta_2^*) = \frac{\tilde{f}(\underline{\theta},\theta_2^*)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)\lambda}\left(S_{11}\dot{Q}_1(0,\theta_2^*)+\lambda\dot{Q}_2(0,\theta_2^*)-2\dot{\theta}_1(0)\right) \quad \text{(A.28)}$$

where $S_{11} = \frac{\partial^2\tilde{S}}{\partial q_1^2}(Q(0,\theta_2^*))$, $S_{22} = \frac{\partial^2\tilde{S}}{\partial q_2^2}(Q(0,\theta_2^*))$.

This system admits a unique solution in $(\dot{Q}_1(0,\theta_2^*), \dot{Q}_2(0,\theta_2^*))$, which proves local existence, since

$$\begin{vmatrix} -1-\gamma & -\gamma\frac{S_{22}}{\lambda} \\ \epsilon\frac{S_{11}}{\lambda} & \epsilon-1 \end{vmatrix} = 2+\gamma\epsilon\left(\frac{S_{11}S_{22}}{\lambda^2}-1\right) \neq 0$$

where $\gamma = 1 + \frac{\tilde{f}(\underline{\theta},\theta_2^*)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)} < 0$ when the correlation is small enough and $\epsilon = \gamma - 1 = \frac{\tilde{f}(\underline{\theta},\theta_2^*)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)} < 0$. This defines the derivative and the local behavior of at least a solution $(Q_1(y,\theta_2^*), Q_2(y,\theta_2^*))$ as $(Q_1(y,\theta_2^*) = Q_1(0,\theta_2^*) + \dot{Q}_1(0,\theta_2^*)y, Q_2(y,\theta_2^*) = Q_2(0,\theta_2^*) + \dot{Q}_1(0,\theta_2^*)y)$.

Now, solving the system $\theta = \theta(y,\theta_2^*)$ for y small enough yields $(y,\theta_2^*) = (Y(\theta), \Theta_2^*(\theta))$. Using (A.25), we get:

$$y = \theta_1 - \underline{\theta} \text{ and } \theta_2 - \theta_2^* = \beta(\theta_2^*)(\theta_1 - \underline{\theta})$$

where $\beta(\theta_2^*) = -\frac{f(\underline{\theta})\left(\int_{\underline{\theta}}^{\theta_2^*} \tilde{f}_\theta(x|\underline{\theta})dx\right)}{F(\theta_2^*)\tilde{f}_\theta(\underline{\theta}|\theta_2^*)}$. This system can be uniquely inverted for θ_1 close enough to $\underline{\theta}$ since the derivative w.r.t. θ_2^* of the right-hand side of the second equation is non-zero for θ_2 close enough to θ_2^*. Finally, locally around $(\underline{\theta},\theta_2^*)$, we get $q^{SB}(\theta) = Q(Y(\theta), \Theta_2^*(\theta))$ for a solution to (26) such that $Q(0,\theta_2^*) = q^{SB}(\underline{\theta},\theta_2^*)$.

Finally, tedious but straightforward computations show that the derivatives $\frac{\partial q_i}{\partial \theta_1}(\theta)$ and $\frac{\partial q_i}{\partial \theta_2}(\theta)$ for $i = 1,2$ satisfy Assumptions 3 and 4 provided λ is small enough and $\frac{\partial \varphi}{\partial \theta_i}(\theta_i,\theta_{-i}) \geq |\frac{\partial \varphi}{\partial \theta_{-i}}(\theta_i,\theta_{-i})|$. ■

• **Proof of Corollary 2:** Denote $q^{FB}(\underline{\theta},\underline{\theta}) = \frac{\mu-\underline{\theta}}{1-\lambda}$ the first-best output at $(\underline{\theta},\underline{\theta})$, and the new variables $y_i(x_i,x_{-i}) = q_i^{SB}(\theta_i,\theta_{-i}) - q^{FB}(\underline{\theta},\underline{\theta})$ and $x_i = \theta_i - \underline{\theta}$. We are looking for an analytic solution to (26) in the neighborhood of $(\underline{\theta},\underline{\theta})$. Up to terms of order more than 2, (26) (for $i = 1$) can be rewritten in the neighborhood of $(\underline{\theta},\underline{\theta})$ as:

$$(1+l(\theta_1-\underline{\theta}))(\mu-q_1^{SB}+\lambda q_2^{SB}-\theta_1)-(\theta_1-\underline{\theta})-m(\theta_1-\underline{\theta})^2 = l\lambda(\theta_1-\underline{\theta})(\theta_2-\underline{\theta})\left(\frac{\partial q_2^{SB}}{\partial \theta_1} - \frac{\partial q_2^{SB}}{\partial \theta_2}\right)$$

which yields with the new variables

$$(1 + lx_1)(-y_1 + \lambda y_2 - x_1) - x_1 - mx_1^2 = l\lambda x_1 x_2 \left(\frac{\partial y_2}{\partial x_1} - \frac{\partial y_2}{\partial x_2}\right) \tag{A.29}$$

and a similar equation is obtained by permuting indices.

We look for a symmetric analytic solution of the form:

$$y_i(x_i,x_{-i}) = a_1 x_i + a_2 x_{-i} + b_1 x_i^2 + b_2 x_{-i}^2 + b_3 x_i x_{-i} + o_i(||x||^2),$$

where $o_i(||x||^2)$ $(i = 1,2)$ is of order more than 2. Inserting this expression into (A.29) and identifying the coefficients yields:

$$a_1 = -\frac{2}{1-\lambda^2}, \quad a_2 = -\frac{2\lambda}{1-\lambda^2}, \quad b_1 = \frac{l-m}{1-\lambda^2}, \quad b_2 = \frac{\lambda(l-m)}{1-\lambda^2}, \quad b_3 = -\frac{2l\lambda}{1-\lambda^2}.$$

This yields the expression of the solution in the text. Assumption 3 is easily checked. Note that $|a_1| > |a_2|$ so that Assumption 4 holds. ■

• **Proof of Proposition 4:** Let us write the principal's objective function as:

$$W(q(\cdot),t(\cdot)) = (\nu^2+\alpha)(S(q(\underline{\theta},\underline{\theta})) - 2t(\underline{\theta},\underline{\theta}))$$

$$+2(\nu(1-\nu)-\alpha)(S(q(\underline{\theta},\bar{\theta})) - t(\bar{\theta},\underline{\theta})) - t(\underline{\theta},\bar{\theta})) + ((1-\nu)^2+\alpha)(S(q(\bar{\theta},\bar{\theta})) - 2t(\bar{\theta},\bar{\theta})).$$

For α small enough, intuition suggests that the relevant Bayesian incentive constraint is that of an efficient agent and the relevant participation constraint that of an inefficient one. Those constraints can be written respectively as:

$$(\nu^2+\alpha)(t(\underline{\theta},\underline{\theta}) - \underline{\theta}q(\underline{\theta},\underline{\theta})) + (\nu(1-\nu)-\alpha)(t(\underline{\theta},\bar{\theta}) - \underline{\theta}q(\underline{\theta},\bar{\theta}))$$

$$\geq (\nu^2+\alpha)(t(\bar{\theta},\underline{\theta}) - \underline{\theta}q(\bar{\theta},\underline{\theta})) + (\nu(1-\nu)-\alpha)(t(\bar{\theta},\bar{\theta}) - \underline{\theta}q(\bar{\theta},\bar{\theta})). \tag{A.30}$$

$$(\nu(1-\nu)-\alpha)(t(\bar{\theta},\underline{\theta}) - \bar{\theta}q(\bar{\theta},\underline{\theta})) + ((1-\nu)^2+\alpha)(t(\bar{\theta},\bar{\theta}) - \bar{\theta}q(\bar{\theta},\bar{\theta})) \geq 0. \tag{A.31}$$

Neglecting the Bayesian incentive constraint of an inefficient agent and the participation constraint of an inefficient one, the principal's problem so relaxed becomes thus:

$$\max_{\{q(\cdot),t(\cdot)\}} W(q(\cdot),t(\cdot)) \text{ subject to (29), (30), (A.30) and (A.31).}$$

This set of constraints define a convex set with non-empty interior. Denoting respectively by β, γ, λ and μ the non-negative multipliers of those constraints, forming the Lagrangean and optimizing with respect to transfers yields the following Karush-Kuhn-Tucker conditions:

$$-2(\nu^2+\alpha) + \lambda(\nu^2+\alpha) - 2\beta = 0$$

$$-2(\nu(1-\nu)-\alpha) + \lambda(\nu(1-\nu)-\alpha) + 2\beta = 0,$$

$$-2(\nu(1-\nu)-\alpha) - \lambda(\nu^2+\alpha) + 2\gamma + \mu(\nu(1-\nu)-\alpha) = 0,$$

$$-2((1-\nu)^2+\alpha) - \lambda(\nu(1-\nu)-\alpha) - 2\gamma + \mu((1-\nu)^2+\alpha) = 0.$$

Solving this system yields,

$$\beta = 0, \gamma = \alpha > 0, \lambda = 2 > 0, \mu = \frac{2}{1-\nu}. \tag{A.32}$$

From which we deduce that (29) is slack and (30), (A.30) and (A.31) are all binding at the optimum. Using (30) and (A.31) binding, yields:

$$t(\bar{\theta},\bar{\theta}) - \bar{\theta}q(\bar{\theta},\bar{\theta}) = \frac{\nu(1-\nu)-\alpha}{2(1-\nu)}(S(q(\bar{\theta},\bar{\theta})) - 2\bar{\theta}q(\bar{\theta},\bar{\theta}) - (S(q(\underline{\theta},\bar{\theta})) - 2\bar{\theta}q(\underline{\theta},\bar{\theta})))$$

$$t(\bar{\theta},\underline{\theta}) - \bar{\theta}q(\underline{\theta},\bar{\theta}) = -\frac{(1-\nu)^2+\alpha}{2(1-\nu)}(S(q(\bar{\theta},\bar{\theta})) - 2\bar{\theta}q(\bar{\theta},\bar{\theta}) - (S(q(\underline{\theta},\bar{\theta})) - 2\bar{\theta}q(\underline{\theta},\bar{\theta})))$$

and the expression of the agent's expected rent given by

$$\nu U(\underline{\theta}) = \Delta\theta \left(\left(\nu^2 + \alpha\right) q(\underline{\theta}, \bar{\theta}) + \left(\nu(1-\nu) - \alpha\right) q(\bar{\theta}, \bar{\theta}) \right)$$

$$-\frac{\alpha}{1-\nu} \left(S(q(\bar{\theta}, \bar{\theta})) - 2\bar{\theta} q(\bar{\theta}, \bar{\theta}) - (S(q(\underline{\theta}, \bar{\theta})) - 2\bar{\theta} q(\underline{\theta}, \bar{\theta})) \right).$$

Inserting those values into $W(q(\cdot), t(\cdot))$ and optimizing gives the expression of the optimal outputs in the proposition.

The expression above for an efficient agent's rent already shows that, for α not too large, an efficient agent's rent is strictly positive given that outputs are so that the latter's participation constraint is slack. We have:

$$U^{SB}(\bar{\theta}) = 0 < U^{SB}(\underline{\theta}) = \Delta\theta \left(\left(\nu + \frac{\alpha}{\nu}\right) q^{SB}(\underline{\theta}, \bar{\theta}) + \left(1 - \nu - \frac{\alpha}{\nu}\right) q^{SB}(\bar{\theta}, \bar{\theta}) \right)$$

$$-\frac{\alpha}{\nu(1-\nu)} \left(S(q^{SB}(\bar{\theta}, \bar{\theta})) - 2\bar{\theta} q^{SB}(\bar{\theta}, \bar{\theta}) - (S(q^{SB}(\underline{\theta}, \bar{\theta})) - 2\bar{\theta} q^{SB}(\underline{\theta}, \bar{\theta})) \right). \quad \text{(A.33)}$$

We now check that an inefficient agent's incentive constraint is slack, at least for α not too large. This amounts to verify:

$$0 > (\nu(1-\nu) - \alpha)(t^{SB}(\underline{\theta}, \underline{\theta}) - \bar{\theta} q^{SB}(\underline{\theta}, \underline{\theta})) + ((1-\nu)^2 + \alpha)(t^{SB}(\underline{\theta}, \bar{\theta}) - \bar{\theta} q^{SB}(\underline{\theta}, \bar{\theta}))$$

but this inequality holds strictly for $\alpha = 0$ since then

$$\Delta(\nu q_0^{SB}(\underline{\theta}, \underline{\theta}) + (1-\nu) q_0^{SB}(\underline{\theta}, \bar{\theta})) > \Delta(\nu q_0^{SB}(\bar{\theta}, \underline{\theta}) + (1-\nu) q_0^{SB}(\bar{\theta}, \bar{\theta}))$$

where $S'(q_0^{SB}(\underline{\theta}, \underline{\theta})) = 2\underline{\theta}$, $S'(q_0^{SB}(\underline{\theta}, \bar{\theta})) = \underline{\theta} + \bar{\theta} + \frac{\nu}{1-\nu}\Delta\theta$, and $S'(q_0^{SB}(\bar{\theta}, \bar{\theta})) = 2\bar{\theta} + \frac{2\nu}{1-\nu}\Delta\theta$ and, by continuity, it holds also for α small enough. ■

• **Proof of Proposition 5:** Let us write the principal's objective function as:

$$W(q(\cdot), t(\cdot)) = (\nu^2 + \alpha)(2Sq(\underline{\theta}, \underline{\theta}) - 2t(\underline{\theta}, \underline{\theta}))$$

$$+2(\nu(1-\nu) - \alpha)(S(q(\underline{\theta}, \bar{\theta}) + q(\bar{\theta}, \underline{\theta})) - t(\underline{\theta}, \underline{\theta})) - t(\underline{\theta}, \bar{\theta})) + ((1-\nu)^2 + \alpha)(2Sq(\bar{\theta}, \bar{\theta}) - 2t(\bar{\theta}, \bar{\theta})).$$

For α small enough, intuition suggests again that the relevant Bayesian incentive constraint is that of an efficient agent and the relevant participation constraint that of an inefficient one which can be written still as (A.30) and (A.31). Neglecting again the Bayesian incentive constraint of an inefficient agent and the participation constraint of an inefficient one and the non-manipulability constraint (31), the principal's problem so relaxed becomes thus:

$$\max_{\{q(\cdot), t(\cdot)\}} W(q(\cdot), t(\cdot)) \text{ subject to (32) (A.30) and (A.31).}$$

This set of constraints define a convex set with non-empty interior. Optimizing with respect to transfers, Karush-Kuhn-Tucker conditions take the same form as in (A.32). From which we deduce:

$$t(\bar{\theta},\bar{\theta}) - \bar{\theta}q(\bar{\theta},\bar{\theta}) = \frac{\nu(1-\nu)-\alpha}{1-\nu}(S-\bar{\theta})(q(\bar{\theta},\bar{\theta}) - q(\bar{\theta},\underline{\theta}))$$

$$t(\bar{\theta},\underline{\theta}) - \bar{\theta}q(\underline{\theta},\bar{\theta}) = -\frac{(1-\nu)^2+\alpha}{1-\nu}(S-\bar{\theta})(q(\bar{\theta},\bar{\theta}) - q(\bar{\theta},\underline{\theta}))$$

and the expression of the agent's expected rent given by

$$\nu U(\underline{\theta}) = \Delta\theta\left(\left(\nu^2+\alpha\right)q(\bar{\theta},\underline{\theta}) + (\nu(1-\nu)-\alpha)\,q(\bar{\theta},\bar{\theta})\right)$$

$$-\frac{\alpha}{1-\nu}(S-\bar{\theta})(q(\bar{\theta},\bar{\theta}) - q(\bar{\theta},\underline{\theta})).$$

Inserting those values into $W(q(\cdot),t(\cdot))$ and optimizing taking into account the feasibility constraints that probabilities are non-negative and sum at most to one gives the expression of the optimal probabilities in the proposition. As argument similar to that in the Proof of Proposition 4 shows that, for α not too large, the efficient agent's rent is strictly positive when $q^{SB}(\bar{\theta},\bar{\theta})$ is and that an inefficient agent's incentive constraint is slack.

Consider the case $S \geq \bar{\theta} + \frac{\nu}{1-\nu}\left(\frac{1-\frac{\alpha}{\nu(1-\nu)}}{1+\alpha\frac{1+\nu}{\nu(1-\nu)^2}}\right)\Delta\theta$ so that $q^{SB}(\bar{\theta},\bar{\theta}) = \frac{1}{2}$ (the case $S < \bar{\theta} + \frac{\nu}{1-\nu}\left(\frac{1-\frac{\alpha}{\nu(1-\nu)}}{1+\alpha\frac{1+\nu}{\nu(1-\nu)^2}}\right)\Delta\theta$ is trivial). Payments in the optimal auctions can be rewritten as:

$$t^{SB}(\bar{\theta},\bar{\theta}) = \frac{\bar{\theta}}{2} + \frac{\nu(1-\nu)-\alpha}{2(1-\nu)}(S-\bar{\theta}), t^{SB}(\bar{\theta},\underline{\theta}) = -\frac{(1-\nu)^2+\alpha}{2(1-\nu)}(S-\bar{\theta}) \neq 0$$

so that even a loser pays.

Moreover, taking $t^{SB}(\underline{\theta},\underline{\theta})$ and $t^{SB}(\underline{\theta},\bar{\theta})$ so that

$$S - t^{SB}(\underline{\theta},\bar{\theta}) = \frac{S}{2} - t^{SB}(\underline{\theta},\underline{\theta})$$

so that the first inequality in (31) is an equality, we find:

$$t^{SB}(\underline{\theta},\underline{\theta}) = \frac{\underline{\theta}}{2} + U^{SB}(\underline{\theta}) - \left(1-\nu-\frac{\alpha}{\nu}\right)\left(\frac{S-\underline{\theta}}{2}\right)$$

and

$$t^{SB}(\underline{\theta},\bar{\theta}) = \underline{\theta} + U^{SB}(\underline{\theta}) + \left(\nu+\frac{\alpha}{\nu}\right)\left(\frac{S-\underline{\theta}}{2}\right)$$

where

$$U^{SB}(\underline{\theta}) = \frac{\nu(1-\nu)-\alpha}{2\nu}\Delta\theta - \frac{\alpha}{2\nu(1-\nu)}(S-\bar{\theta}).$$

With those formula, we check that the second inequality in (31) is also an equality. ■

• **Proof of Proposition 6:** The proof is straightforwardly adapted from that of Proposition 1 by replacing the Bayesian incentive compatibility concept by the dominant strategy incentive compatibility concept. We omit the details. ■

• **Proof of Proposition 7:** The simple bilateral contracts exhibited in the proposition are such that the inefficient agents' participation constraints are binding, namely $u_i(\bar{\theta}, \theta_{-i}) = 0$ for all $\theta_{-i} \in \Theta$. These contracts satisfy also incentive compatibility. Moreover, they implement the optimal bilateral quantity schedules. They thus maximize the principal's expected payoff within the set of simple bilateral contracts.

We must check that more complex bilateral mechanisms cannot achieve a greater payoff. Notice that non-manipulability and dominant strategy incentive compatibility imply that there exists functions $H_i(\cdot)$ $(i = 1, 2)$ such that

$$H_i(\theta_i) = S(q_i(\theta_i, \theta_{-i})) - \theta_i q_i(\theta_i, \theta_{-i}) - u_i(\bar{\theta}, \theta_{-i}) - \int_{\theta_i}^{\bar{\theta}} q_i(x, \theta_{-i})dx \quad \forall \theta_{-i}. \tag{A.34}$$

The principal's problem can thus be written

$$\max_{\{q(\cdot),h(\cdot)\}} \sum_{i=1}^{2} E_{\theta_i}(H_i(\theta_i))$$

subject to (A.34), $q_i(., \theta_{-i})$ decreasing and

$$u_i(\bar{\theta}, \theta_{-i}) \geq 0 \ \forall \theta_{-i} \in \Theta.$$

This last constraint is obviously binding at the optimum.

For any acceptable non-manipulable and dominant strategy mechanism which implements a quantity schedule $q_i(\theta_i, \theta_{-i})$, (A.34) implies that the principal can get the same payoff with a non-manipulable mechanism that implements the schedule $q_i(\theta_i) = q_i(\theta_i, \bar{\theta})$. The optimal such output is then $q^{BM}(\theta_i)$. Moreover, such a mechanism can be implemented with a set of simple bilateral contracts linking the principal with each agent A_i, i.e., with corresponding transfers $t_i(\theta_i) = t_i(\theta_i, \bar{\theta})$ and outputs $q_i(\theta_i) = q_i(\theta_i, \bar{\theta})$ which depend only on this agent's type. ■

• **Proof of Proposition 8:** A non-manipulable and collusion-proof mechanism that is Bayesian incentive compatible must satisfy (35) and (15). This yields:

$$E_{\theta_{-i}}\left((S'(q_{-i}(\theta_i, \theta_{-i})) - \theta_{-i}) \frac{\partial q_{-i}}{\partial \theta_i}(\theta_i, \theta_{-i}) | \theta_i\right) = 0. \tag{A.35}$$

Clearly, the second-best $q^{SB}(\theta)$ does not satisfy this condition. An output schedule such that $q_i(\theta_i, \theta_{-i}) \leq q^{FB}(\theta_i)$ and $\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) \geq 0$ can only satisfy (A.35) when

$$\frac{\partial q_i}{\partial \theta_{-i}}(\theta_i, \theta_{-i}) = 0.$$

Non-manipulable and collusion-proof mechanisms are thus necessarily simple bilateral mechanisms. The corresponding solution yields the Baron-Myerson outcome $q^{BM}(\theta_i)$. ■

• **Proof of Proposition 9:** It should be clear that offering the same contracts as in the case of public offers is clearly an optimal equilibrium strategy for the principal within the class of direct revelation mechanisms where he is a priori restricted to offer sub-mechanisms where the only report he makes to agent A_i is on the message he receives from agent A_{-i}. It follows the same steps as the proof of Proposition 2 and is therefore omitted.

The only new question to investigate is whether the principal could deviate to a larger class of mechanisms to communicate with A_i possibly the endogenous information he has on whatever private offers he makes to agent A_{-i}. Denote thus by $\mathcal{P}_i$ any arbitrary compact message space available to the principal to communicate with A_i on top of the type space available to report on A_{-i}'s type, and by $\{\tilde{t}_i(\theta_i, \theta_{-i}, p_i), \tilde{q}_i(\theta_i, \theta_{-i}, p_i)\}_{\{\hat{\theta}_i \in \Theta, \hat{\theta}_{-i} \in \Theta, p_i \in \mathcal{P}_i\}}$ a menu of extended direct mechanisms (extended because we append an extra communication possibility with A_i for the principal), upper-hemi continuous in p_i. Finally, denote by $p(\theta) = (p_1(\theta), p_2(\theta))$ an array of best-responses for the principal, a priori this is a correspondence but slightly abusing notations we will denote the same way any selection within that correspondence. Optimality of the principal's behavior at the last stage of the game requires:

$$(\theta, p(\theta)) \in \arg \max_{\{\hat{\theta} \in \Theta^2, p \in \prod_{i=1}^{2} \mathcal{P}_i\}} \sum_{i=1}^{2} S(\tilde{q}_i(\theta_i, \hat{\theta}_{-i}, p_i)) - \tilde{t}_i(\theta_i, \hat{\theta}_{-i}, p_i) \qquad \text{(A.36)}$$

where the maximum above is achieved by compactness of $\mathcal{P}_i$ and upper-hemi continuity in p_i. Let define the new mechanism $(t_i^S(\theta), q_i^S(\theta)) = (\tilde{t}_i(\theta, p_i(\theta)), \tilde{q}_i(\theta, p_i(\theta)))$. The optimality condition (A.36) can be rewritten without loss of generality as:

$$\theta \in \arg \max_{\hat{\theta} \in \Theta^2} \sum_{i=1}^{2} S(q_i^S(\theta_i, \hat{\theta}_{-i})) - t_i^S(\theta_i, \hat{\theta}_{-i}). \qquad \text{(A.37)}$$

This shows that there is no point in enlarging the set of mechanisms available to the principal. ■