Mechanism Design with Private Communication[1]

Vianney Dequiedt[2] and David Martimort[3]

This version: November 7, 2006

Abstract: This paper assumes that communication between the principal and each of his agents is *private*. First, this assumption simplifies significantly mechanisms and institutions. Second, it restores continuity with respect to the information structure but still maintains the useful role of correlation to better extract the agents' information rent. We first prove a *Revelation Principle with private communication* that characterizes the set of implementable allocations which cannot be manipulated by the principal by means of simple *non-manipulability constraints.* Equipped with this tool, we investigate optimal non-manipulable mechanisms in various environments (unrelated projects, auctions, team productions). We also demonstrate a *Taxation Principle with private communication* and draw some links between our framework and the common agency literature.

Keywords: Mechanism Design, Private Communication.

JEL Classification : D82.

[1]We thank seminar participants in Marseille LAGV 2006 and Lisbon ASSET 2006 Conferences, Northwestern, Paris-PSE, Paris-Institut Poincaré, Erasmus-Rotterdam, Toulouse, University of British Columbia, and also Olivier Compte, Philippe Jehiel, Peter Klibanoff, Frédéric Koessler, Thomas Palfrey, Mike Peters, Andy Postlewaite, Jérôme Pouyet and Jean Tirole for helpful comments and discussions. The usual disclaimer applies.

[2]INRA-GAEL, University of Grenoble, France. e-mail: dequiedt@grenoble.inra.fr

[3]TSE and EHESS, France. e-mail: martimor@cict.fr

1 Introduction

Over the last thirty years or so, the theory of mechanism design has been viewed as the most powerful tool to understand how complex organizations and institutions are shaped. By means of the Revelation Principle,[1] this theory offers a full characterization of the set of implementable allocations in contexts where information is decentralized and privately known by agents at the periphery of the organization. Once this first step of the analysis is performed and once a particular welfare criterion is specified at the outset, one can find an optimal incentive feasible allocation and look for practical mechanisms that could implement this outcome.

Although this methodology has been successful to understand auction design, regulation theory, optimal organizations of the firm, etc... it has also faced severe critiques coming from various fronts. The first line of critiques followed the works of Riordan and Sappington (1988), Crémer and McLean (1985, 1988), Johnson, Pratt and Zeckhauser (1990), D'Aspremont, Crémer and Gerard-Varet (1990), Matsushima (1991) and McAfee and Reny (1992). In various contexts, those authors have all argued that private information is costless for an organization. As long as the agents' types are correlated, a clever mechanism designer can design complex lotteries to induce costless information revelation and fully extract the agents' surplus if needed. Without correlation, privately informed agents earn instead information rents and optimal mechanisms must generally reach a genuine trade-off between rent extraction and allocative efficiency which disappears when types are (even slightly) correlated. This lack of continuity of the optimal mechanism with respect to the information structure is clearly troublesome and a significant impediment to the *"Wilson Doctrine"* which argues that mechanisms should be robust to small perturbations of the game modelling. Clearly, the received theory of mechanism design fails to pass this test.

Although related to the first critique above, the second source of scepticism on the relevance of the received theory points out that mechanisms are in practice much simpler than predicted by the theory. In real-world organizations, the scope for yardstick competition and relative performance evaluations seems quite limited. Agents hardly receive contracts which are so dependent on what their peers might claim. Multilateral contracting in complex organizations seems closer to a superposition of simple bilateral contracts between the principal and each of his agents, although how it differs has to a large extent not yet been explored theoretically.[2]

[1] Gibbard (1973) and Green and Laffont (1977) among others.

[2] Payments on financial and electricity markets depend on how much an agent wants to buy from an asset and, of course, of the equilibrium price of that asset but rarely on the whole vector of quantities requested by others as the theory would predict. Similarly, incentive payments within firms do not look like complex lotteries.

Finally, an often heard criticism of the mechanism design literature points out that communication between the principal and his agents may not be as transparent as assumed. In the canonical framework for Bayesian collective choices under asymmetric information due to Myerson (1991, Chapter 6.4), all communication between the principal and his agents is public. This facilitates the implementation of the allocation recommended by the mechanism and makes credible that the principal sticks to complex rewards and punishments.[3] The flip-side of more opaque institutions is that the principal may act opportunistically and manipulate himself the agents' messages if he finds it worth. Lack of transparency and opportunistic behavior on the principal's side go hands in hands.

The model developed below responds to those criticisms and goes towards modelling weaker institutions than currently assumed in standard mechanism design theory. To do so, we relax the assumption that communication between the principal and each of his agents is public. Instead, we assume that communication is *private.* Doing so buys us two important things. First, it simplifies significantly mechanisms and shows the major role played by nonlinear prices in such environments. Second, it restores continuity with respect to the information structure but still maintains the useful role of correlation as a means to better (but not fully) extract the agents' information rent.

• **Simplicity of mechanisms and institutions:** When communication between the principal and his agents is private, the former might have strong incentives to manipulate what he has learned from one agent to punish arbitrarily others and reap the corresponding punishments. With *private communication*, the set of incentive feasible mechanisms which cannot be manipulated by the principal is thus severely restricted. We first prove a *Revelation Principle with private communication* which characterizes this set. For a given implementation concept (Bayesian or dominant strategy) characterizing the agents' behavior there is no loss of generality in restricting the analysis to mechanisms that cannot be manipulated by the principal. This set of mechanisms that can be implemented with private communication is characterized by means of simple *non-manipulability constraints.*

Equipped with this tool, we investigate the form of optimal non-manipulable mechanisms in various environments of increasing complexity.

In the simple case where agents run different projects on behalf of the principal, the only externality between them is an informational one: Their costs are correlated. Non-manipulability constraints have then strong implications on the form of feasible contracts. To avoid manipulations, the principal makes the agent's residual claimant for the return of his own project through a *sell-out contract* whose entry fee depends on the agent's report on his type only. With such contract, the principal commits himself to be indifferent

[3] This should be contrasted with the case of moral hazard where agents are first asked to report confidentially their types to the principal who then recommends to them some actions which depend only on their own announced types. See Myerson (1982).

between all possible outputs that a given agent may produce.

To avoid manipulations by the principal, non-manipulable mechanisms must limit the informational role of what has been learned from others in determining the compensation and output of a given agent. In that context, *nonlinear prices* play a significant role and a *Taxation Principle with private communication* holds. Taking into account the non-manipulability constraints is actually equivalent to imposing that the principal proposes menus of nonlinear prices to the agents and then picks his most desired quantity vector. Complex organizations are then run by contracts which look very much like bilateral ones. Nevertheless, in Bayesian environments, the optimal mechanism still *strictly* dominates the simple superposition of bilateral contracts.

Equipped with this Taxation Principle, we develop techniques to characterize non-manipulable mechanisms. The key observation is that, under private communication, the variables available for contracting between the principal and any of the agents are not observable by others. In other words, non-manipulability constraints can also be understood as incentive constraints on the principal's side preventing him from lying on what he has learned from contracting with others. We can then use standard mechanism design techniques to derive optimal non-manipulable mechanisms in various contexts.

In the case of multi-unit auctions, there is also a negative externality between competing bidders on top of the informational one. The optimal mechanism is an all-pay auction. The buyer (principal) selects first the most efficient seller, i.e., the one who pays the highest entry fee, and then offers him a sell-out contract. Again, the principal is indifferent between all possible outputs that this winning agent could produce but now, on top of that, the principal does not want to manipulate the identity of who produces.

Finally, we consider a team production context where agents exert efforts which are perfect complements. Nonlinear contracts are now more complex: Each agent only gets a fraction of the overall return of the team activity but still pays an entry fee contingent only on what he reports on his type.

• **Continuity of the optimal mechanism:** In all those environments and even when the agents' types are correlated, there always exists a genuine trade-off between rent extraction and efficiency at the optimum. Of course, how this trade-off affects contract design depends on the level of correlation but it does so in an intuitive way. Correlation makes it easier to extract the agents' information rent. When correlation vanishes, the optimal mechanism implements an allocation that comes close to that obtained for independent types *but* without the non-manipulability constraint. Non-manipulability constraints do not bind in the limit of no correlation. With independent types, there always exists an implementation of the second-best which is non-manipulable by the principal. Continuity of the optimal mechanism with respect to the information structure is thus restored when non-manipulability constraints are taken into account. More generally, standard techniques used to perform second-best analysis in settings with independent types can be rather straightforwardly adapted to the case of correlation. In particular, a *generalized virtual cost* that takes into account the correlation of types can be defined and plays the same role as in the independent type case.

Section 2 discusses the relevant literature. Section 3 presents our general model and exposes a few polar cases of interest for the rest of the analysis. In Section 4, we develop a very simple example highlighting the role of private communication in constraining mechanisms. Section 5 proves the Revelation and Taxation Principles with private communication. Equipped with these tools, we characterize optimal mechanisms in the case of unrelated projects (Section 6), with general production externalities (Section 7), multi-unit auctions (Section 8), and teams (Section 9). Section 10 concludes and proposes alleys for further research. All proofs are relegated to an Appendix.

2 Literature Review

The strong results on the benefits of correlated information pushed forward by Riordan and Sappington (1988), Crémer and McLean (1985, 1988), Johnson, Pratt and Zeckhauser (1990), D'Aspremont, Crémer and Gerard-Varet (1990), Matsushima (1991) and McAfee and Reny (1992) 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 (Dana (1993), Demski and Sappington (1988)), or limited enforceability (Compte and Jehiel (2006)). Here instead, the benefits of using correlated information is undermined by incentive constraints on the principal's side.

A second approach points out that correlated information may not be as generic as suggested by the earlier literature. Enriching the information structure may actually lead to a significant simplification of mechanisms. Neeman (2004) argues 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.[4] Bergeman and Morris (2005) argue that modelling higher order beliefs leads to ex post implementation whereas Chung and Ely (2005) show that a maxmin principal may want to rely on dominant strategy. Although important, these approaches lead also to somewhat extreme results since Bayesian mechanisms have to be

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

given up.[5,6] Our approach stills relaxes the common knowledge requirement assumed in standard mechanism design but private communication does so in a simple and tractable way. As a result, optimal mechanisms keep much of the features found in the case of independent types and Bayesian implementation keeps some of its force.

A last approach to avoid the full surplus extraction in correlated environments consists in considering collusive behavior. Laffont and Martimort (2000) showed that mechanisms extracting entirely all the agents' surplus are not robust to horizontal collusion between the agents.[7] Key to this horizontal collusion possibility is the fact that the agents can coordinate their strategies in any grand-mechanism offered by the designer. This coordination is facilitated when communication is public. Hence, our focus on private communication points at another polar case which leaves less scope for such horizontal collusion. Finally, Gromb and Martimort (2005) propose a specific model of expertise involving both moral hazard in information gathering and adverse selection and show that private communication between the principal and each of his experts opens the possibility for some vertical collusion which is harmful for the organization.

Our characterization of non-manipulable mechanisms by means of a simple Taxation Principle is reminiscent of the common agency literature which has already forcefully stressed the role of nonlinear prices as means of describing feasible allocations.[8] This resemblance comes at no surprise. Under private communication and centralized mechanism design, 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 the principals' offers. However, and in sharp contrast, there is still a bit of commitment in the game we are analyzing here in the sense that the principal first chooses the menu of nonlinear prices available to the informed agents. In a true common agency game, informed agents would be offering mechanisms first and there would be a priori no restriction in possible deviations. Although minor a priori, this difference between our model and the common agency framework will significantly simplify 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

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

[6]Also, if the aim of study 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 types distributions from observing past performances.

[7]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.

[8]Bernheim and Whinston (1986), Stole (1991), Martimort (1992 and 2005), Mezzetti (1997), Martimort and Stole (2002, 2003, 2005), Peters (2001 and 2003). Most of the time private information is modeled on the common agent's side in this literature (an exception is Martimort and Moreira (2005)).

mechanism design.[9] Martimort (2005) discusses this point and argues that one should look for minimal departures of the centralized mechanism design framework which go towards modelling multi-contracting settings. The non-manipulability constraint modeled below can precisely be viewed as such a minimal departure. Once this step is performed, one gets also an important justification for what can be mostly viewed as an ad hoc assumption generally made under common agency: Under complete information, Bernheim and Whinston (1986) suggested indeed that principals should offer the so-called *truthful contributions* which are similar to the *"sell-out"* contracts implied by non-manipulability.

The last branch of the literature related to our work is the IO literature on bilateral contracting (Hart and Tirole (1990), O'Brien and Shaffer (1992), McAfee and Schwartz (1994), Segal (1999) and Segal and Whinston (2003) among others). Those papers analyze complete information environments with secret bilateral contracting between a principal (manufacturer) and his agents (retailers). They also focus on some form of opportunism on the principal's side coming from the fact that bilateral contracts with agents are secret. Our framework differs mostly because of our focus on asymmetric information.

3 The Model

• **Preferences and Information:** We consider an organization made of one principal (P) and n agents (A_i for $i = 1, ..., n$).[10] Agent A_i produces a good in quantity q_i on the principal's behalf. The vector of goods (resp. transfers) is denoted by $q = (q_1, ..., q_n)$ (resp. $t = (t_1, ..., t_n)$). By a standard convention, A_{-i} denotes the set of all agents except A_i and similar notations are used for all other variables. Players 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 vector of goods q (resp. transfers t) belongs to some set $\mathcal{Q} = \Pi_{i=1}^{n} \mathcal{Q}_i \subset \mathbb{R}_+^n$ (resp. $\mathcal{T} = \Pi_{i=1}^{n} \mathcal{T}_i \subset \mathbb{R}^n$).

The efficiency parameter θ_i is A_i's private information. It belongs to a set $\Theta = [\underline{\theta}, \bar{\theta}]$. A vector of types is denoted $\theta = (\theta_1, ..., \theta_n)$. Types are jointly drawn from the common knowledge non-negative and atomless density function $\tilde{f}(\theta)$ whose support is Θ^n. For future reference, we will also denote the marginal density and the corresponding

[9] The most noticeable difficulty being of course the multiplicity of equilibria.

[10] Our framework can be extended in a straightforward manner to settings with more than two agents at the cost of some notational burden.

cumulative distribution, respectively, as:[11]

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

The principal's surplus function $\tilde{S}(\cdot)$ is increasing in each of its arguments q_i and concave in q. For simplicity, we shall also assume that $\tilde{S}(\cdot)$ is symmetric.

This formulation encompasses three cases of interest to whom we shall devote more attention in the sequel, specially in the case of two agents:

- *Unrelated 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 assumed to be increasing and concave with the Inada condition $S'(0) = +\infty$ and $S(0) = 0$.
- *Perfect substitutability:* $\tilde{S}(\cdot)$ is in fact a function of the total production $q_1 + q_2$ only: $\tilde{S}(q_1, q_2) = S(q_1 + q_2)$ for some increasing and concave function $S(\cdot)$ still satisfying the above conditions.
- *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 unrelated projects, the only externality between agents is an informational one and goes through the possible correlation of their cost parameters. This correlation may help the principal to better design incentives for truthful behavior. Perfect substitutability arises instead in the context of a procurement auction for an homogenous good. Perfect complementarity occurs in a team production context.[12]

• **Mechanisms:** In standard mechanism design, messages are *public*, i.e., the reports made by A_i on his type is observed by all agents A_{-i} before the given allocation requested by the mechanism gets implemented. We focus instead on the case of *private communication.* Each agent A_i privately communicates with the principal some message m_i. Then, the principal releases a report $\hat{m}_i^k$ to any agent A_k $(k \neq i)$ before implementing the requested transfers and quantity for this agent. Agent A_i observes just the private message m_i that he sends to the principal and the report $\hat{m}_{-i}$ he receives from the principal on the messages m_{-i} the latter has himself received from all the other agents A_{-i}.[13]

A mechanism is a pair $(g(\cdot), \mathcal{M})$. The outcome function $g(\cdot) = (g_1(\cdot), ..., g_n(\cdot))$ is itself a vector of outcome functions. $g_i(\cdot)$ maps the communication space $\mathcal{M} = \Pi_{i=1}^n \mathcal{M}_i$

[11] In the case of independent types, $\tilde{f}(\theta) = \Pi_{i=1}^n f(\theta_i)$.

[12] By a quick change of set-up, perfect substitutability is also the relevant case to treat standard auctions while perfect complementarity is the relevant case to treat public good problems.

[13] Note that the principal may a priori send to two different agents different messages concerning the report he received from a third one.

into the set $\Delta(\mathcal{Q}_i \times \mathcal{T}_i)$ of (possibly random) allocations available for agent A_i. The outcome function $g_i(\cdot)$ associates to any message vector $m = (m_i, \hat{m}_{-i})$ from the joint communication space $\mathcal{M} = \mathcal{M}_i \times \mathcal{M}_{-i}$ an output $q_i(m_i, \hat{m}_{-i})$ and a transfer $t_i(m_i, \hat{m}_{-i})$ for agent A_i. When the allocations is 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.

To avoid inferences by A_i on the true report made by agents A_{-i} to the principal just by checking whether the transfer and output given to those agents are consistent with his own private report to the principal and on the message he received, we assume that the outputs and transfers (q_{-i}, t_{-i}) are not observed by A_i. For minimal departure from standard mechanism design theory, we assume that the mechanism $(g_{-i}(\cdot), \mathcal{M}_{-i})$ is observable by A_i.[14]

With private communication, the principal once informed on an agent's report might manipulate this report to extract more from others if he finds it attractive. Of course, in the background the Court of Law that can observe the private messages m sent by all agents to the principal but enforce the allocations contingent on the released messages $\hat{m}$ is corruptible and colludes with the principal. In this sense, our modelling captures a case for weak institutions where the opacity of transactions leaves scope for that gaming.

• **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, all agents simultaneously accept or refuse this mechanism. If agent A_i refuses, he gets no transfer ($t_i = 0$) and produces nothing ($q_i = 0$) so that he obtains a payoff normalized to zero. Fourth, agents privately and simultaneously send the vector of messages m to the principal. Fifth, and this is the novelty of our modelling, the principal privately reports the messages $\hat{m}_{-i}$ to A_i. Finally, the corresponding outputs and transfers for agent A_i are implemented according to the $(m_i, \hat{m}_{-i})$.

For most of the paper, our equilibrium concept is perfect Bayesian equilibrium (thereafter PBE).[15]

• **Benchmark:** Without already entering into the details of the analysis, let us consider the case of public messages. If types are correlated, a by-now standard result in the literature, is that the first-best outcome can be either achieved (with discrete types) or arbitrarily approached (with a continuum of types). In sharp contrast with what economic intuition commends, there is no trade-off between efficiency and rent extraction in such correlated environments. In the case of unrelated projects, for instance, the (symmetric)

[14]However, this assumption plays little role in the analysis which would be identical under the alternative assumption of secret offers, provided agents hold passive beliefs. Section 9 below investigates further this possibility.

[15]Except in Section 6.2 where we study also dominant strategy implementation as the implementation concept for the agents' behavior.

first-best output requested from each agent trades off the marginal benefit of production against its marginal cost, namely:

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

When types are instead independently distributed, the first-best outcome can no longer be costlessly implemented. Because asymmetric information gives information rents to the agents and those rents are viewed as costly by the principal, there is now a genuine trade-off between efficiency and rent extraction. The marginal benefit of production must be equal to the *virtual* marginal cost. With unrelated projects, the (symmetric) second-best output is therefore given by the so-called *Baron-Myerson* outcome[16] for each agent:

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

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 solution.[17]

This Baron-Myerson outcome is also obtained when the principal contracts separately with each agent on the basis of the latter's report only. This would also be the solution if the principal was a priori restricted to use bilateral contracts with each agent even in settings with correlated types. If types are correlated, the discrepancy between (1) and (2) measures then the loss when going from a multilateral contracting environment to a bilateral contracting one.

4 A Simple Example

To fix ideas and already give some preliminary insights on the general analysis that will be performed later on, let us consider a very simple example 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 gross surplus that accrues to the principal when consuming this unit is S. The seller's cost may take two values $\theta \in \{\underline{\theta}, \bar{\theta}\}$ (where $\Delta\theta = \bar{\theta} - \underline{\theta} > 0$) with respective probabilities ν and $1 - \nu$. The following conditions hold:

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

The right-hand side inequality simply means that trade is efficient with both types of seller under complete information. The left-hand side inequality instead captures the

[16]Baron and Myerson (1982).

[17]Otherwise, bunching may arise at the optimal contract. See Laffont and Martimort (2002, Chapter 3) for instance.

fact that trade is no longer efficient with the high cost seller when there is asymmetric information. The buyer makes then an optimal take-it-or-leave-it offer to the seller at a price $\underline{\theta}$. Only an efficient seller accepts this offer and trades.

Let us now suppose 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 informative on the agent's type, and more specifically,

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

Let assume that the signal σ is publicly verifiable. The price paid by the buyer for one unit of the good should in full generality be a function of the seller's report on his cost as well as the realized value of the signal. Let denote by $t(\theta, \sigma)$ this price.

Looking for transfers that would implement the first-best production decision, 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 that both types produce) can be written as:

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

It is well known from the work of Riordan and Sappington (1988) that the buyer can extract all surplus from the seller and implement the first-best outcome by properly designing the price schedule: It suffices (among many other possibilities) to set price lotteries which bind all incentive and participation constraints:

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

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

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 price scheme above can no longer be used since it is 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 the reward. To avoid those manipulations by the principal, the price must be independent of the realized signal:

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

With this *non-manipulability constraint*, we are back to the traditional screening model *without* ex post information. Given (3), it is suboptimal for the buyer to procure the good in all states of nature.

This simple example illustrates the consequences of having the principal manipulate information which, if otherwise public, would be used for screening purposes. In the sequel, information is no longer exogenously produced but is learned from contracting with another agent. Second, the non-manipulability is derived rather than assumed. Moreover, and again in sharp 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), the non-manipulability of a mechanism by the principal may require distorting both outputs and transfers. Although the kind of lotteries used above lose much of their content, they might still have some value if output is accordingly distorted.

5 Revelation and Taxation Principles with Private Communication

5.1 The Revelation Principle

Let us come back to our general model. To start the analysis, we look for a full characterization of the set of allocations that can be achieved as PBEs of the overall contracting game where the principal first offers a private communication mechanism $(g(\cdot), \mathcal{M})$ (using a priori any arbitrary communication space $\mathcal{M}$) and, second, may then manipulate the report of an agent when releasing those reports to others.

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

For a fixed mechanism $(g(\cdot), \mathcal{M})$, let us define the continuation PBEs that such mechanism induce as follows:

Definition 1 *: A continuation PBE for any arbitrary mechanism $(g(\cdot), \mathcal{M})$ is a triplet $\{m^*(\cdot), \hat{m}^*(\cdot), d\mu(\theta|m)\}$ 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 manipulation strategy $\hat{m}^*(\cdot)$*

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

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

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

- *The principal's posterior beliefs $d\mu(\theta|m)$ on the agents' types follow Bayes's rule whenever possible (i.e., when $m \in \sup m^*(\cdot)$) and are arbitrary otherwise.*

Given a mechanism $(g(\cdot), \mathcal{M})$, a continuation PBE $\{m^*(\cdot), \hat{m}^*(\cdot), d\mu(\theta|m)\}$ induces an allocation $a = g \circ \hat{m}^* \circ m^*$ which maps Θ^n on $\Delta(\mathcal{Q} \times \mathcal{T})$.

Definition 2 *: A mechanism $(g(\cdot), \mathcal{M})$ is non-manipulable if and only if $\hat{m}^*(m) = m$, for all $m \in \sup m^*(\cdot)$ at a continuation PBE.*[18]

Definition 3 *: A direct mechanism $(\bar{g}(\cdot), \Theta^2)$ is truthful if and only if $m^*(\theta) = \theta$, for all $\theta \in \Theta$ at a continuation PBE.*

We are now ready to state:

Proposition 1 *: **The Revelation Principle with Private Communication.** Any allocation $a(\cdot)$ achieved at a continuation PBE of any arbitrary mechanism $(g(\cdot), \mathcal{M})$ with private communication can also be implemented as a truthful and non-manipulable continuation PBE of a direct mechanism $(\bar{g}(\cdot), \Theta^n)$.*

The Bayesian incentive compatibility constraints describing the agents' behavior are written as usual:

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

The following *non-manipulability* constraints stipulate that the principal will not misrepresent to one agent what he has learned from another agent's report:

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

In the sequel, we analyze the impact of the non-manipulability constraint (7) on optimal mechanisms in different contexts.

[18]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 PBEs.

5.2 The Taxation Principle

Beforehand, we propose an alternative formulation of the problem which clarifies the impact of private communication and uncovers a link between our analysis and the common agency literature. We show below that non-manipulable mechanisms can equivalently be implemented through the following three-stage *modified common agency game*:

• At stage 1, the principal offers menus of nonlinear prices $\{T_i(q_i, \hat{\theta}_i)\}_{\hat{\theta}_i \in \Theta}$ which stipulate a payment for agent A_i as a function of how much he produces and which type he reports.[19]

• At stage 2, agents report simultaneously and non-cooperatively their types and thus pick schedules among the offered menus. The menu is truthful and each agent chooses truthfully the schedule corresponding to his own type.

• At stage 3, the principal chooses how much output to request from each agent.

Replacing a direct mechanism with menus of nonlinear prices is the essence of the standard Taxation Principle.[20] The specific point here is the principal's non-commitment encapsulated in the game form above. The principal optimally chooses the agents' outputs *ex post* conditionally on what he has learned from their truthful reports within the range of outputs specified by the nonlinear prices those agents respectively choose from. This aspect of the game is clearly reminiscent of the common agency literature where the player at the nexus of all contracts optimally reacts to the others' choices. However, and in sharp contrast, there is still a bit of commitment in the game we are analyzing here since the principal initially chooses the menu of nonlinear prices available to the informed agents. In common agency games, not only the players moving first would be the informed agents but there would be a priori no restriction in the deviations that they could envision. Here such restrictions are implicit in the fact that the principal already designs the available menu of possible schemes from which agents choose.

Proposition 2 *: **The Taxation Principle.***

- *Any allocation $a(\theta)$ achieved at a continuation PBE of a non-manipulable direct Bayesian mechanism $(\bar{g}(\cdot), \Theta)$ with private communication can alternatively be implemented as a continuation PBE of a modified common agency game which requires each agent A_i to choose truthfully a nonlinear price from menus $\{T_i(q_i, \hat{\theta}_i)\}_{\hat{\theta}_i \in \Theta}$ and then the principal to choose outputs.*

- *Conversely, any allocation $a(\theta)$ achieved at a continuation PBE of a modified common agency game which has agents choosing truthfully nonlinear prices from menus*

[19]To simplify exposition, we focus on the case of deterministic menus. The case where the principal proposes a menu of measures over price-output allocations can be addressed similarly.
[20]Rochet (1985).

$\{T_i(q_i, \hat{\theta}_i)\}_{\hat{\theta}_i \in \Theta}$ and then the principal choosing outputs can alternatively be implemented as a truthful continuation PBE of a non-manipulable direct Bayesian mechanism $(\bar{g}(\cdot), \Theta)$ with private communication.

Proposition 2 shows the exact nature of the non-manipulability constraint: The principal can only use personalized nonlinear prices to reward the agents. Of course, those schedules are designed in an incentive compatible way. More complex mechanisms are manipulable and thus cannot be used in any credible way by the principal. The Taxation Principle above also shows that non-manipulability does not necessarily imply bilateral contracting. Picking the outputs q after agents having chosen nonlinear prices still allows the principal to somewhat exploit informational and production externalities if any.

6 Unrelated Projects

6.1 Bayesian Implementation

To familiarize ourselves with the non-manipulability constraint, let us start with the simplest case where only two agents work on projects without any production externality. The principal's gross surplus function is separable:

$$\tilde{S}(q_1, q_2) = \sum_{i=1}^{2} S(q_i).$$

Written in terms of direct mechanisms, the non-manipulability constraint (7) yields that there exists an arbitrary function $h_i(\theta_i)$ such that:

$$S(q_i(\theta_i, \theta_{-i})) - t_i(\theta_i, \theta_{-i}) = h_i(\theta_i) \tag{8}$$

Equation (8) 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 nonlinear price which achieves this objective is a *sell-out contract*:

$$T_i(q, \theta_i) = S(q) - h_i(\theta_i). \tag{9}$$

Everything happens thus as if agent A_i had to pay upfront an amount $h_i(\theta_i)$ to have the right to produce on the principal's behalf. Then, the agent enjoys all returns $S(q)$ on the project he is running for the principal. The principal's payoff in his relationship with A_i is $h_i(\theta_i)$ which does not depend on the amount produced. Of course, fixed-fees are adapted so that participation by all types is ensured.

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

$$U_i(\theta_i) = \underset{\theta_{-i}}{E}\left(S(q_i(\theta_i, \theta_{-i})) - \theta_i q_i(\theta_i, \theta_{-i})|\theta_i\right) - h_i(\theta_i). \tag{10}$$

Individual rationality implies:

$$U_i(\theta_i) \geq 0 \qquad \forall i, \quad \forall \theta_i \in \Theta. \tag{11}$$

Bayesian incentive compatibility can be written as:

$$U_i(\theta_i) = \arg\max_{\hat{\theta}_i \in \Theta_i} \underset{\theta_{-i}}{E}\left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i})|\theta_i\right) - h_i(\hat{\theta}_i) \quad \forall i, \ \forall \theta_i \in \Theta. \tag{12}$$

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 in the expression of the incentive constraint exactly as the function $h_i(\cdot)$ in (12). This similarity makes the analysis of the set of non-manipulable incentive compatible allocations look very close to that with independent types.

Assuming differentiability of $q_i(\cdot)$,[21] simple revealed preferences arguments show that $h_i(\cdot)$ is itself differentiable. The local first-order condition for Bayesian incentive compatibility becomes thus:[22]

$$\dot{h}_i(\theta_i) = \underset{\theta_{-i}}{E}\left((S'(q_i(\theta_i, \theta_{-i})) - \theta_i)\frac{\partial q_i(\theta_i, \theta_{-i})}{\partial \theta_i}|\theta_i\right) \quad \forall i, \ \forall \theta_i \in \Theta; \tag{13}$$

Consider thus any output schedule $q_i(\cdot)$ which is monotonically decreasing in θ_i and which lies below the first-best. Then (13) shows that necessarily, the $h_i(\cdot)$ function that implements this output schedule is necessarily also decreasing in θ_i. In other words, less efficient types are requested to pay lower up-front payments. The Bayesian incentive constraint (13) captures then the trade-off faced by an agent with type θ_i. By exaggerating his type, this agent will have to pay a lower up-front payment. However, he will also produce less and enjoy 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. (13) becomes:

$$\dot{U}_i(\theta_i) = -\underset{\theta_{-i}}{E}\left(q_i(\theta_i, \theta_{-i})|\theta_i\right) + \underset{\theta_{-i}}{E}\left((S(q_i(\theta_i, \theta_{-i})) - \theta_i q_i(\theta_i, \theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right). \tag{14}$$

[21] Because conditional expectations depend on A_i's type, one cannot also derive from revealed preferences arguments that $q_i(\cdot)$ is itself monotonically decreasing in θ_i.

[22] We postpone the analysis of the global incentive compatibility constraints to the Appendix.

To better understand the right-hand side of (14), consider an agent with type θ_i willing to mimic a less efficient type $\theta_i + d\theta_i$. By doing so, this agent produces the same amount than this less efficient type at a lower marginal cost. This gives a first source of information rent to type θ_i which is worth:

$$\underset{\theta_{-i}}{E}\left(q_i(\theta_i,\theta_{-i})|\theta_i\right) d\theta_i.$$

Note that this source of rent is there whether there is correlation or not.

By mimicking this less efficient type, type θ_i affects also how the principal interprets the information contained in the other agent's report to adjust θ_i's own production. The corresponding marginal rent is the second term on the right-hand side of (14):

$$-\underset{\theta_{-i}}{E}\left((S(q_i(\theta_i,\theta_{-i})) - \theta_i q_i(\theta_i,\theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) d\theta_i.$$

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

$$-\underset{\theta_{-i}}{E}\left(\frac{\partial q_i(\theta_i,\theta_{-i})}{\partial\theta_i}|\theta_i\right) + \underset{\theta_{-i}}{E}\left((S'(q_i(\theta_i,\theta_{-i})) - \theta_i)\frac{\partial q_i(\theta_i,\theta_{-i})}{\partial\theta_i}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) \geq 0$$
$$\forall i = 1,2,\ \forall\theta_i \in \Theta. \quad (15)$$

The optimal non-manipulable mechanism corresponds to an allocation $\{(q_i(\theta), U_i(\theta_i))_{i=1,2}\}$ which 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 (11) to (15).

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

Assumption 1 *Monotonicity of the generalized virtual cost:*

$$\varphi(\theta_i,\theta_{-i}) = \theta_i + \frac{\frac{F(\theta_i)}{f(\theta_i)}}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\frac{F(\theta_i)}{f(\theta_i)}}$$

is always non-negative, strictly increasing in θ_i and decreasing in θ_{-i}.

This assumption 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 (15) but which remains a useful ingredient for it.

Assumption 1 is related to the following three other assumptions:

Assumption 2 *Weak correlation:*[23]

$$\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i) \text{ is close enough to zero, for all } \theta \in \Theta^2.$$

Assumption 3 *Monotone hazard rate property (MHRP):*

$$\frac{d}{d\theta}\left(\frac{F(\theta_i)}{f(\theta_i)}\right) > 0 \text{ for all } \theta_i \in \Theta.$$

Assumption 4 *Monotone likelihood ratio property (MLRP):*

$$\frac{\partial}{\partial\theta_{-i}}\left(\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right) \geq 0 \text{ for all } \theta \in \Theta^2.$$

This last assumption is in fact implied by Assumption 1. Assumptions 3 and 4 are standard in Incentive Theory. They help to build intuition on some of the results below.

Proposition 3 : ***Unrelated Projects.*** *Assume that Assumptions 1 and 2 both hold. 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*

$$S'(q^{SB}(\theta)) = \varphi(\theta_i, \theta_{-i}), \tag{16}$$

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

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

[23]We have also analyzed the case of a strong correlation for a model with discrete types. If correlation is strong enough, the non-manipulability constraints have less bite and the first-best can still be costlessly achieved in the limit of a very strong correlation. Results are available upon request.

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

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

As already stressed, there is a strong similarity between incentive constraints for a non-manipulable Bayesian mechanism and for independent types without the non-manipulability constraint. This similarity suggests that the trade-off between efficiency and rent extraction that occurs under independent types carries over in our context even with correlation. This intuition is confirmed by equation (16) which highlights the output distortion capturing this trade-off even with correlation.

With independent types, the right-hand sides of (2) and (16) are the same. The principal finds useless the report of one agent to better design the other's incentives. He must give up some information rent to induce information revelation anyway. Outputs are accordingly distorted downward to reduce those rents and the standard Baron-Myerson distortions follow. The important point to notice is that the optimal multilateral contract with unrelated projects and independent types can be implemented with a pair of bilateral contracts which are *de facto* non-manipulable by the principal. The non-manipulability constraint has no bite in this case.

When types are instead correlated, the agents' rent can be (almost) fully extracted in this context with a continuum of types[24] and the first-best output can be implemented at no cost. Of course, this result relies on the use of complex lotteries linking an agent's payment to what the other reports. Those schemes are manipulable and thus no longer used with private communication.

A similar logic to that of Section 4 applies here with an added twist. Indeed, in our earlier example, non-manipulability constraints imposed only a restriction on transfers since output was fixed at one unit. When output may also vary, non-manipulability constraints impose only that the principal's payoff remains constant over all possible transfer-output pairs that he offers to an agent. This still allows the principal to link also agent A_i's payment to what he learns from agent A_{-i}'s report as long as A_i's output varies accordingly. Doing so, the principal may still be able to incorporate some of the benefits of correlated information in the design of contracts. The multilateral contract signed with both agents does better than a pair of bilateral contracts.[25]

To understand the nature of the output distortions and the role of the correlation, it is useful to compare the solution found in (16) with the standard Baron-Myerson formula (2) which corresponds also to the optimal mechanism had the principal contracted separately with each agent. As already noticed, this pair of bilateral contracts is of course

[24] McAfee and Reny (1992).

[25] If we switch to the interpretation in terms of nonlinear prices, retaining control on the quantity produced by each agent allows the principal to still somewhat exploit the informational externalities.

non-manipulable since each agent's output and payment depend only on his own type. Whether communication is public or private does not matter. Let us see how those bilateral contracts affect the agents' information rent. Using (14), we observe that the second term on the right-hand side is null for a bilateral contract implementing $q^{BM}(\theta_i)$ since

$$\underset{\theta_{-i}}{E}\left(\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) = 0. \tag{17}$$

By departing from the Baron-Myerson outcome, one affects this second term and reduce the agent's information rent. Think now of the principal as using A_{-i}'s report to improve his knowledge of agent A_i's type. Suppose that the principal starts from the 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 "maximum likelihood estimator," the principal should infer how likely it is that A_i lies on his type by simply observing A_{-i}'s report. From Assumption 4 and condition (17), 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 from A_{-i}, $\theta_{-i} = \theta^*_{-i}(\theta_i)$. Nothing a priori unknown has been learned from A_{-i}'s report in that case. The only concern of the principal remains reducing the first-term on the right-hand side of (14). For this type θ_i, the optimal output is still equal to the Baron-Myerson solution.

Think now of an observation $\theta_{-i} > \theta^*_{-i}(\theta_i)$. Because Assumption 4 holds, it is much likely that the principal infers from such observation that A_i is less efficient than what it pretends to be. 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 in the Baron-Myerson solution. Instead, a signal $\theta_{-i} < \theta^*_{-i}(\theta_i)$ is more likely to confirm the agent's report if he exaggerates his type. Curbing these incentives requires increasing further the distortion beyond the Baron-Myerson solution.

Corollary 1 *: Under the assumptions of Proposition 3, the following output ranking holds*

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

where $\theta^*_{-i}(\theta_i)$ is increasing.

Remark 1: With correlated types it is no longer true that the local second-order condition given by equation (15) is always sufficient to guarantee global incentive compatibility even if the agents' utility function satisfies the Spence-Mirrlees condition. However, Assumption 2 ensures that the mechanism identified in Proposition 3 is globally incentive compatible so that our approach remains valid.[26] ■

[26]See the Appendix for details.

Remark 2: To get a simpler design of the optimal mechanism, we might impose also the following property:

Assumption 5 *Best-Predictor Property (BPP):*

$$\tilde{f}_{\theta_i}(\theta_i|\theta_i) = 0.$$

Given the report made by A_{-i}, the most likely type for A_i is this report itself. With this property, the optimal output equals the Baron-Myerson outcome only when reports are the same (i.e., $\theta_i = \theta^*_{-i}(\theta_i)$.) ■

6.2 Dominant Strategy and Bilateral Contracting

The previous section has shown that some form of multilateral contracting remains optimal even with non-manipulability. In this section, we strengthen the implementation concept and require that agents play dominant strategies in the mechanism offered by the principal. We ask then whether such strengthening makes multilateral contracts look more like a set of bilateral contracts.

Notice first that the notions of private communication and non-manipulability are independent of the implementation concept that is used to describe the agents' behavior. Our framework can thus be straightforwardly adapted to dominant strategy implementation. For any arbitrary mechanism $(g(\cdot), \mathcal{M})$, a dominant strategy continuation equilibrium is then defined as follows:

Definition 4 *: A continuation dominant strategy equilibrium is a triplet* $\{m^*(\cdot), \hat{m}^*(\cdot), d\mu(\theta|m)\}$ *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 manipulation strategy* $\hat{m}^*(\cdot)$

$$m_i^*(\theta_i) \in \arg\max_{m_i \in \mathcal{M}_i} t_i(m_i, \hat{m}^*_{-i}(m_i, m_{-i})) - \theta_i q_i(m_i, \hat{m}^*_{-i}(m_i, m_{-i})), \forall m_{-i} \in \mathcal{M}_{-i}; \tag{18}$$

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

$$\hat{m}^*(m) \in \arg\max_{(\hat{m}_{-1}, ..., \hat{m}_{-n}) \in \Pi_{i=1}^n \mathcal{M}_{-i}} \tilde{S}(q_1(m_1, \hat{m}_{-1}), ..., q_n(m_n, \hat{m}_{-n})) - \sum_{i=1}^n t_i(m_i, \hat{m}_{-i}). \tag{19}$$

- *The principal's posterior beliefs on the agents' types are derived following Bayes's rule whenever possible (i.e., when $m \in \sup m^*(\cdot)$) and are arbitrary otherwise.*

We immediately adapt our previous findings to get:

Proposition 4 *: **The Revelation Principle for Dominant Strategy Implementation with Private Communication.** Any allocation $a(\cdot)$ achieved at a dominant strategy equilibrium of any arbitrary mechanism $(g(\cdot), \mathcal{M})$ with private communication can alternatively be implemented as a truthful and non-manipulable dominant strategy equilibrium of a direct mechanism $(\bar{g}(\cdot), \Theta^2)$.*

Under dominant strategy implementation, non-manipulability for unrelated projects is still characterized by the condition:

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

The analysis of the dominant strategy incentive and participation constraints is standard. 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 received by an agent with type θ_i when the other agent reports being θ_{-i}, dominant strategy incentive compatibility amounts to the following implementability conditions:

$$q_i(\theta_i, \theta_{-i}) \text{ weakly decreasing in } \theta_i, \text{ for all } \theta_{-i},$$

and

$$u_i(\theta_i, \theta_{-i}) = u_i(\bar{\theta}, \theta_{-i}) + \int_{\theta_i}^{\bar{\theta}} q_i(u, \theta_{-i}) du. \tag{20}$$

We also strengthen the participation condition and impose ex post participation constraints:

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

Proposition 5 *Under dominant strategy implementation and ex post participation, the optimal non-manipulable mechanism can be achieved with a pair of bilateral contracts implementing the Baron-Myerson outcome for each agent, $(t_i^{BM}(\theta_i), q_i^{BM}(\theta_i))$ such that*

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

With dominant strategy and non-manipulability, informational externalities can no longer be exploited and the principal cannot do better than offering bilateral contracts.

Therefore, the Baron-Myerson outcome becomes optimal even with correlated types in such an environment.

Remark 3: Bilateral contracts are suboptimal if we do not impose non-manipulability even under dominant strategy implementation and ex post participation. In that case, when the required monotonicity conditions are satisfied, the optimal quantities are

$$S'(q_i(\theta_i, \theta_{-i})) = \theta_i + \frac{\tilde{F}(\theta_i \mid \theta_{-i})}{\tilde{f}(\theta_i \mid \theta_{-i})} \tag{21}$$

and the optimal mechanism yields a strictly higher payoff than a pair of bilateral contracts. Non-manipulability and dominant strategy implementability are clearly two different concepts with quite different implications. One restriction does not imply the other. These restrictions justify simple bilateral contracts only when taken in tandem. ■

7 Characterizing Non-Manipulability

When there are no externalities between the agents' projects, the non-manipulability constraints are separable in the agents' identities and it is straightforward to derive from the non-manipulability constraints the form of non-manipulable mechanisms. With more general surplus functions $\tilde{S}$, it is no longer possible to isolate directly the consequences of non-manipulability on agent A_i's schedules. Nevertheless, we now propose a general approach that enables us to derive second-best distortions in those more general environments. For simplicity, we still focus on the case of two agents only.

Using again the Taxation Principle derived in Proposition 2, non-manipulability constraints can generally be written as:

$$q(\theta) \in \arg\max_{q \in \mathcal{Q}} \tilde{S}(q) - \sum_{i=1}^{2} T_i(q_i, \theta_i). \tag{22}$$

This formulation is attractive since the optimality conditions above look very much like an incentive compatibility constraint on the principal's side. Keeping q_{-i} as fixed, the optimality condition satisfied by q_i is the same as that one should write to induce this principal to publicly reveal his private information q_{-i}. This remark being made, one can proceed as usual in mechanism design and characterize direct revelation mechanisms $\{t_i(\hat{q}_{-i}|\theta_i); q_i(\hat{q}_{-i}|\theta_i)\}_{\hat{q}_{-i} \in \mathcal{Q}}$ which induce truthful revelation of the piece of private information $\hat{q}_{-i}$. Of course, this parameter is not exogenously given as in standard adverse selection problem, but is derived endogenously from the equilibrium behavior. Starting then from such direct revelation mechanism, we can then use standard techniques and reconstruct a non-manipulable nonlinear price $T_i(q_i, \theta_i)$ by simply "eliminating" $\hat{q}_{-i}$ from the expressions obtained for $t_i(\hat{q}_{-i}|\theta_i)$ and $q_i(\hat{q}_{-i}|\theta_i)$.

Lemma 1 *The direct revelation mechanism* $\{t_i(\hat{q}_{-i}|\theta_i); q_i(\hat{q}_{-i}|\theta_i)\}_{\hat{q}_{-i}\in\mathcal{Q}}$ *associated to a non-manipulable nonlinear price* $T_i(q_i, \theta_i)$ *is such that:*

- $q_i(q_{-i}|\theta_i)$ *is monotonically increasing (resp. decreasing) in* q_{-i} *and thus a.e. differentiable when the agents' efforts are complements, i.e.,* $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} > 0$, *(resp. substitutes, i.e.,* $\frac{\partial^2 \tilde{S}}{\partial q_1 \partial q_2} < 0$*).*

- $t_i(q_{-i}|\theta_i)$ *is a.e. differentiable in* q_{-i} *with*

$$\frac{\partial t_i}{\partial q_{-i}}(q_{-i}|\theta_i) = \frac{\partial \tilde{S}}{\partial q_i}(q_i(q_{-i}|\theta_i), q_{-i})\frac{\partial q_i}{\partial q_{-i}}(q_{-i}|\theta_i). \tag{23}$$

- *Consider any differentiability point where* $\frac{\partial q_i}{\partial q_{-i}}(q_{-i}|\theta_i) \neq 0$ *and denote* $\tilde{q}_{-i}(q_i, \theta_i)$ *the inverse function of* $q_i(q_{-i}|\theta_i)$. *The non-manipulable nonlinear price* $T_i(q_i, \theta_i)$ *is differentiable at such point and its derivative satisfies:*

$$\frac{\partial T_i}{\partial q_i}(q_i, \theta_i) = \frac{\partial \tilde{S}}{\partial q_i}(q_i, \tilde{q}_{-i}(q_i, \theta_i)). \tag{24}$$

From (24), a simple integration yields the expression of $T_i(q_i, \theta_i)$ as:

$$T_i(q_i, \theta_i) = \int_0^{q_i} \frac{\partial \tilde{S}}{\partial x}(x, \tilde{q}_{-i}(x, \theta_i))dx - H_i(\theta_i) \tag{25}$$

where $H_i(\theta_i)$ is some arbitrary function. Of course, in equilibrium, conjectures must be correct and we should have:

$$\tilde{q}_{-i}(q_i(\theta), \theta_i) = q_{-i}(\theta) \quad \forall \theta.$$

Equation (24) is rather general and can be used to recover some important polar cases:

- *Unrelated Projects:* This case is straightforward since $\frac{\partial \tilde{S}}{\partial x}(x, \tilde{q}_{-i}(x, \theta_i)) = S'(x)$. Direct integration of (25) yields (9).

- *Perfect Substitutability:* Suppose that $\tilde{S}(q_1, q_2) = S(q_1 + q_2)$ so that the agents' outputs are perfect substitutes as in the case of a multi-unit auction with the winner taking all market shares.[27] Then, conjecturing that $\tilde{q}_{-i}(x, \theta_i) = 0$ when $x > 0$, non-manipulable nonlinear prices are again given by sell-out contracts:

$$T_i(q_i, \theta_i) = S(q_i) - H_i(\theta_i) \tag{26}$$

  where $H_i(\theta_i)$ is some arbitrary function.

[27] This case will be studied with more details in Section 8 below.

Using (25), we can also express the agent's incentive compatibility constraint as follows:

$$U_i(\theta_i) = \arg\max_{\hat{\theta}_i \in \Theta} \underset{\theta_{-i}}{E} \left(\int_0^{q_i(\hat{\theta}_i, \theta_{-i})} \frac{\partial \tilde{S}}{\partial x}(x, \tilde{q}_{-i}(x, \hat{\theta}_i))dx - \theta_i q_i(\hat{\theta}_i, \theta_{-i}) | \theta_i \right) - H_i(\hat{\theta}_i). \quad (27)$$

From this we can derive the optimal second-best distortions:

Proposition 6 *Assume that Assumptions 1, 2 and 5 hold altogether and that $\tilde{S}(\cdot)$ is strictly concave in (q_1, q_2). Then, the optimal non-manipulable mechanism entails outputs such that:*

$$\frac{\partial \tilde{S}}{\partial q_i}(q^{SB}(\theta)) = \varphi(\theta_i, \theta_{-i}), \quad (28)$$

provided $q_i^{SB}(\theta)$ is non-increasing in θ_i and non-decreasing (resp. non-increasing) in θ_{-i} if outputs are substitutes (resp. complements) as requested by Lemma 1.

This is again a generalized Baron-Myerson formula. The marginal benefit of one activity is equal to its generalized virtual cost.

Remark 4: With more than two agents, the difficulty is that each nonlinear price must screen a multidimensional vector of private information. We leave that issue for further research. ■

Remark 5: As an example, consider the following surplus function for the principal:

$$\tilde{S}(q_1, q_2) = \mu(q_1 + q_2) - \frac{q_1^2}{2} - \frac{q_2^2}{2} - \lambda(q_1 - q_2)^2$$

for some parameter $\mu > 0$ and $\lambda > 0$ so that optimal outputs remain non-negative. Using (28) above, it is straightforward to check that, in the limit of λ very large, i.e., when agents produce outputs which are almost perfect complements for the principal, both agents produce the same amount given by:

$$q^{SB}(\theta) = \mu - \frac{\varphi(\theta_i, \theta_{-i}) + \varphi(\theta_{-i}, \theta_i)}{2}. \quad (29)$$

The principal's marginal benefit of production is equal to the sum of the agents' generalized virtual costs. In Section 9 below, we will give up this limit argument and tackle directly the case of a team production process. ■

8 Multi-Unit Auctions

Auction design provides a nice area of application of our theory. The private communication hypothesis seems indeed quite relevant to study auctions organized on the internet.

In light of the recent development of such trading mechanisms, it is certainly a major objective to extend auction theory in that direction.[28] We now adapt the general framework of Section 7 to a multi-unit auction framework. Doing so raises new issues coming from the fact that the principal's objective is no longer strictly concave in (q_1, q_2). The principal's gross surplus from consuming $q = q_1 + q_2$ units of the good can be written as $S(q)$ where $S'(0) = +\infty$, $S' > 0$, $S'' < 0$ and $S(0) = 0$. [29]

Proposition 7 *: Assume that Assumptions 1, 2, 3, 4 and 5 hold altogether. Then, the optimal non-manipulable multi-unit auction mechanism entails:*

- *The most efficient agent always produces an output given by the modified "Baron-Myerson" formula*
$$S'(q^{SB}(\theta)) = \varphi(\theta_i, \theta_{-i}) \quad \text{for } \theta_{-i} \geq \theta_i \tag{30}$$
with $q^{SB}(\theta_i, \theta_i) = q^{BM}(\theta_i)$*;*

- *The optimal nonlinear price is a sell-out contract defined as*
$$T^{SB}(q, \theta_i) = S(q) - h^{SB}(\theta_i), \quad \forall q \in \ range(q^{SB}(\theta_i, \cdot)),\ \forall \theta_i \tag{31}$$
where
$$h^{SB}(\theta_i) = E_{\theta_{-i}}(S(q^{SB}(\theta_i, \theta_{-i})) - \theta_i q^{SB}(\theta_i, \theta_{-i}))|\theta_i)$$
$$- \int_{\theta_i}^{\bar{\theta}} E_{\theta_{-i}} \left(q^{SB}(x, \theta_{-i}) - (S(q^{SB}(x, \theta_{-i})) - \theta_i q^{SB}(x, \theta_{-i})) \frac{\tilde{f}_{\theta_i}(\theta_{-i}|x)}{\tilde{f}(\theta_{-i}|x)} |x \right) dx.$$

Several features of the optimal auction are worth to be stressed. First, the optimal multi-unit auction is efficient in our symmetric environment; the right to produce is given to the most efficient agent. Second, conditionally on winning, the agent produces an output which is modified to take into account what the principal learns from the losing agent's report. However, output distortions are always less than in the Baron-Myerson outcome. Indeed, the mere fact that an agent wins the auction conveys only "good news" to the principal; the other agent's cost parameter is always greater. Third, the principal offers a menu of (symmetric) nonlinear schedules which are sell-out contracts from which agents pick their most preferred choices. The agent having revealed the lowest cost parameter produces all output accordingly in this winner-takes-all context. Finally, even the losing agent pays an entry fee although he does not produce himself. The optimal mechanism is an all-pay auction.

[28] The private communication hypothesis is also a consistent way to give a more active role to the auctioneer and build a general model of "shill bidding."

[29] The Inada condition ensures that it is always optimal to induce a positive production even in the second-best environment that we consider so that the issue of finding an upper bound on the set of types who may actually produce no longer arises. For the case of unit-auction and the characterization of the reserve price in this case, see Dequiedt and Martimort (2006a).

9 Team Production

In a team context, agents provide efforts which are perfect complements in the production process. We will denote $q = \min(q_1, q_2)$ the organization's output and by $S(q)$ the principal's benefit from producing q units of output. We assume also that $S'(0) = +\infty$, $S' > 0$, $S'' < 0$ with $S(0) = 0$.[30]

As a benchmark, consider the case where types are independently distributed. Then, both agents produce the same amount and the marginal benefit of such production $q^{LS}(\theta_1, \theta_2)$ must be traded off against the sum of the agents' virtual costs of effort:

$$S'(q^{LS}(\theta)) = \sum_{i=1}^{2} \theta_i + \frac{F(\theta_i)}{f(\theta_i)}. \tag{32}$$

More generally, for a given (symmetric) output schedule $q(\cdot)$ offered to the agents, we rewrite the non-manipulability constraints (7) as:

$$\theta \in \arg \max_{(\hat{\theta}_1, \hat{\theta}_2) \in \Theta^2} S\left(\min(q_1(\theta_1, \hat{\theta}_2), q_2(\hat{\theta}_1, \theta_2))\right) - \sum_{i=1}^{2} t_i(\theta_i, \hat{\theta}_{-i}). \tag{33}$$

In this context with perfect complementarity, it is quite natural to look for incentive schemes such that both the agents' efforts and payments are non-decreasing in both types. Then, the principal has no incentives to lie to either agent in such a way that this agent produce more effort than the final output. We should have the following constraint on possible lies:

$$q_1(\theta_1, \hat{\theta}_2) = q_2(\hat{\theta}_1, \theta_2).$$

Alternatively, imposing such an equality can also be justified when both agents only observe the output of the organization. Using the *modified common agency* formulation proposed in Proposition 2, we observe that the output $q(\theta) = q_1(\theta) = q_2(\theta)$ should now solve:

$$q(\theta) \in \arg \max_{q} S(q) - \sum_{i=1}^{2} T_i(q, \theta_i). \tag{34}$$

Again, the non-separability of the non-manipulability constraint raises some technical issues. To make some progress, we slightly depart from the previous observability assumptions and consider now the case where agent A_i does not observe the nonlinear price $T_{-i}(q, \theta_{-i})$ offered by the principal to agent A_{-i}. As usual in the literature on secret contract offers, we are assuming passive beliefs out of the equilibrium path, i.e., agent A_i does not change his beliefs on A_{-i}'s scheme if he himself receives an offer different from that he expects in equilibrium.[31]

[30] The Inada condition again ensures that it is worth always contracting with both agents so that the issue of "shutting-down" the worst types do not arise.

[31] Segal (1999) for instance.

Guided by the intuition built in Section 7, the nonlinear price $T_i(q, \theta_i)$ can still be recovered from a direct revelation mechanism that would now induce the principal to reveal everything not known by A_i to this agent. A_i correctly infers in equilibrium A_{-i}'s own contract and all information unknown to A_i amount only to A_{-i}'s type.

When designing a nonlinear schedule $T_i(q, \theta_i)$ to extract the principal's endogenous information on A_{-i}, one must take into account that A_i forms conjectures on the equilibrium output $q^e(\theta_1, \theta_2)$ and the nonlinear price $T^e_{-i}(q, \theta_{-i})$ offered to A_{-i} (which is also non-observable by A_i, here). Let define $\phi_{-i}(\theta_i, q)$ such that:

$$q = q^e(\theta_i, \phi_{-i}(\theta_i, q)).$$

Of course, at equilibrium expectations are correct and we have:

$$\phi_{-i}(\theta_i, q(\theta)) = \theta_{-i} \quad \forall \theta.$$

Inducing truthtelling from the principal on A_{-i}'s type requires to use a nonlinear price $T_i(q, \theta_i)$ which solves:

$$T_i(q, \theta_i) = \int_0^q \left(S'(x) - \frac{\partial T^e_{-i}}{\partial x}(x, \phi_{-i}(\theta_i, x)) \right) dx - H(\theta_i) \tag{35}$$

where $H(\theta_i)$ is some arbitrary function.[32]

To characterize the optimal non-manipulable mechanism, we proceed as in the previous sections and find:

Proposition 8 *: When Assumptions 1 and 2 hold, there exists a non-manipulable mechanism for the team which entails:*

- *A symmetric output $q^{SB}(\theta_1, \theta_2)$ such that:*

$$S'(q^{SB}(\theta_1, \theta_2)) = \sum_{i=1}^{2} \varphi(\theta_i, \theta_{-i}), \tag{36}$$

  as long as $q^{SB}(\theta_1, \theta_2)$ is decreasing in both arguments;

- *The marginal payment to A_i is equal to his generalized virtual cost parameter*

$$\frac{\partial T_i^{SB}}{\partial q}(q^{SB}(\theta_i, \theta_{-i}), \theta_i) = \varphi(\theta_i, \theta_{-i}) \tag{37}$$

  which is non-decreasing in θ_i and non-increasing in θ_{-i}.

[32] The class of mechanisms satisfying (35) is non-empty. It contains for instance the separable nonlinear prices $T(q, \theta)$ of the form $T(q, \theta) = \frac{1}{2}S(q) - h(\theta)$, where $h(\cdot)$ is some arbitrary function.

- *Both agents always get a positive information rent except when inefficient*

$$U^{SB}(\theta_i) \geq 0 \quad (\textit{with } = 0 \textit{ at } \theta_i = \bar{\theta} \textit{ only}).$$

The logic of the argument here is very similar to that made earlier although the output distortions differ somewhat due to the specificities of the team problem. Non-manipulability constraints require that each agent's payment make him somewhat internalize the principal's objective function. Because of the team problem, each agent can only partially internalize the principal's objective and his marginal payment is only a fraction of the principal's marginal benefit of production. Equation (37) shows that the marginal reward to agent A_i decreases as A_{-i} (resp. A_i) becomes less (resp. more) efficient. As a consequence, the agents' shares of the production process reflect their relative efficiency.

In this team production framework, the output distortions necessary to reduce both agents' information rents must be compounded as it can be seen on (36) which generalizes the limiting case found on a particular example in (29). Also using (36), we observe that the optimal output converges towards the solution (32) as correlation diminishes. This confirms again that the non-manipulability constraints have no bite in the limit of no correlation.

10 Conclusion

This paper has investigated the consequences of relaxing the assumption of public communication in an otherwise standard mechanism design environment. Doing so paves the way to a tractable theory which responds to some of the most often heard criticisms addressed to the mechanism design methodology. Even in correlated information environments, considering the non-manipulability of mechanisms restores a genuine trade-off between efficiency and rent extraction which leads to a standard second-best analysis. In several environments of interest (unrelated projects, auctions, team production, more general production externalities), we analyzed this trade-off and characterized optimal non-manipulable mechanisms.

Each of the 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, 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. We conjecture that simple institutions like market mechanisms will perform extremely well if one insists on the non-manipulability of mechanisms.[33]

Non-manipulable public good mechanisms may also be attractive as a way out of the paradox that arises without this constraint if one considers on one hand the free-riding problem that arises in large populations with independent types and, on the other hand, the fact that the first-best is costlessly achieved as soon as there is a little bit of correlation among agents.[34]

The introduction of 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 the 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 what is the scope for horizontal collusion between the agents in the environments depicted in this paper. Indeed, since an agent's output and information rents still depend on what the other claims, there is scope for collusion in Bayesian environments whereas relying on dominant and non-manipulable mechanisms destroys this possibility.

In practice, the degree of transparency of communication in an organization may be intermediate between what we have assumed here and the more usual postulate of public communication. We conjecture that reputation-like arguments on the principal's side may 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

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

[33]For some preliminary steps in that direction in the case of auctions, see Dequiedt and Martimort (2006a).

[34]Dequiedt and Martimort (2006b) adapt the framework of the present paper to a public good environment.

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.

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 (2005): "Foundations for Dominant Strategy Mechanisms," *forthcoming Review of Economic Studies*.

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.

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

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

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.

Dequiedt, V. and D. Martimort (2006a): "Non-Manipulable Auctions," mimeo *in preparation* TSE and Grenoble.

Dequiedt, V. and D. Martimort (2006b): "Public Good Provision by Non-Manipulable Lindhal Prices," mimeo *in preparation* TSE and Grenoble.

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

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.

Gromb, D. and D. Martimort (2005): "Collusion and the Organization of Delegated Expertise," mimeo IDEI Toulouse.

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.

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

Laffont, J.-J. and D. Martimort (2000): "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309-342.

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

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

Martimort, D. and H. Moreira (2005): "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.

— (2005): "On the Robustness of Truthful 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.

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

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.

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.

Rochet, J-C., (1985), "The Taxation Principle and Multitime Hamilton-Jacobi Equations", *Journal of Mathematical Economics,* 14: 113-128.

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 Chicago University.

Appendix

• **Proof of Proposition 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. We make explicit the corresponding equilibrium conditions below.

• P updates his beliefs on the agents' types following Bayes' rule whenever possible, i.e, when $m \in$ supp $m^*(\cdot)$. 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 publicly reveals 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:

$$(\hat{m}^*_{-1}(m),...,\hat{m}^*_{-n}(m))$$
$$\in \arg\max_{(\hat{m}_{-1},...,\hat{m}_{-n})\in\Pi_{i=1}^n\mathcal{M}_{-i}} \tilde{S}(q_1(m_1,\hat{m}_{-1})),...,q_n(m_n,\hat{m}_n)) - \sum_{i=1}^n t_i(m_i,\hat{m}_{-i}). \tag{A.2}$$

Let us turn now to 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 general mechanism $(g(\cdot), \mathcal{M})$ by another general 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 and truthful mechanism $(\bar{g}(\cdot), \Theta)$.

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

$$\tilde{t}_i(m_i, m_{-i}) = t_i(m_i, \hat{m}_i^*(m_i, m_{-i})) \text{ and } \tilde{q}_i(m_i, m_{-i}) = q_i(m_i, \hat{m}_{-i}^*(m_i, m_{-i})) \text{ for } i = 1, ..., n. \tag{A.3}$$

Lemma 2 : *$(\tilde{g}(\cdot), \mathcal{M})$ 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 (A.2), we have:

$$\tilde{S}(q_i(m_i, \hat{m}_{-i}^*(m)), q_{-i}(m_{-i}, \hat{m}_{-(-i)}^*(m))) - t_i(m_i, \hat{m}_{-i}^*(m))$$
$$\geq \tilde{S}(q_i(m_i, \tilde{m}_{-i}), q_{-i}(m_{-i}, \hat{m}_{-(-i)}^*(m))) - t_i(m_i, \tilde{m}_{-i}) \quad \forall \tilde{m}_{-i} \in \mathcal{M}_{-i}.$$

Using the definition of $\tilde{g}(\cdot)$ given in (A.3), we get:

$$\tilde{S}(\tilde{q}(m)) - \tilde{t}_i(m) \geq \tilde{S}(q_i(m_i, \hat{m}_{-i}^*(m')), \tilde{q}_{-i}(m)) - t_i(m_i, \hat{m}_{-i}^*(m')) \quad \forall m' = (m_i, m'_{-i}) \in \mathcal{M}. \tag{A.4}$$

Then (A.4) becomes:

$$\tilde{S}(\tilde{q}(m)) - \tilde{t}_i(m) \geq \tilde{S}(\tilde{q}_i(m_i, m'_{-i}), \tilde{q}_{-i}(m_{-i}, m_{-(-i)})) - \tilde{t}_i(m_i, m'_{-i}) \quad \forall m'_{-i} \in \mathcal{M}_{-i}. \tag{A.5}$$

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:

$$m_i^*(\theta_i) \in \arg\max_{m_i} \underset{\theta_{-i}}{E} \left(\tilde{t}_i(m_i, m_{-i}^*(\theta_{-i})) - \theta_i q_i(m_i, m_{-i}^*(\theta_{-i})) | \theta_i \right). \tag{A.6}$$

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

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

$$\bar{t}_i(\theta) = \tilde{t}_i(m^*(\theta)) \text{ and } \bar{q}_i(\theta) = \tilde{q}_i(m^*(\theta)) \quad \text{for } i = 1, ..., n. \tag{A.7}$$

Lemma 3 *: $\bar{g}(\cdot)$ is truthful in Bayesian incentive compatibility and not manipulable.*

Proof: First consider the non-manipulability of the mechanism $\bar{g}(\cdot)$. From (A.5), we get:

$$\tilde{S}\left(\bar{q}(\theta)\right) - \bar{t}_i(\theta) \geq \tilde{S}\left(\tilde{q}_i(m_i^*(\theta_i), m'_{-i}), \tilde{q}_{-i}(m^*(\theta))\right) - \tilde{t}_i(m_i^*(\theta_i), m'_{-i}) \quad \forall m'_{-i} \in \mathcal{M}_{-i}. \quad \text{(A.8)}$$

Taking $m'_{-i} = m^*_{-i}(\theta'_{-i})$, (A.8) becomes

$$\tilde{S}\left(\bar{q}(\theta)\right) - \bar{t}_i(\theta) \geq \tilde{S}(\bar{q}_i(\theta_i, \theta'_{-i}), \bar{q}_{-i}(\theta)) - \bar{t}_i(\theta_i, \theta'_{-i}) \quad \forall \theta'_{-i} \in \Theta^{n-1}. \quad \text{(A.9)}$$

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

Turning to (A.6), it is immediate to check that the agents' Bayesian incentive constraints can be written as:

$$\theta_i \in \arg\max_{\hat{\theta}_i} \underset{\theta_{-i}}{E}\left(\bar{t}_i(\hat{\theta}_i, \theta_{-i}) - \theta_i \bar{q}_i(\hat{\theta}_i, \theta_{-i}) | \theta_i\right). \quad \text{(A.10)}$$

■

• **Proof of Proposition 2:** Let consider the non-manipulability constraint (7) and define a nonlinear price $T_i(\hat{q}_i, \theta_i)$ as $T_i(\hat{q}_i, \theta_i) = t_i(\theta_i, \hat{\theta}_{-i})$ for $\hat{q}_i = q_i(\theta_i, \hat{\theta}_{-i})$. (This definition is non-ambiguous since, still from (7), all transfers $t_i(\theta_i, \hat{\theta}_{-i})$ corresponding to the same output $q_i(\theta_i, \hat{\theta}_{-i})$ are the same. Note that $T_i(\cdot, \theta_i)$ is defined over the range of $q_i(\theta_i, \cdot)$ that we denote $rg(q_i(\theta_i, \cdot))$. For any $\hat{q}_i \in rg(q_i(\theta_i, \cdot))$ and $\hat{q}_{-i} \in rg(q_{-i}(\theta_{-i}, \cdot))$, the non-manipulability constraint (7) can be rewritten as:

$$\tilde{S}(q(\theta)) - \sum_{i=1}^{n} T_i(q_i(\theta), \theta_i) = arg \max_{\hat{q} \in \prod_{i=1}^{n} rg(q_i(\theta_i, \cdot))} \tilde{S}(\hat{q}) - \sum_{i=1}^{n} T_i(\hat{q}_i, \theta_i), \quad \text{(A.11)}$$

(A.11) is an optimality condition for the principal.

It is straightforward to check the agents' Bayesian incentive compatibility constraints:

$$\theta_i \in \arg\max_{\hat{\theta}_i} \underset{\theta_{-i}}{E}\left(T_i(q_i(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i) - \theta_i q_i(\hat{\theta}_i, \theta_{-i}) | \theta_i\right). \quad \text{(A.12)}$$

The modified common agency game $\{T_i(q_i, \hat{\theta}_i)\}_{\hat{\theta}_i \in \Theta}$ is thus Bayesian incentive compatible.

Conversely, consider any equilibrium quantities $q(\theta)$ of the modified common agency game and the nonlinear prices $T_i(q_i, \theta_i)$ that sustain this equilibrium. These nonlinear prices satisfy equations (A.11) and (A.12). Define a direct mechanism with transfers $t_i(\theta_i, \theta_{-i}) = T_i(q_i(\theta_i, \theta_{-i}), \theta_i)$ and outputs $q_i(\theta_i, \theta_{-i})$. Equation (A.11) implies

$$\tilde{S}(q(\theta)) - \sum_{i=1}^{n} t_i(\theta) \geq \tilde{S}(q_1(\theta_1, \hat{\theta}_{-1}), ..., q_n(\theta_n, \hat{\theta}_{-n})) - \sum_{i=1}^{n} t_i(\theta_i, \hat{\theta}_{-i}), \quad \forall (\theta_i, \hat{\theta}_{-i}) \in \Theta^n. \quad \text{(A.13)}$$

Hence, this direct mechanism is non-manipulable.

Equation (A.12) implies

$$\underset{\theta_{-i}}{E}\left(t_i(\theta_i,\theta_{-i})-\theta_i q_i(\theta_i,\theta_{-i})|\theta_i\right) \geq \underset{\theta_{-i}}{E}\left(t_i(\hat{\theta}_i,\theta_{-i})-\theta_i q_i(\hat{\theta}_i,\theta_{-i})|\theta_i\right) \quad \forall(\theta_i,\hat{\theta}_i)\in\Theta^2, \tag{A.14}$$

which ensures Bayesian incentive compatibility. ■

• **Proof of Proposition 3:** First, let us suppose that (11) is binding only at $\theta_i = \bar{\theta}$. Integrating (14) , we get

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

Therefore, we obtain:

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

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

Integrating by parts yields

$$\underset{\theta_i}{E}(U_i(\theta_i)) = U_i(\bar{\theta}) + \underset{\theta}{E}\left(\frac{F(\theta_i)}{f(\theta_i)}\left(q_i(\theta) - (S(q_i(\theta)) - \theta_i q_i(\theta))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right)\right). \tag{A.15}$$

Of course minimizing the agents' information rent requires to set $U_i(\bar{\theta}) = 0$ when the right-hand side in (14) is negative; something that will be checked later. Inserting (A.15) into the principal's objective function and optimizing pointwise yields (16).

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 $U_i(\theta_i)$: From Assumption 2 ($\tilde{f}_\theta$ is small enough), the second term on the right-hand side of (14) is small relative to the first one and $U_i(\cdot)$ is strictly decreasing.

Second-order conditions: Let us come back to condition (15). For $q^{SB}(\theta_i,\theta_{-i})$ this condition 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 obviously holds under the assumptions of Proposition 3.

Global incentive compatibility: The global incentive compatibility condition writes as:

$$U_i(\theta_i) \geq U_i(\hat{\theta}_i) + \underset{\theta_{-i}}{E}\left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i})|\theta_i\right) - \underset{\theta_{-i}}{E}\left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \hat{\theta}_i q_i(\hat{\theta}_i, \theta_{-i})|\hat{\theta}_i\right).$$

Using the first-order condition, the above constraint rewrites as:

$$\begin{array}{l} \int_{\theta_i}^{\hat{\theta}_i} \underset{\theta_{-i}}{E}\left(q_i(x, \theta_{-i}) - (S(q_i(x, \theta_{-i})) - x q_i(x, \theta_{-i}))\frac{\tilde{f}_{\theta_i}(\theta_{-i}|x)}{\tilde{f}(\theta_{-i}|x)}|x\right) dx \geq \\ \underset{\theta_{-i}}{E}\left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \theta_i q_i(\hat{\theta}_i, \theta_{-i})|\theta_i\right) - \underset{\theta_{-i}}{E}\left(S(q_i(\hat{\theta}_i, \theta_{-i})) - \hat{\theta}_i q_i(\hat{\theta}_i, \theta_{-i})|\hat{\theta}_i\right), \end{array} \quad \text{(A.16)}$$

When $q_i(\cdot)$ is the second-best schedule and for a fixed strictly positive marginal density $f(\cdot|\cdot)$, both sides of the inequality are continuous functions of the degree of correlation, where correlation is measured by the function $\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)$ and where continuity is with respect to the *supnorm*. For independent types, i.e., $\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i) = 0$, the above inequality becomes

$$\int_{\theta_i}^{\hat{\theta}_i} q^{BM}(x)dx \geq (\hat{\theta}_i - \theta_i)q^{BM}(\hat{\theta}_i), \quad \text{(A.17)}$$

which is clearly satisfied (with a strict inequality as soon as $\hat{\theta}_i \neq \theta_i$) and $q_i^{BM}(\theta_i)$ is strictly decreasing in θ_i. Moreover, under these hypothesis, the local second-order condition, which is also a continuous function of the degree of correlation, strictly holds for independent types since:

$$\frac{\partial q^{BM}}{\partial \theta_i}(\theta_i) < 0.$$

Therefore, a continuity argument shows that global incentive compatibility is satisfied for $\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)$ sufficiently small. ■

• **Proof of Proposition 4:** 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 5:** The 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 bilateral contracts.

We must check that a multilateral mechanism cannot achieve a greater payoff. Non-manipulability and dominant strategy incentive compatibility imply that there exists functions $h_i(\cdot)$ $(i = 1, 2)$ such that

$$h_i(\theta_i) = S(q_i(\theta_i, \theta_{-i})) - \theta_i q_i(\theta_i, \theta_{-i}) - u_i(\bar{\theta}, \theta_{-i}) - \int_{\theta_i}^{\bar{\theta}} q_i(x, \theta_{-i})dx \quad \forall \theta_{-i}, \quad \text{(A.18)}$$

and the program of the principal can be written

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

subject to (A.18), $q_i(.,\theta_{-i})$ 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.18) 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 bilateral contract with A_i, i.e., with transfers $t_i(\theta_i) = t_i(\theta_i,\bar{\theta})$ which depend only on the type of this agent. ■

• **Proof of Lemma 1:** The proof is standard and is thus omitted. See for instance Laffont and Martimort (2002, Chapters 3 and 9). ■

• **Proof of Proposition 6:** Using (25) for differentiable outputs, we obtain:

$$\dot{U}_i(\theta_i) = -\underset{\theta_{-i}}{E}\left(q_i(\theta_i,\theta_{-i})|\theta_i\right)$$

$$+\underset{\theta_{-i}}{E}\left(\left(\int_0^{q_i(\theta_i,\theta_{-i})} \frac{\partial \tilde{S}}{\partial x}(x,\tilde{q}_{-i}(x,\theta_i))dx - \theta_i q_i(\theta_i,\theta_{-i})\right)\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}|\theta_i\right) \quad \forall i=1,2,\ \forall \theta_i \in \Theta. \tag{A.19}$$

The rent is decreasing when Assumption 2 holds and thus (11) is binding at $\bar{\theta}$. This yields the following expression of A_i's expected rent:

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

$$-\underset{\theta}{E}\left(\left(\int_0^{q_i(\theta)} \frac{\partial \tilde{S}}{\partial x}(x,\tilde{q}_{-i}(x,\theta_i))dx - \theta_i q_i(\theta)\right)\frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right).$$

Inserting these expected rents into the principal's objective function yields the following optimization problem:

$$\max_{\{q(\cdot)\}} \underset{\theta}{E}\Bigg(\tilde{S}(q(\theta)) - \sum_{i=1}^{2}\left(\theta_i + \frac{F(\theta_i)}{f(\theta_i)}\right)q_i(\theta)$$

$$+\sum_{i=1}^{2}\left(\int_0^{q_i(\theta)} \frac{\partial \tilde{S}}{\partial x}(x,\tilde{q}_{-i}(x,\theta_i))dx - \theta_i q_i(\theta)\right)\frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\Bigg).$$

Optimizing with respect to output this strictly concave objective and taking into account that, at the solution, expectations are correct so that $\tilde{q}_{-i}(q_i(\theta), \theta_i) = q_{-i}(\theta)$ yields (28).

By the same continuity argument as previously, the global incentive compatibility conditions for the agents' incentive problem are still satisfied when Assumption 2 holds. Indeed, $q_i(\theta_i, \theta_{-i})$ is strictly non-increasing in θ_i and non-decreasing (resp. non-increasing) in θ_{-i} if outputs are substitutes (resp. complements) as requested by Lemma 1. ■

• **Proof of Proposition 7:** The first steps follow those of the Proof of Proposition 6 with the specification of the nonlinear price given in (26). The principal's optimization problem becomes:

$$\max_{\{q(\cdot)\}} \underset{\theta}{E}\Bigg(S(\sum_{i=1}^{2} q_i(\theta)) - \sum_{i=1}^{2}\left(\theta_i + \frac{F(\theta_i)}{f(\theta_i)}\right) q_i(\theta)$$

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

Agent A_i with type θ_i produces all output (i.e., $\tilde{q}_{-i}(q_i(\theta), \theta_i) = 0$) if and only if the following condition holds:

$$\theta_i + \frac{\frac{F(\theta_i)}{f(\theta_i)}}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \frac{F(\theta_i)}{f(\theta_i)}} < \theta_{-i} + \frac{\frac{F(\theta_{-i})}{f(\theta_{-i})}}{1 + \frac{\tilde{f}_{\theta_{-i}}(\theta_i|\theta_{-i})}{\tilde{f}(\theta_i|\theta_{-i})} \frac{F(\theta_{-i})}{f(\theta_{-i})}} \tag{A.20}$$

When Assumptions 4 and 5 both hold, $\theta_{-i} > \theta_i$ implies

$$\theta_i + \frac{\frac{F(\theta_i)}{f(\theta_i)}}{1 + \frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)} \frac{F(\theta_i)}{f(\theta_i)}} < \theta_i + \frac{F(\theta_i)}{f(\theta_i)} < \theta_{-i} + \frac{F(\theta_{-i})}{f(\theta_{-i})}$$

where the right-hand side inequality follows from Assumption 3. Hence, we get

$$\theta_{-i} + \frac{F(\theta_{-i})}{f(\theta_{-i})} < \theta_{-i} + \frac{\frac{F(\theta_{-i})}{f(\theta_{-i})}}{1 + \frac{\tilde{f}_{\theta_{-i}}(\theta_i|\theta_{-i})}{\tilde{f}(\theta_i|\theta_{-i})} \frac{F(\theta_{-i})}{f(\theta_{-i})}}$$

from using again Assumptions 4 and 5. Finally, (A.20) holds. The optimal auction is efficient and the optimal output allocation is given by (30). ■

• **Proof of Proposition 8:** The proof follows the same lines as before. Given that the transfer schedule satisfies (35), the agents' information rent can thus be written as:

$$U_i(\theta_i) = \max_{\hat{\theta}_i} \left\{ \underset{\theta_{-i}}{E} \left(\int_0^{q(\hat{\theta}_i, \theta_{-i})} \left(S'(x) - \frac{\partial T^e_{-i}}{\partial x}(x, \phi_{-i}(x, \hat{\theta}_i)) \right) dx - \theta_i q(\hat{\theta}_i, \theta_{-i}) | \theta_i \right) - H(\hat{\theta}_i) \right\}.$$

Using the Envelope Theorem yields:

$$\dot{U}_i(\theta_i) = -\underset{\theta_i}{E}\left(q(\theta)|\theta_{-i}\right)$$

$$+\underset{\theta_{-i}}{E}\left(\left(\int_0^{q(\theta)}\left(S'(x) - \frac{\partial T^e_{-i}}{\partial x}(x, \phi_{-i}(x,\theta_i))\right)dx - \theta_i q(\theta)\right)\frac{\tilde{f}_\theta(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\Big|\theta_i\right).$$

Under Assumption 2, the information rent of an agent A_i is decreasing with his type and the participation constraint is binding only at $\bar{\theta}$. This yields the expression of A_i's expected rent:

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

$$-\underset{\theta}{E}\left(\left(\int_0^{q(\theta)}\left(S'(x) - \frac{\partial T^e_{-i}}{\partial x}(x, \phi_{-i}(x,\theta_i))\right)dx - \theta_i q(\theta)\right)\frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_\theta(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right).$$

Inserting these expected rents into the principal's objective function yields the following optimization problem:

$$\max_{\{q(\cdot)\}}\underset{\theta}{E}\Bigg(S(q(\theta)) - \left(\sum_{i=1}^2 \theta_i + \frac{F(\theta_i)}{f(\theta_i)}\right)q(\theta)$$

$$+\sum_{i=1}^2\left(\int_0^{q(\theta)}\left(S'(x) - \frac{\partial T^e_{-i}}{\partial x}(x, \phi_{-i}(x,\theta_i))\right)dx - \theta_i q(\theta)\right)\frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\Bigg).$$

Optimizing pointwise yields:

$$\left(1 + \sum_{i=1}^2 \frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right)S'(q(\theta))$$

$$= \sum_{i=1}^2 \theta_i\left(1 + \frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\right) + \frac{F(\theta_i)}{f(\theta_i)} + \frac{F(\theta_i)}{f(\theta_i)}\frac{\tilde{f}_{\theta_i}(\theta_{-i}|\theta_i)}{\tilde{f}(\theta_{-i}|\theta_i)}\frac{\partial T_{-i}}{\partial q}(q(\theta), \theta_{-i}) \qquad \text{(A.21)}$$

where we have taken into account that expectations about the nonlinear price $T_i(q, \theta_i)$ are correct in equilibrium.

Also, the first-order condition for (34) can be written as:

$$S'(q(\theta)) = \sum_{i=1}^2 \frac{\partial T_i}{\partial q}(q(\theta), \theta_i). \qquad \text{(A.22)}$$

We are looking for a pair $\left(\frac{\partial T_1}{\partial q}(q(\theta), \theta_1), \frac{\partial T_2}{\partial q}(q(\theta), \theta_2)\right)$ which solves (A.21) and (A.22). The pair of marginal contributions given in (37) does the job.

If $q^{SB}(\cdot)$ is decreasing in θ_i (which is true for a sufficiently small degree of correlation), the second-order condition of the agent's problem holds and global incentive compatibility is ensured. ■