The Organization of Regulated Production: Complementarities, Correlation and Collusion*

Jos Jansen[†], Doh-Shin Jeon[‡], and Domenico Menicucci[§]

December 29, 2006

Abstract

We analyze the choice between vertical separation (VS) and vertical integration (VI) when two regulated firms produce complementary inputs with correlated costs and are protected by ex post break-even constraints. First, in the absence of collusion the regulator prefers VI (VS) for negative and weak positive (respectively, strong positive) correlation. Second, if the firms can collude under VS and know all costs, then VS is equivalent to VI. However, if firms collude under asymmetric information, then collusion does not affect the choice between VS and VI, since the regulator takes advantage of the transaction costs created by asymmetric information.

Key words: Regulation, Vertical Separation, Vertical Integration, Collusion, Asymmetric Information.
JEL Classification: D8, L2, L5

*We thank Patrick Bolton, Antonio Cabrales, Eric van Damme, Jean-Jacques Laffont, Jean-Charles Rochet, Joel Shapiro, Yossi Spiegel, Vincent Verouden, two anonymous referees, and the editor, David Martimort, for helpful comments. Jansen thanks Tilburg University and the Social Science Research Center Berlin (WZB) for their support. Jeon gratefully acknowledges the financial support from the Spanish Ministry of Science and Technology under BEC2003-00412 and the Ramon y Cajal grant.

†European University Institute, Italy. Jos.Jansen@EUI.eu
‡Universitat Pompeu Fabra and CREA, Spain. doh-shin.jeon@upf.edu (corresponding author)
§Università degli Studi di Firenze, Italy. domenico.menicucci@dmd.unifi.it

1 Introduction

In most regulated network industries the production of a final good requires the production of more than one complementary inputs. For example, generation and distribution can be distinguished as two distinct activities for utilities such as electricity, gas, and water. Traditionally, final goods were supplied by a regulated vertically integrated monopolist who produces both inputs. Since the 1980s several countries have introduced or are considering structural reforms of these industries.

In this setting a regulator makes a choice between vertical separation (VS, henceforth) and vertical integration (VI, henceforth) of complementary production activities. That is, either all inputs are produced by one vertically integrated monopolist, or each input is produced by an independent input supplier. A change of the industry's organization changes the incentives of the firms in the industry (Armstrong *et al.*, 1994). In this paper we study how asymmetric information about the costs of production affects the firms' incentives, and in effect the optimal organization of regulated production.[1] In particular, we address the following two questions. First, how is the choice affected by correlation of production costs? Second, how is the choice affected by collusion opportunities between vertically separated firms?

An example where this analysis could be particularly relevant is in the water industry. The supply of water involves water collection, treatment, and distribution. Similar production stages can be distinguished in the sewage system. The industry often consists of regulated, local monopolies. Although VI remains the norm in most countries, Garcia *et al.* (2006) observe 32 vertically separated utilities (i.e. 17 production, and 15 distribution utilities) in a balanced panel of 203 class A-C water utilities regulated by the Public Service Commission of Wisconsin between 1997 and 2000. Asymmetric information about labor efficiency created substantial distortions in the output levels (and consumers' surplus) of class A water utilities in California during the 1980s and 1990s (Wolak, 1994, Brocas *et al.*, 2006). Moreover, Wolak (1994) finds that his model of asymmetric information has a greater explanatory power than a related model of symmetric information. Labor efficiency is likely to be correlated across inputs. On the one hand, local conditions, such as education levels, climatic conditions, and geographic diversity, are common shocks to efficiency. On the other hand, the labor efficiency of each individual input may also be affected by idiosyncratic shocks, since the production of each particular input needs

[1] Although our leading example is the incentive regulation of network industries, we would like to emphasize that our results can also be applied to the internal organization of a multi-divisional firm which has two divisions producing complementary inputs.

different, specialized skills.

In our setting, two perfectly complementary inputs are produced. The input cost parameters (types) are correlated, and have a discrete distribution. Each agent (firm) has private information about his cost parameter(s). In addition, we assume that the principal (the regulator) should guarantee the agents a minium rate of return, i.e. break-even constraints should be satisfied ex post.

First, we characterize the optimal mechanisms in the absence of collusion for a given organizational choice. The optimal mechanisms are determined by a trade-off between efficiency and rent extraction. Under VS the optimal quantity schedule is decreasing in the total cost of production, if correlation is negative or weakly positive, while it is non-monotonic in the case of strong positive correlation. In the latter case, the rent obtained by an efficient agent is mainly determined by the quantity produced at the intermediate state in which one agent is efficient and the other is inefficient. The introduction of a large downward distortion in the quantity produced at this state of nature enables the principal to extract the information rent at a relatively low cost of efficiency.

Under VI the principal proposes a menu of contracts to a consolidated agent who has complete information on both cost parameters (Baron and Besanko, 1992, Dana, 1993, and Gilbert and Riordan, 1995). This gives a standard single-agent screening problem with a standard solution. The optimal quantity schedule is always weakly monotonic, and in particular exhibits partial pooling for intermediate and high total costs of production, if correlation is positive and very strong.

The optimal organization of production trades off two effects (Mookherjee and Tsumagari, 2004): *internalization of bidding externalities* and *loss of control*. Bidding externalities are present under VS when a higher cost reported by one agent modifies the quantity produced by the other agent, and thus the latter's rent. For negative or weak positive correlation the quantity profiles are monotonic, and the externality is negative. A cost overstatement by one agent reduces the quantity produced, and thereby reduces the rent that the other agent can obtain by overstating his cost. Under VI such a negative externality is internalized, which makes it less costly to induce truth-telling there. Consequently, VI dominates VS for negative or weakly positive cost correlation. This result extends the results found by Baron and Besanko (1992), Gilbert and Riordan (1995), Mookherjee and Tsumagari (2004) and Severinov (2005), who focus on the case of independently distributed types.

Loss of control plays a role when correlation is strongly positive. It represents the fact that the principal faces a more stringent set of incentive compatibility constraints under VI than under VS. The consolidated agent knows both cost parameters and coordinates his reports, while each agent under VS chooses his cost report independently from the other. This reduces the set of feasible mechanisms under VI to those with (weakly) monotonic quantity schedules. When correlation is positive and strong, the quantity schedule under VI exhibits partial pooling for intermediate and high aggregate costs. Therefore, the implementation of this schedule under VS does not create bidding externalities. However, the effect of loss of control is present. When correlation is positive and strong, the optimal quantity schedule under VS is non-monotonic, which is infeasible under VI. Therefore, VS dominates VI for strong and positive correlation.

After comparing VI and VS without collusion, we study how the comparison is affected by introducing collusion under VS. Drawing on the methodology developed by Laffont and Martimort (1997, 2000), we model collusion under asymmetric information between the two agents by a side-contract offered by a benevolent and *uninformed* third-party. The third-party can use side-payments to implement joint manipulations of reports. In our setting the agents have the incentive to coordinate their reports in order to internalize bidding externalities and to acquire more control.

As a benchmark, we consider collusion between vertically separated agents under complete information. Since informed colluding agents have identical incentives and opportunities as a consolidated agent, we find that VI is equivalent to VS if collusion occurs under complete information.

As a main result, we show that if collusion occurs under asymmetric information, the optimal outcome under VS without collusion can be implemented in a collusion-proof way at (approximately) no loss. This implies that collusion does not affect the choice between VI and VS as long as it occurs under asymmetric information.

On the one hand, if the quantity schedule under VS without collusion is monotonic (i.e. correlation is negative or weakly positive), then the optimal mechanism under no collusion leaves no room for collusion in the sense that, even if agents were completely informed about the costs, they cannot realize any gain from a joint manipulation of reports. Therefore, collusion has no effect on the organization of production if correlation is negative or weakly positive.

On the other hand, if the quantity schedule is non-monotonic (i.e. strongly positive correlation), then the agents can realize some gain from joint manipulation of reports if there is complete information between themselves. In other words, there is room for collusion. In particular, if one agent is efficient and the other is inefficient, the agents have an incentive to report that both are inefficient. Such a manipulation of reports increases the quantity of production, which increases the rent of the efficient agent while it does not hurt the inefficient agent. However, the fact that the manipulation increases the production of the inefficient agent creates an incentive problem within the coalition. The efficient agent's incentive to pretend to be inefficient to the third-party, who organizes the coalition, is larger in the presence of the manipulation than in its absence, since the efficient agent's rent is increasing in the quantity produced by an inefficient agent. Therefore, in order to implement the report manipulation, an efficient agent has to receive a higher rent than he would obtain in the absence of the manipulation. We show that this additional rent (the transaction cost created by asymmetric information) is larger than the gains from collusion, implying that the agents fail to collude.

Our paper contributes to two strands of literature. First, we contribute to the literature studying the choice between VS and VI when two firms produce complementary products (Baron and Besanko, 1992, Gilbert and Riordan, 1995, Mookherjee and Tsumagari, 2004 and Severinov, 2005).[2] We extend their result that VI dominates VS obtained in the case of independent types to the case of negative and weak positive correlation. By contrast, we show that VS dominates VI in the case of strong positive correlation.[3] Dana (1993) compares integration and separation when agents have correlated types and are protected by ex post break even constraints in a two-type setting. His setting differs from ours in two ways. In Dana, the produced goods are final goods and are sold in independent markets, i.e. the products are neither substitutes nor complements. Despite this technological difference, also he finds that integration dominates separation if and only if correlation is negative or weakly positive. Moreover, Dana does not consider collusion under separation.

Second, we contribute to the literature studying mechanism design when agents collude under asymmetric information, pioneered by Laffont and Martimort (1997, 2000). In particular, our paper extends Laffont and Martimort (2000) by considering ex post break-even constraints. In addition, Laffont and Martimort limit their analysis to the two polar cases of weak positive correlation and almost perfect positive correlation. Whereas full rent-extraction emerges without ex post break-even constraints in the absence of collusion (Crémer and McLean, 1988), Laffont and Martimort show that collusion prevents the principal from achieving the first-best outcome. However, in their optimal collusion-proof mechanism, asymmetric information does not generate any transaction cost except in the

[2] Another related paper is Dequiedt and Martimort (2004), in which one of the input suppliers can learn the other's private information at a cost. The paper studies when the buyer wants to induce the suppliers to share information and coordinate, as opposed to contracting at arm's length.

[3] Furthermore, Mookherjee and Tsumagari (2004) and Severinov (2005) consider the case in which two inputs are substitutes as well and show that VS dominates VI in this case.

limit case of almost perfect correlation. We show that when the agents are protected by ex post break-even constraints, collusion is irrelevant since the principal optimally exploits the transaction costs among colluding firms.

The result that the principal can achieve her payoff without collusion in a collusion-proof way is also obtained by Pouyet (2002), Jeon (2005), Jeon and Menicucci (2005) and Che and Kim (2006a). The first three papers follow the methodology of Laffont and Martimort by explicitly characterizing the set of collusion-proof mechanisms in some specific settings.[4] Che and Kim (2006a) use an elegant and constructive approach by presenting an optimal mechanism which is also collusion-proof in a wide range of settings. However, they do not allow for ex post break-even constraints and, in correlated environments, they need to assume that there are at least three agents. As a consequence, their approach seems to have limited usefulness in our setting under strong positive correlation, in which the optimal collusion-proof mechanism can be found only by explicitly characterizing the coalition incentive constraints under asymmetric information, unlike in Che and Kim. Finally, we focus on collusion in terms of manipulation of reports, as Laffont and Martimort (1997, 2000) do, and do not consider collusion in terms of coordination of acceptance or rejection of the principal's mechanism as some recent papers study (Dequiedt 2004, Quesada 2004, Pavlov 2004, and Che and Kim 2006b).

The paper is organized as follows. In Section 2, we present the basic model of VS without collusion. In Section 3, we study the optimal mechanisms under VS, and under VI, and we compare the two. In Section 4, we study the effect of collusion under VS. Section 5 concludes the paper. The proofs are relegated to the Appendix.

2 The model of vertical separation

This section describes the model of vertical separation (VS). We consider a regulator (the principal) P and two regulated firms (agents) A^1 and A^2 producing complementary goods. The production process consists of two stages, an intermediary and a final stage. In stage one, A^1 produces quantity $q \geq 0$ of an intermediate good which is used as input by A^2 to produce in the final stage, using a one-to-one production technology, quantity q of a final

[4] Pouyet (2002) extends Laffont and Martimort (2000) to the case in which two agents with correlated types produce perfect substitutes but does not consider ex post break-even constraints. In Jeon (2005) and Jeon and Menicucci (2005), the agents' types are independently distributed. The first paper extends the adverse selection model of Laffont and Martimort (1997) by adding moral hazard (effort). The second considers a setting of monopolistic screening in which colluding buyers can engage in arbitrage.

good. The principal chooses the output level q, and the transfers t^1 and t^2 she gives to A^1 and A^2, respectively, to induce them to produce.

Firm A^i, for $i \in \{1,2\}$, has a constant marginal cost θ^i and therefore $\theta^i q$ is his production cost. Each firm observes privately his marginal cost. We suppose that θ^1 and θ^2 are drawn from the common support $\Theta \equiv \{\theta_L, \theta_H\}$; let $\Delta \equiv \theta_H - \theta_L > 0$. We say that an agent with cost θ_L is an L-type, and an agent with cost θ_H is an H-type. The joint distribution of (θ^1, θ^2) is symmetric, and is common knowledge. For each state of nature $(\theta^1, \theta^2) \in \Theta \times \Theta$, let $p(\theta^1, \theta^2)$ denote the probability of having the state (θ^1, θ^2), with $p(\theta_L, \theta_H) = p(\theta_H, \theta_L)$ in order to make θ^1 and θ^2 symmetrically distributed. For expositional simplicity, we introduce the following notation:

$$p(\theta_L, \theta_L) = p_{LL}, \quad p(\theta_L, \theta_H) = p(\theta_H, \theta_L) = p_{LH}, \quad p(\theta_H, \theta_H) = p_{HH},$$

with $p_{mn} > 0$ for all $m, n \in \{L, H\}$, and $p_{LL} + 2p_{LH} + p_{HH} = 1$. Let $p(\theta^1|\theta^2)$ denote the conditional distribution of θ^1 given θ^2. We allow θ^1 and θ^2 to be positively or negatively correlated, and let $\rho \equiv p_{LL}p_{HH} - p_{LH}^2$ represent the degree of correlation.[5]

Let $S(q)$ denote the consumer surplus from consuming q units of the final good, with $S'(q) > 0 > S''(q)$ for any $q > 0$. In addition, $S(q)$ satisfies the Inada condition $S'(0) = \infty$, $\lim_{q \to +\infty} S'(q) = 0$.[6] The regulator's objective is given by:[7]

$$W \equiv S(q) - \sum_{i=1}^{2} t^i.$$

According to the revelation principle, we can restrict our attention to direct revelation grand-mechanisms without loss of generality. A direct grand-mechanism M takes the following form, in which $\widehat{\theta}^i$ is A^i's report to P about his cost parameter θ^i:

$$M = \left\{ t^1(\widehat{\theta}^1, \widehat{\theta}^2), t^2(\widehat{\theta}^1, \widehat{\theta}^2), q(\widehat{\theta}^1, \widehat{\theta}^2) \right\}.$$

Since the two agents are ex ante symmetric, we consider only symmetric mechanisms; this means that we impose $t^1(\theta^1, \theta^2) = t^2(\theta^2, \theta^1)$ and $q(\theta^1, \theta^2) = q(\theta^2, \theta^1)$ for any (θ^1, θ^2).[8]

[5] The independent case is included in our framework as a particular case with measure zero.

[6] The assumptions $S'(0) = \infty$ and $\lim_{q \to +\infty} S'(q) = 0$ jointly imply that that P chooses $q > 0$ and finite for any realization of cost parameters (no corner solutions).

[7] As in Gilbert and Riordan (1995), Mookerjee and Tsumagari (2004), and Severinov (2005), we do not include the firms' profits in the objective function.

[8] This restriction entails no loss of generality since (i) when there is no collusion, one of the optimal mechanisms is symmetric (the proof is available from the authors); (ii) in the setting with collusion, we prove below that a suitable symmetric mechanism yields P the same payoff as in the absence of collusion.

For expositional simplicity, we introduce the following notation: for the transfers,

$$\begin{aligned} t_{LL} &= t^1(\theta_L,\theta_L) = t^2(\theta_L,\theta_L), \quad t_{LH} = t^1(\theta_L,\theta_H) = t^2(\theta_H,\theta_L), \\ t_{HL} &= t^1(\theta_H,\theta_L) = t^2(\theta_L,\theta_H), \quad t_{HH} = t^1(\theta_H,\theta_H) = t^2(\theta_H,\theta_H); \end{aligned}$$

and for the quantity to produce,

$$q_{LL} = q(\theta_L,\theta_L), \quad q_{LH} = q(\theta_L,\theta_H) = q(\theta_H,\theta_L) = q_{HL}, \quad q_{HH} = q(\theta_H,\theta_H).$$

Each agent is risk neutral and his reservation utility is normalized to zero regardless of his type. After observing θ^i, A^i should reject or accept the regulator's offer of a grand-mechanism. We assume that A^i's acceptance decision binds and that the grand-mechanism must satisfy the ex post break-even constraint for any firm, in every state of nature, under truthful reporting. This condition is formally defined as follows:

$$t^i(\theta^1,\theta^2) - \theta^i q(\theta^1,\theta^2) \geq 0 \qquad \text{for any } (\theta^1,\theta^2) \in \Theta \times \Theta, \quad \text{for } i = 1, 2.$$

Note that when the condition holds, each agent makes a non-negative profit as long as he reports truthfully, regardless of the report made by the other agent.[9] The requirement to satisfy the ex post break-even constraints can be justified when the regulator is constrained to guarantee a minimum rate of return in each state of nature. Rate of return regulation was the norm before price caps or profit sharing were introduced (e.g. see the introduction in Laffont and Tirole, 1993), and continues to be used in several cases.

The ex post break-even constraints provide the agents with some protection and in particular allow us to depart from the mechanisms *à la* Crémer and McLean (1988) that achieve the first-best outcome whenever the types are correlated. Note that such mechanisms require P to combine almost infinite rewards with almost infinite penalties when correlation is very weak, which are rarely observed in reality.

Under VS and without collusion, the timing of the game is given as follows:

1. Nature draws each agent's cost parameter. Each A^i learns only θ^i.

[9] Notice that if there is no collusion, imposing the ex post break-even constraints is equivalent to assuming limited liability *à la* Sappington (1983), in the sense that each agent has the option of terminating his relationship with P at any time before incurring the production cost, and without paying any fine. However, this would allow agents to collude on participation once they accepted a grand-mechanism, while collusion on participation is impossible in our setting since each agent's acceptance of M is binding. It would be interesting to study collusion on participation under limited liability *à la* Sappington, but this is beyond the scope of this paper.

2. The principal proposes a grand-mechanism M.

3. Each agent accepts or rejects M. If at least one agent rejects, both agents get the reservation utility (zero), and the game ends. Otherwise, the game continues.

4. Each agent A^i makes a report $\widehat{\theta}^i \in \Theta$ in M.

5. Production and transfers are enforced according to M.

3 Optimal organization in the absence of collusion

In this section we analyze the optimal organization of production in the absence of collusion. First, we characterize as a benchmark the first-best solution. Second, we analyze the optimal mechanisms under VS and under VI. Finally, we compare the two mechanisms.

3.1 Benchmark: first-best

If P has complete information about the production costs, then the grand-mechanism M only needs to satisfy the following ex post break-even constraints (for $m, n \in \{L, H\}$):

$$(BE_{mn}) \quad t_{mn} - \theta_m q_{mn} \equiv u_{mn} \geq 0, \tag{1}$$

where u_{mn} is the rent that an m-type earns when the other agent reports an n-type.

Welfare maximization under the ex post break-even constraints yields the following intuitive first-best solution. First, the first-best transfers extract all rents from the agents, i.e. $u_{mn}^{FB} = 0$ for all $m, n \in \{L, H\}$. Second, the first-best output levels, q_{mn}^{FB}, are efficient, i.e. $S'(q_{mn}^{FB}) = \theta_m + \theta_n$ for $m, n \in \{L, H\}$. Finally, an informed principal is indifferent between VS and VI, since the first-best allocation can be implemented under both.

3.2 Vertical separation

In this section we derive the optimal grand-mechanism under VS when the agents behave non-cooperatively, and the principal only knows the cost distribution.

Again, the grand-mechanism M should satisfy the ex post break-even constraints (1) for $m, n \in \{L, H\}$. In order to induce truth-telling, M should satisfy the incentive compatibility constraints for an m-type (with $m, k \in \{L, H\}$ and $k \neq m$):

$$(BIC_m) \quad \sum_{n \in \{L,H\}} \frac{p_{mn}}{p_{mL} + p_{mH}} (t_{mn} - \theta_m q_{mn}) \geq \sum_{n \in \{L,H\}} \frac{p_{mn}}{p_{mL} + p_{mH}} (t_{kn} - \theta_m q_{kn}) \tag{2}$$

The principal maximizes her expected payoff W^S below, where the superscript S stands for separation, subject to (1) and (2) for $m, n \in \{L, H\}$:

$$W^S = \sum_{m\in\{L,H\}} \sum_{n\in\{L,H\}} p_{mn}[S(q_{mn}) - t_{mn} - t_{nm}] \tag{3}$$

The next proposition characterizes the optimal grand-mechanisms under VS when there is no collusion.

Proposition 1 *The optimal grand-mechanism(s) M^* under VS are as follows:*

(i) The optimal quantity schedule $(q^*_{LL}, q^*_{LH}, q^*_{HH})$ satisfies $S'(q^*_{mn}) = \theta^*_{mn}$, with

$$\theta^*_{mn} = \left(\theta_m + \frac{\Pr[\theta^1 < \theta_m, \theta^2 = \theta_n]}{p(\theta_m, \theta_n)}\Delta\right) + \left(\theta_n + \frac{\Pr[\theta^1 = \theta_m, \theta^2 < \theta_n]}{p(\theta_m, \theta_n)}\Delta\right),$$

for $m, n \in \{L, H\}$, and is decreasing ($q^*_{LL} > q^*_{LH} \geq q^*_{HH}$) for $\rho \leq \rho^* \equiv p_{LH}(p_{HH} + p_{LH})$, but non-monotonic ($q^*_{LL} > q^*_{HH} > q^*_{LH}$) for $\rho > \rho^*$.[10]

(ii) The L-type's incentive constraint and both the H-type's ex post break-even constraints are binding. The following transfer scheme, for instance, implements the optimal quantity schedule (with $m \in \{L, H\}$):
(a) for an H-type: $t^*_{Hm} = \theta_H q^*_{mH}$;
(b) for an L-type:

$$t^*_{Lm} = \theta_L q^*_{Lm} + \Delta q^*_{mH}, \qquad \text{for } \rho \leq \rho^*,$$

$$t^*_{Lm} = \theta_L q^*_{Lm} + \Delta q^*_{LH} + \frac{p_{LH}}{p_{LL}}\Delta(q^*_{HH} - q^*_{mH}), \qquad \text{for } \rho > \rho^*.$$

In both cases, the principal has a residual degree of freedom in the choice of (t_{LL}, t_{LH}).

Since the ex post break-even constraints have to be satisfied, an H-type has to be given at least zero rent in every state of nature (proposition 1 iia). This forces P to concede a positive rent to an L-type in order to eliminate this agent's incentive to exaggerate his cost, and satisfy the upward incentive constraint (BIC_L), see proposition 1 (iib). That is, the mechanisms described by Crémer and McLean (1988), which extract the full surplus from the agents, are never feasible here.

By using the binding constraints (BE_{HH}), (BE_{HL}) and (BIC_L) we find the following total expected rent payment:

$$\sum_{m\in\{L,H\}} \sum_{n\in\{L,H\}} p_{mn}(u_{mn} + u_{nm}) = 2p_{LH}\left(\frac{p_{LL}}{p_{LH}}\Delta\right) q_{LH} + p_{HH}\left(2\frac{p_{LH}}{p_{HH}}\Delta\right) q_{HH}. \tag{4}$$

[10] It may be interesting to note that our case of strong positive correlation $\rho > \rho^*$ exactly corresponds to the case of strong positive correlation in Armstrong and Rochet (1999) in which they characterize the solution of a two-dimensional screening problem. We are grateful to Jean-Charles Rochet for this remark.

This expression illustrates that the payment of information rents to an L-type increases the cost of an H-type's input. In particular, (4) shows that the cost of supplying an extra unit of input by an H-type in state (θ_L, θ_H) equals the unit cost of production θ_H plus the unit cost of providing incentives $\frac{p_{LL}}{p_{LH}}\Delta$, i.e. the virtual cost of type H in state (θ_L, θ_H) is $\theta_H + \frac{p_{LL}}{p_{LH}}\Delta$. Similarly, the marginal cost of supplying the input from an H-type in state (θ_H, θ_H) is increased by $\frac{p_{LH}}{p_{HH}}\Delta$.[11] This gives P an incentive to reduce the informational rents by choosing the output levels q^*_{LH} and q^*_{HH} below the first-best output levels q^{FB}_{LH} and q^{FB}_{HH}, respectively. Therefore, the optimal quantity schedule is determined by a trade-off between efficiency and rent extraction. On the other hand, the cost of supplying an L-type's input is not affected by rent payments, since the downward incentive constraint (BIC_H) is slack in the optimum. Thus, the virtual cost for a type L is simply θ_L.

The ranking of second-best quantity levels is the reverse of the ranking of the sum of virtual costs. If correlation is negative or weakly positive ($\rho \leq \rho^*$), the optimal quantity schedule is decreasing in the sum of the two agents' cost parameters ($q^*_{LL} > q^*_{LH} \geq q^*_{HH}$). In the case of strong positive correlation ($\rho > \rho^*$), instead, the sum of virtual costs is non-monotonic (i.e. $\theta^*_{LL} < \theta^*_{HH} < \theta^*_{LH}$), which yields a non-monotonic quantity schedule: $q^*_{LL} > q^*_{HH} > q^*_{LH}$ (see proposition 1 i). If the costs are strongly positively correlated (i.e. p_{LH} is close to zero), then a large downward distortion in output q_{LH} becomes profitable for P. A reduction of q_{LH} reduces the expected information rent in (4) and creates a relatively small expected efficiency loss, since the probability with which the loss is incurred is close to zero. Intuitively, if p_{LH} is close to zero and an L-type reports θ_H, then his probability to produce q_{LH} is much higher than the probability to produce q_{HH}.

Finally, we notice that Bayesian implementation leaves P one degree of freedom in choosing (t_{LL}, t_{LH}).[12] When $\rho \leq \rho^*$, the degree of freedom can be used to find the mechanism described in Proposition 1 that satisfies the following ex post incentive compatibility constraints (for $m, n, k \in \{L, H\}$):

$$t_{mn} - \theta_m q_{mn} \geq t_{kn} - \theta_m q_{kn} \tag{5}$$

When the optimal quantity schedule is non-monotonic, instead, it is not possible to satisfy (1) and (5) for $m, n, k \in \{L, H\}$ within the class of optimal mechanisms. We will show

[11] Precisely, the cost mark-up depends on the so-called inverse hazard rate for input i. The hazard rate of input i is the conditional probability of drawing a certain cost value θ_m for input i given that the cost of this input is smaller or equal to θ_m, i.e. $h_{mn} \equiv \Pr[(\theta^i, \theta^j) = (\theta_m, \theta_n) | \theta^i \leq \theta_m, \theta^j = \theta_n]$.

[12] Bayesian implementation requires that (t_{LL}, t_{LH}) make (BIC$_L$) binding and satisfy (BE$_{LH}$), (BE$_{LL}$) and (BIC$_H$).

later on that the degree of freedom in the choice of (t_{LL}, t_{LH}) turns out to be useful when agents can collude.

3.3 Vertical integration

In this section we study the optimal mechanism under VI. As in Baron and Besanko (1992) and Gilbert and Riordan (1995), the vertically integrated firm is represented by a consolidated agent who has complete information on (θ^1, θ^2) and maximizes the total transfer received by the two firms minus the total production cost $(\theta^1 + \theta^2)q$. Thus, the principal proposes a menu of contracts to the consolidated agent as in a standard single-agent, one-dimensional screening problem with type $\theta^1 + \theta^2$. We denote this menu as $M^I = \{(q^I_{LL}, t^I_{LL}), (q^I_{LH}, t^I_{LH}), (q^I_{HL}, t^I_{HL}), (q^I_{HH}, t^I_{HH})\}$, in which $q^I_{LH} = q^I_{HL}$ and $t^I_{LH} = t^I_{HL}$. The timing of the game under VI is similar to the timing under VS presented in section 2. Mechanism M^I should satisfy the following participation constraints for $m, n \in \{L, H\}$:

$$(PC^I) \quad t^I_{mn} - (\theta_m + \theta_n)q^I_{mn} \equiv u^I_{mn} \geq 0. \tag{6}$$

Note that the ex post break-even constraints are satisfied as long as (PC^I) is satisfied for $m, n \in \{L, H\}$. M^I should also satisfy the following incentive constraints for any $m, n, \hat{m}, \hat{n} \in \{L, H\}$:

$$(IC^I) \quad u^I_{mn} \geq t^I_{\hat{m}\hat{n}} - (\theta_m + \theta_n)q^I_{\hat{m}\hat{n}}. \tag{7}$$

The regulator maximizes the following objective subject to (6) and (7):

$$W^I = \sum_{m\in\{L,H\}} \sum_{n\in\{L,H\}} p_{mn}(S(q^I_{mn}) - t^I_{mn}).$$

The optimal mechanism for this setting is well known in the literature (e.g. see section 3.1 in Laffont and Martimort, 2002), and is characterized below.

Proposition 2 *The optimal mechanism under VI is characterized as follows*
(a) The optimal quantity schedule $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$ *is such that:*
If $\rho \leq \rho^{**} \equiv p_{LH}(2 - p_{LH})$, *then* $S'(q^{*I}_{mn}) = \theta^I_{mn}$, *for* $m, n \in \{L, H\}$, *with*

$$\theta^I_{mn} = \theta_m + \theta_n + \frac{\Pr[\theta^1 + \theta^2 < \theta_m + \theta_n]}{\Pr[\theta^1 + \theta^2 = \theta_m + \theta_n]}\Delta$$

If $\rho > \rho^{**}$, *then* $S'(q^{*I}_{LL}) = 2\theta_L$ *and* $S'(q^{*I}_{LH}) = S'(q^{*I}_{HH}) = E\left\{\theta^I_{mn}\middle|\, \theta^1 + \theta^2 \neq \theta_L + \theta_L\right\}$.
(b) The participation constraint of the HH*-type and the local upward incentive constraints bind:* $t^{*I}_{HH} = 2\theta_H q^{*I}_{HH}$, $t^{*I}_{LH} = (\theta_L + \theta_H)q^{*I}_{LH} + \Delta q^{*I}_{HH}$ *and* $t^{*I}_{LL} = 2\theta_L q^{*I}_{LL} + \Delta(q^{*I}_{LH} + q^{*I}_{HH})$.

In contrast to the second-best solution under VS, incentive compatibility under VI implies the monotonicity of the quantity schedule, i.e. $q_{LL}^{*I} \geq q_{LH}^{*I} \geq q_{HH}^{*I}$. The binding incentive and participation constraints under VI result in the following expected information rent payment (proposition 2 b):

$$\sum_{m\in\{L,H\}} \sum_{n\in\{L,H\}} p_{mn} u_{mn}^I = 2p_{LH}\left(\frac{p_{LL}}{2p_{LH}}\Delta\right) q_{LH}^I + p_{HH}\left(\frac{1-p_{HH}}{p_{HH}}\Delta\right) q_{HH}^I, \tag{8}$$

The right hand side of this expression illustrates the virtual cost mark-up under VI, as the right hand side of (4) for the case of VS. Now the mark-up is related to the hazard rate of the sum of costs. In particular, the monotonicity constraint binds when $\theta_{LH}^I > \theta_{HH}^I$, which is equivalent to $\rho > \rho^{**}$, and in this case we have $S'(q_{LH}^{*I}) = S'(q_{HH}^{*I}) = \frac{2p_{LH}}{1-p_{LL}}\theta_{LH}^I + \frac{p_{HH}}{1-p_{LL}}\theta_{HH}^I$ (proposition 2 a).

3.4 Vertical separation versus vertical integration

In this section we examine the principal's choice between VS and VI. Consolidating two agents into a single agent has two effects (see e.g. Mookherjee and Tsumagari, 2004): *internalization of bidding externalities* and *loss of control*. First, under VS independent reporting by agents creates bidding externalities. Specifically, a higher cost reported by one agent modifies the quantity produced by the other agent if he also exaggerates his cost, and thereby changes his rent from reporting untruthfully. Under VI the consolidated agent internalizes this externality. Second, under VI the principal faces a more stringent set of incentive compatibility constraints than under VS (loss of control). The consolidated agent knows (θ^1, θ^2) and exercises complete control over reporting, while each separated agent has only partial control. This reduces the set of feasible mechanisms under VI. The trade-off between these two effects determines the optimal organization of production, which is characterized below.

Proposition 3 *(i) The principal prefers VI if $\rho \leq \rho^*$ while she prefers VS if $\rho \geq \rho^{**}$. If $\rho^* < \rho < \rho^{**}$, her preference between VS and VI depends on $S(\cdot)$.*

(ii) Suppose that $S(q) = aq - \frac{b}{2}q^2$ with $a(>0)$ sufficiently large and $b > 0$. Then there exists a $\widetilde{\rho} \equiv \frac{p_{LH}}{3}(4 - 3p_{LH} - 2p_{LL}) \in (\rho^*, \rho^{**})$ such that the principal prefers VI to VS if and only if $\rho < \widetilde{\rho}$.

First, if $\rho \leq \rho^*$, then the optimal quantity schedule $(q_{LL}^*, q_{LH}^*, q_{HH}^*)$ under VS, described by Proposition 1, is implementable under VI. Therefore, there is no effect from loss of control. However, the effect from internalization of bidding externalities is present and positive. If an L-type agent exaggerates his cost, then this reduces the gain from cost exaggeration by the other agent, since the quantity schedule is monotonic. Under VI the consolidated agent internalizes this negative externality, which enables the principal to save informational rents. More formally, we illustrate the difference in expected rent payments under VS and VI for implementing a schedule (q_{LL}, q_{LH}, q_{HH}) by taking the difference between (4) and (8):

$$\sum_{m \in \{L,H\}} \sum_{n \in \{L,H\}} p_{mn}(u_{mn} + u_{nm} - u^I_{mn}) = p_{LL}(q_{LH} - q_{HH})\Delta \tag{9}$$

For a low cost correlation, the optimal quantity schedule under VS is monotonic (i.e. $q^*_{LH} > q^*_{HH}$) and can be implemented under VI at a lower cost since the expected rent difference in (9) is positive. This implies that VI dominates VS.

Second, if $\rho \geq \rho^{**}$, then the optimal quantity schedule under VI can be implemented under VS. In this case the effect from internalization of bidding externalities is not present. If an L-type agent exaggerates his cost, then this does not affect the gain from cost exaggeration by the other agent, since q^{*I}_{LH} and q^{*I}_{HH} are equal (see proposition 2 a). The equality of expected rent payments under VS and VI is illustrated in (9), which is zero for $q^{*I}_{LH} = q^{*I}_{HH}$. However, the effect from loss of control is negative for VI. The principal can reach a higher payoff by choosing a non-monotonic quantity schedule under VS, which is infeasible under VI.[13] Therefore, VS dominates VI for $\rho \geq \rho^{**}$.

Finally, if $\rho^* < \rho < \rho^{**}$, both effects emerge. On the one hand, it is less expensive to implement the monotonic schedule $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$ under VI than under VS, i.e. the effect from internalizing bidding externalities is positive. On the other hand, the principal can enhance her payoff under VS by choosing a schedule that is infeasible under VI, i.e. the effect from loss of control is negative. The trade-off between these two conflicting effects makes the general comparison ambiguous. Proposition 3(ii) shows that in the case of quadratic surplus function, there is a cut-off level of correlation $\widetilde{\rho}$ between ρ^* and ρ^{**} such that VI dominates VS if and only if $\rho < \widetilde{\rho}$.

[13] For example, when the correlation is almost perfect (i.e. $p_{LH} \simeq 0$), the principal can almost achieve the first-best outcome under VS by choosing q^*_{LH} close to zero. By contrast, under VI, the principal cannot get close to the first-best, since a consolidated agent with cost $2\theta_L$ would have an incentive to report $2\theta_H$ and obtain a rent of $2\Delta q^{FB}_{HH} > 0$.

4 Vertical separation and collusion

In this section we study the optimal mechanism under VS when the agents collude and investigate how collusion affects the comparison between VS and VI.

4.1 Coalition formation and timing

We model collusion between A^1 and A^2 by a (direct) side-contract S offered by a benevolent and *uninformed* third-party, T (e.g. see Laffont and Martimort, 1997, 2000).[14] Given M designed by P, the third-party maximizes the sum of the agents' rents by implementing joint manipulations of reports into M subject to incentive compatibility constraints, ex post break-even constraints, and acceptance constraints that induce the agents to accept S instead of playing M non-cooperatively. In order to satisfy these constraints, T can use balanced side transfers between the agents.

A side-contract contains two elements. First, the manipulation of report function, $\phi(\cdot)$, maps any pair of reports made by the agents to T (in S) into the set of (possibly stochastic) reports to P (in M). Second element is a monetary side-transfer, $y^i(\cdot)$, from A^i to T. Since T is not a source of money, the following budget balance constraints must hold:

$$\sum_{i=1}^{2} y^i(\theta^1, \theta^2) = 0, \quad \text{for any } (\theta^1, \theta^2) \in \Theta \times \Theta.$$

Thus, the side-contract takes the following form, where $\widetilde{\theta}^i$ is A^i's report to T about θ^i:

$$S = \left\{ \phi(\widetilde{\theta}^1, \widetilde{\theta}^2), y^1(\widetilde{\theta}^1, \widetilde{\theta}^2), y^2(\widetilde{\theta}^1, \widetilde{\theta}^2) \right\}.$$

We assume that S must satisfy the following ex post break-even constraints:[15]

$$t^i(\phi(\theta^1, \theta^2)) - y^i(\theta^1, \theta^2) - \theta^i q(\phi(\theta^1, \theta^2)) \geq 0 \quad \text{for any } (\theta^1, \theta^2) \in \Theta \times \Theta. \tag{10}$$

[14] Any outcome of an extensive form game of bargaining can be achieved by a side-contract designed by such a third-party. Hence, our modelling strategy is a shortcut which characterizes the upperbound of what can be achieved by collusion.

[15] This assumption allows us to capture a situation in which both P and T face similar constraints. This is important because if T does not need to satisfy ex post break-even constraints while P does, then the weakly collusion-proofness principle below does not apply. Therefore, there may be outcomes which P can achieve by letting collusion occur, but cannot achieve in a collusion-proof way; in short, collusion may be beneficial to P. A setting in which collusion is beneficial is theoretically interesting, but we think that it is not very relevant in our regulatory context.

The constraints can be justified by financial constraints for the firms. Assume that the firms have no assets to start with and cannot borrow money to cover a loss (as in Laffont and Martimort, 2002, p.296). Then a firm is forced to default on the grand-mechanism and produce nothing if S violates (10). As a consequence, the firm will face sanctions from the regulator in addition to losing its reputation. Furthermore, the default can be regarded as evidence of collusion (recall that M satisfies the ex post break-even constraints). As collusion is illegal, the default would provoke criminal investigations and the managers of the firms might be imprisoned. Since the third party wants to maximize the firms' payoffs, he will not design S which violates (10). An alternative justification is based on limited enforceability of S. More precisely, suppose that the manager of a firm can renege ex post on the side-contract if respecting the latter's terms makes his firm incur a loss (see e.g. Laffont and Martimort, 2002, p.64). In this situation, by reneging on S and playing M, the firm can get a non-negative profit and hence the manager's temptation to renege is strong.

The timing of the game with collusion under VS is as before, except that stage 4 is replaced by the following steps:

4(a). The third-party offers a side-contract S.

4(b). Each agent accepts or rejects S.

4(c). If both agents accept S, then each agent reports to T (in S) and subsequently side-transfers to T and reports to P (in M) are enforced according to S. Otherwise, M is played non-cooperatively and reports are made directly to P (in M).

4.2 Benchmark: collusion under complete information

In this subsection we study the benchmark of collusion under complete information. For this purpose, we consider the timing described in the previous subsection but assume that the third-party knows the true type of each agent, as in Baron and Besanko (1999). Therefore, a side-mechanism does not need to satisfy any incentive constraint. Once a side-contract is accepted by both agents, they will behave as if they were one consolidated agent. Therefore, there is no loss of generality in restricting attention to the set of direct mechanisms inducing truth-telling of the consolidated agent. This implies that a grand-mechanism should satisfy the following coalition incentive constraint

$$t_{mn} + t_{nm} - (\theta_m + \theta_n)q_{mn} \geq t_{\hat{m}\hat{n}} + t_{\hat{n}\hat{m}} - (\theta_m + \theta_n)q_{\hat{m}\hat{n}} \tag{11}$$

for every state of nature, that is for $m, n \in \{L, H\}$ and $\hat{m}, \hat{n} \in \{L, H\}$. This is a condition analogous to (7), but it applies to a mechanism M rather than to a menu of contracts M^I. Moreover, the ex post break-even constraints (10) together with truth-telling and budget balance imply the participation constraint

$$t_{mn} + t_{nm} - (\theta_m + \theta_n)q_{mn} \geq 0$$

for $m, n \in \{L, H\}$; this is a condition analogous to (6).

Therefore, P's expected payoff under VS with collusion cannot be higher than the payoff under VI. In fact, P can achieve the expected payoff of VI when the agents collude under complete information by choosing the following grand-mechanism.

$$M^C \quad : \quad \begin{array}{l} q^C_{mn} = q^{*I}_{mn}, \qquad t^C_{mm} = \frac{1}{2}t^{*I}_{mm}, \text{ for } m, n \in \{L, H\}, \\ t^C_{HL} = \theta_H q^{*I}_{LH}, \qquad t^C_{LH} = t^{*I}_{LH} - t^C_{HL}. \end{array}$$

The grand-mechanism M^C is incentive feasible, and implements the optimal quantity schedule under VI by paying the same sum of transfers to the agents as under VI. Therefore, P achieves the same expected payoff through M^C as under VI, which we state below.

Proposition 4 *If vertically separated agents can collude under complete information, then VS is equivalent to VI.*

4.3 Weak collusion proofness

In the rest of this section we consider collusion under asymmetric information. In this subsection we define weak collusion proofness, and show that P cannot gain from designing a grand mechanism which is not weakly collusion-proof.

After T proposes S, a two-stage game is played between A^1 and A^2. First, the agents accept or refuse S (stage 4b), and then they report either to P or to T depending on their previous acceptance decisions (stage 4c). We call this two-stage game the game of coalition formation. We are interested in Perfect Bayesian equilibria of this game in which both agents accept S, i.e. there is no learning about types along the equilibrium path. However, before accepting or rejecting S, an agent needs to know what is going to happen off-the equilibrium path. As in Laffont and Martimort (1997, 2000), we assume the following about off-the-equilibrium-path beliefs and events.

Assumption WCP (weak collusion proofness) Given an incentive compatible M, if A^i rejects S (an off-the-equilibrium-path event) then the other agent still has prior beliefs about θ^i and the truthful equilibrium is played in M.

In order to define weakly collusion-proof grand mechanisms, we need to introduce some definitions. We let $U^{\bullet}_m$ $(m = L, H)$ denote the expected payoff of an m-type in the truthful equilibrium of M. Thus, $U^{\bullet}_m$ plays the role of the reservation utility for an m-type when he decides whether to accept S or not. Let ϕ_{mn} and y^i_{mn} denote $\phi(\theta_m, \theta_n)$ and $y^i(\theta_m, \theta_n)$, respectively, for $m, n \in \{L, H\}$ and $i = 1, 2$.

Definition 1 *A side-contract $S^{\bullet} = \{\phi^{\bullet}(\cdot), y^{1\bullet}(\cdot), y^{2\bullet}(\cdot)\}$ is coalition-interim-efficient with respect to a grand-mechanism $M = \{t^1(\cdot), t^2(\cdot), q(\cdot)\}$ providing the reservation utilities $\{U^{\bullet}_L, U^{\bullet}_H\}$ if and only if it is a solution of the following program:*

$$\max_{\phi, y^1, y^2} \sum_{m \in \{L,H\}} \sum_{n \in \{L,H\}} p_{mn}[t^1(\phi_{mn}) + t^2(\phi_{mn}) - (\theta_m + \theta_n)q(\phi_{mn})] \tag{12}$$

subject to

$$U^1_m = \sum_{n \in \{L,H\}} \frac{p_{mn}}{p_{mL} + p_{mH}}[t^1(\phi_{mn}) - y^1_{mn} - \theta_m q(\phi_{mn})], \quad \textit{for any } m \in \{L, H\};$$
$$U^2_m = \sum_{n \in \{L,H\}} \frac{p_{nm}}{p_{Lm} + p_{Hm}}[t^2(\phi_{nm}) - y^2_{nm} - \theta_m q(\phi_{nm})], \quad \textit{for any } m \in \{L, H\};$$

$$(BIC^{1S}) \quad U^1_m \geq \sum_{n \in \{L,H\}} \frac{p_{mn}}{p_{mL} + p_{mH}}[t^1(\phi_{kn}) - y^1_{kn} - \theta_m q(\phi_{kn})], \quad \textit{for any } m, k \in \{L, H\};$$
$$(BIC^{2S}) \quad U^2_m \geq \sum_{n \in \{L,H\}} \frac{p_{nm}}{p_{Lm} + p_{Hm}}[t^2(\phi_{nk}) - y^2_{nk} - \theta_m q(\phi_{nk})], \quad \textit{for any } m, k \in \{L, H\};$$

$$(BIR^S) \quad U^i_m \geq U^{\bullet}_m, \quad \textit{for any } m \in \{L, H\};$$
$$(BB) \quad y^1_{mn} + y^2_{nm} = 0, \quad \textit{for any } m, n \in \{L, H\};$$
$$(Ex\ post\ BE) \quad \begin{array}{l} t^1(\phi_{mn}) - y^1_{mn} - \theta_m q(\phi_{mn}) \geq 0 \\ t^2(\phi_{nm}) - y^1_{nm} - \theta_m q(\phi_{nm}) \geq 0 \end{array}, \quad \textit{for any } m, n \in \{L, H\}.$$

A side-contract is coalition-interim-efficient if it maximizes the sum of the agents' expected profits subject to incentive, acceptance, budget balance, and ex post break-even constraints. The acceptance constraint (BIR^S) is defined in Bayesian terms and guarantees that A^i's payoff by accepting S is not smaller than the reservation utility $U^{\bullet}_m$ he obtains by playing M non-cooperatively if S is rejected.

Definition 2 *The null side-contract S^0 is the side-contract where there is no manipulation of report $(\phi(\cdot) = I_d(\cdot))$ and no side transfers between agents $(y^1(\cdot) = y^2(\cdot) = 0)$.*

If the third party chooses $S = S^0$, then M is not affected by collusion since S^0 implements no manipulation of reports. This motivates the following definition of weakly collusion-proof grand-mechanisms.

Definition 3 *An incentive compatible grand-mechanism M providing the reservation utilities $\{U_L^\bullet, U_H^\bullet\}$ is said to be weakly collusion-proof when S^0 is coalition-interim-efficient with respect to M.*

We are now able to state a result which allows us to restrict attention to weakly collusion proof mechanisms.

Proposition 5 *There is no loss of generality in restricting the principal to offer weakly collusion-proof mechanisms to characterize the outcome of any perfect Bayesian equilibria of the game of grand-mechanism offer cum coalition formation such that a collusive equilibrium occurs on the equilibrium path.*

Proposition 5 states that all the outcomes that can be implemented by allowing collusion to happen can be mimicked by P in a collusion-proof way and implies that P cannot realize a higher payoff when the agents can collude than when they do not. Indeed, when T proposes S^0, (i) Ex post BE and the Bayesian incentive constraints (BIC^S) in the side mechanism reduce to (1) and (2) for $m, n \in \{L, H\}$, respectively; (ii) the acceptance constraints (BIR^S) are automatically satisfied with equality. Hence, under collusion, P needs to choose M to maximize her payoff subject to (1)-(2) for $m, n \in \{L, H\}$, *and* the conditions that make M weakly collusion-proof.[16] Since collusion imposes extra constraints on P's optimization problem, we obtain the following corollary.

Corollary 1 *Under VS, the principal's payoff when agents can collude under asymmetric information is never larger than the payoff without collusion.*

This corollary gives an upper bound for P's payoff. In the next subsection we show that it is possible to (almost) reach this upper bound.

4.4 The optimal weakly collusion proof mechanism

We first state the results.

[16] These are the conditions which deter manipulations of reports. The conditions are explained later on in this section and, with more details, in the proof of Proposition 6.

Proposition 6 *Consider VS and suppose that the agents can collude under asymmetric information and that* Δ *is sufficiently small when* $\rho > \rho^*$*. Then, for any given* ρ*, there exists a grand mechanism which is both optimal (or almost optimal) in the absence of collusion and weakly collusion proof:* M' *if* $\rho \leq \rho^*$ *and* $M''(\varepsilon)$ *with* $\varepsilon(> 0)$ *close to 0 if* $\rho > \rho^*$*. This mechanism is essentially unique if* $\rho > \rho^*$*.*[17]

We will define M' and $M''(\varepsilon)$ later on. The proposition states that under VS, collusion does not hurt P. Therefore, we get the following result:

Proposition 7 *Suppose that under VS, the agents can collude under asymmetric information and that* Δ *is sufficiently small when* $\rho > \rho^*$*. Then, for any given* ρ*, the principal's preference between VS and VI is not affected by the agents' ability to collude under VS, and is described by Proposition 3.*

In what follows, we explain the result of Proposition 6 with minimum technical details. We distinguish the case of $\rho \leq \rho^*$ from the case of $\rho > \rho^*$. In each case, we provide a weakly collusion-proof optimal mechanism.

4.4.1 The case of $\rho \leq \rho^*$

If $\rho \leq \rho^*$, one optimal collusion-proof mechanism M' is defined below; notice that M' differs from M^* in Proposition 1 only in terms of t_{LL} and t_{LH}.

$$M' \quad : \quad \begin{array}{l} q'_{mn} = q^*_{mn}, \qquad t'_{Hn} = t^*_{Hn}, \qquad \text{for all } m, n \in \{L, H\}, \\ t'_{LL} = t^*_{LL} - \frac{X}{p_{LL}}, \qquad t'_{LH} = t^*_{LH} + \frac{X}{p_{LH}}, \quad \text{with } X \equiv \frac{p_{LL}p_{LH}}{1-p_{HH}}(q^*_{LH} - q^*_{HH})\Delta \end{array}$$

By Proposition 1, M' is optimal in the absence of collusion. Furthermore, it is easy to see that M' satisfies (11). This implies that the objective function (12) for the third party is maximized by truth-telling (i.e., by $\phi(.) = Id$) and thus the null side mechanism S^0 is coalition-interim-efficient because it is feasible and maximizes (12). This shows that M' is weakly collusion proof. Basically, when the quantity schedule is monotonic, P can use the residual degree of freedom in the choice of (t_{LL}, t_{LH}) to satisfy (11) without incurring any additional cost.

When a mechanism M violates (11), we say that M exhibits *room for collusion* because some manipulation of reports would occur if (θ^1, θ^2) were complete information among the agents. We have shown that M' leaves no room for collusion when $\rho \leq \rho^*$.

[17] The mechanism is essentially unique in the sense it is unique given ε, and it is profitable to choose ε close to 0. However, as we shall explain below, $M''(0)$ is not collusion-proof.

4.4.2 The case of $\rho > \rho^*$

When $\rho > \rho^*$, we can easily see that if a grand-mechanism M is optimal under no collusion, then it leaves room for collusion. The optimal quantity schedule under VS in the absence of collusion is such that $q^*_{LH} < q^*_{HH}$ when $\rho > \rho^*$. However, P cannot implement a non-monotonic quantity schedule and satisfy (11), since incentive compatibility in a one-agent setting implies that the monotonicity condition $q_{LH} \geq q_{HH}$ must be satisfied as we have seen in the case of VI.[18]

The fact that any mechanism M that is optimal in the absence of collusion leaves room for collusion does not necessarily imply that collusion reduces the payoff of the principal. Since the agents collude under asymmetric information, they have to solve an incentive problem within the coalition. This may make it impossible for them to exploit the potential gain from collusion. Indeed, as long as Δ is sufficiently small, we can find $M''(\varepsilon)$ which is (almost) optimal in the absence of collusion and weakly collusion proof, even though it fails to satisfy (11). Precisely, consider

$$M''(\varepsilon) \ : \ \begin{array}{l} q''_{mn} = q^*_{mn}, \quad t''_{HL} = t^*_{HL} + \varepsilon, \quad t''_{HH} = t^*_{HH} \\ t''_{LL} = \theta_L q^*_{LL} + \Delta q^*_{HH} + \varepsilon, \quad t''_{LH} = \theta_L q^*_{LH} + \Delta q^*_{HH} - \dfrac{p_{LL}}{p_{LH}} \Delta (q^*_{HH} - q^*_{LH}) \end{array}$$

with $\varepsilon (> 0)$ small. The assumption that Δ is small implies that q^*_{HH} is close enough to q^*_{LH} that $M''(\varepsilon)$ satisfies break-even constraint (BE_{LH}). We notice that $M''(0)$ is optimal under no collusion (by Proposition 1) but $M''(\varepsilon)$ is not so. However, we explain below that $M''(\varepsilon)$ is weakly collusion proof, unlike $M''(0)$. The difference of expected payoff for P between $M''(0)$ and $M''(\varepsilon)$ is $2(p_{LL} + p_{LH})\varepsilon$. Since ε can be chosen arbitrarily small, the supremum of P's payoff when agents can collude is equal to her payoff when agents cannot collude.

We can get some intuition for the weak collusion proofness of $M''(\varepsilon)$ by analyzing the incentive and acceptance constraints of a side-contract. In the side-contract that is optimal with respect to $M''(\varepsilon)$, the incentive constraint (BIC^S_L) is binding. In particular, the information rent of an L-type, which gives him an incentive to report truthfully to T, increases with the quantity produced by an H-type. Therefore, T evaluates an H-type's input at a virtual cost which is larger than the real cost θ_H. This can create distortions in T's decisions to manipulate reports (compared to the case in which the manipulation decisions are taken under complete information), as we explain below.

[18] Precisely, if $\rho > \rho^*$, then mechanism M^* in Proposition 1 violates (11) for $\theta^1 + \theta^2 = \theta_L + \theta_H$, and $\widehat{\theta}^1 + \widehat{\theta}^2 = 2\theta_H$.

As in the design of M by P, the virtual cost of an L-type agent in the design of S is equal to his real cost θ_L. Therefore, T has no incentive to manipulate the cost reports of an LL coalition if and only if the following coalition incentive constraints (CIC) are satisfied:

$$(CIC_{LL,LH}) \qquad 2t_{LL} - 2\theta_L q_{LL} \geq t_{HL} + t_{LH} - 2\theta_L q_{LH}; \tag{13}$$

$$(CIC_{LL,HH}) \qquad 2t_{LL} - 2\theta_L q_{LL} \geq 2t_{HH} - 2\theta_L q_{HH}. \tag{14}$$

It can be verified that $M''(\varepsilon)$ satisfies both constraints.

For an LH coalition, the coalition incentive constraints can be written as follows:

$$(CIC_{LH,LL}) \quad t_{HL} + t_{LH} - (\theta_L + \theta_H + \mu_{LH})q_{LH} \geq 2t_{LL} - (\theta_L + \theta_H + \mu_{LH})q_{LL} \tag{15}$$

$$(CIC_{LH,HH}) \quad t_{HL} + t_{LH} - (\theta_L + \theta_H + \mu_{LH})q_{LH} \geq 2t_{HH} - (\theta_L + \theta_H + \mu_{LH})q_{HH} \tag{16}$$

where $\theta_L + \theta_H + \mu_{LH}$ is the virtual cost of production in state (θ_L, θ_H) from the point of view of T. We show in the proof of Proposition 6 that since (BIC_L) binds in $M''(\varepsilon)$, P can choose μ_{LH} in $[0, \frac{p_{LL}}{p_{LH}}\Delta)$. We notice that if μ_{LH} is close to $\frac{p_{LL}}{p_{LH}}\Delta$, then $\theta_L + \theta_H + \mu_{LH}$ is close to θ^*_{LH}, the virtual cost in state (θ_L, θ_H) from P's point of view when there is no collusion. Although, $M''(0)$ satisfies $(CIC_{LH,LL})$ strictly and $(CIC_{LH,HH})$ with equality if $\mu_{LH} = \frac{p_{LL}}{p_{LH}}\Delta$, $M''(0)$ is actually not weakly collusion proof because the value $\frac{p_{LL}}{p_{LH}}\Delta$ is not feasible for μ_{LH}. In $M''(\varepsilon)$ we add a small but positive ε to t''_{HL} and to t''_{LL} (so that (BIC_L) is still binding), and thus (15)-(16) are satisfied for μ_{LH} smaller than $\frac{p_{LL}}{p_{LH}}\Delta$ but close to this value. Furthermore, in the proof of proposition 6 we show that $M''(\varepsilon)$ satisfies all the coalition incentive constraints for an HH coalition. Thus, we conclude that $M''(\varepsilon)$ is weakly collusion proof.

We notice that if $\mu_{LH} = 0$ then $(CIC_{LH,LL})$ and $(CIC_{LH,HH})$ are identical to (11) with $m = L$, $n = H$. In that case, $M''(\varepsilon)$ satisfies $(CIC_{LH,LL})$ but violates $(CIC_{LH,HH})$. The fact that $M''(\varepsilon)$ violates $(CIC_{LH,HH})$ if $\mu_{LH} = 0$ means that there is a potential gain from manipulating reports from LH to HH. However, when the agents collude under asymmetric information, P can induce T to evaluate the production cost of an LH coalition in terms of a virtual cost which is larger than the real cost $\theta_L + \theta_H$ and close to θ^*_{LH}. This in turn penalizes a manipulation that increases the quantity produced by the coalition from q^*_{LH} to q^*_{HH}, and the third party finds the manipulation not profitable.

Alternatively, we can explain the failure to implement a report manipulation from LH to HH by comparing the gains from the manipulation and the transaction costs of implementing the manipulation. Given $M''(\varepsilon)$, consider a side-contract S'' with manipulation of reports $\phi''(\theta_L, \theta_H) = \phi''(\theta_H, \theta_L) = (\theta_H, \theta_H)$, and no joint manipulation otherwise. Since the agents are ex ante symmetric, we restrict ourselves without loss of generality to a symmetric side-contract S''. Hence, S'' specifies a side-payment $\widehat{y}''$ from an L-type to an H-type only when the agents announce (θ_L, θ_H) or (θ_H, θ_L) to T. On the one hand, the expected cost of implementing S'' can be computed as follows. First, the acceptance constraint of an H-type agent puts a lower bound on the side-payment $\widehat{y}''$. The acceptance constraint $(BIR_H^{S''})$ requires that the H-type agent's expected profit from accepting S'', $\frac{p_{LH}}{p_{LH}+p_{HH}}\widehat{y}''$, is greater than or equal to his expected profit in the absence of collusion, $U_H^{\bullet} = \frac{p_{LH}}{p_{LH}+p_{HH}}\varepsilon$, which gives $\widehat{y}'' \geq \varepsilon$. The expected payoff of an L-type agent from overstating his cost to the third party, given truthful reporting by the other agent, equals: $\Delta q_{HH}^* + \frac{p_{LL}}{p_{LL}+p_{LH}}\widehat{y}''$, which is at least $\Delta q_{HH}^* + \frac{p_{LL}}{p_{LL}+p_{LH}}\varepsilon$. Since the payoff of an L-type agent in the truthful equilibrium of $M''(\varepsilon)$ is $U_L^{\bullet} = \Delta q_{HH}^* + \frac{p_{LL}}{p_{LL}+p_{LH}}[\varepsilon - \Delta(q_{HH}^* - q_{LH}^*)]$, we conclude that the implementation of S'' increases an L-type's rent by at least $\frac{p_{LL}}{p_{LL}+p_{LH}}\Delta(q_{HH}^* - q_{LH}^*)$. Therefore, the expected rent increase paid by T is greater than or equal to $2p_{LL}\Delta(q_{HH}^* - q_{LH}^*)$, which we define as the transaction costs in side-contracting generated by asymmetric information. On the other hand, the expected gain from manipulating the reports through S'' is $2p_{LH}[2t''_{HH} - (\theta_H + \theta_L)q_{HH}^* - (t''_{LH} + t''_{HL} - (\theta_H + \theta_L)q_{LH}^*)]$, which equals: $2p_{LL}\Delta(q_{HH}^* - q_{LH}^*) - 2p_{LH}\varepsilon$. Since the gains from the manipulation are smaller than the transaction costs, the manipulation cannot be implemented.[19]

In summary, only the rules that give an L-type more than the whole gains from collusion are incentive compatible within the coalition. This makes collusion fail because there is not enough left to compensate H-type's loss from the manipulation. Therefore, P can implement a non-monotonic quantity schedule in a collusion-proof way. In contrast, if collusion takes place under complete information, P cannot implement a non-monotonic quantity schedule.

4.5 Ratifiability

In the game of coalition formation, after S^0 is proposed by the third party, a two stage game starts. In the first stage each A^i accepts or rejects S^0; in the second stage, agents report types either in M if some agent vetoed S^0,[20] or in S^0 otherwise. In any case,

[19] Given the inequality $\widehat{y}'' \geq \varepsilon$ derived from $(BIR_H^{S''})$, a simpler (but perhaps less insightful) way to see that the manipulation from LH to HH cannot be implemented exploits $(BIC_L^{S''})$, which is $p_{LL}(\Delta q_{HH}^* + \varepsilon) + p_{LH}(\Delta q_{HH}^* - \widehat{y}'') \geq p_{LL}(\Delta q_{HH}^* + \widehat{y}'') + p_{LH}\Delta q_{HH}^*$, or $\frac{p_{LL}}{p_{LL}+p_{LH}}\varepsilon \geq \widehat{y}''$. Since $\frac{p_{LL}}{p_{LL}+p_{LH}}\varepsilon < \varepsilon$, we infer that $(BIC_L^{S''})$ and/or $(BIR_H^{S''})$ is violated.

[20] Here we use M to represent M' when $\rho \leq \rho^*$ and $M''(\varepsilon)$ when $\rho > \rho^*$.

however, M is actually played in the second stage since S^0 is the null side mechanism; the first stage is therefore a sort of cheap-talk stage in which agents may signal their types. We introduced above assumption WCP, under which the truthful Bayesian equilibrium (BNE) of M, supported by prior beliefs, is played when S^0 is vetoed. However, A^i's acceptance decision of S^0 may be affected if he expects the other agent to infer non-prior beliefs from his veto of S^0. We use here the notion of strong ratifiability of a mechanism against itself provided by Cramton and Palfrey (1995) to test whether M' or $M''(\varepsilon)$ is robust to the pre-play announcements at the first stage. In presenting ratifiability we follow Laffont and Martimort (2000) and suppose without loss of generality that it is A^1 the agent who contemplates rejecting S^0.

It is useful to introduce the following notation. We use σ to denote a generic strategy profile in M; $\sigma = (LH, LL)$, for example, is the (pure-strategy) profile in which both types of A^1 report truthfully and A^2 always claims L; $\sigma^\bullet = (LH, LH)$ is the truthful reporting profile. Let $\bar{p}$ represent a belief system of A^2 about θ^1 which may differ from the prior beliefs p. More clearly, $\bar{p}(\theta_m|\theta_n)$ is the probability that A^2 attaches, according to the belief system $\bar{p}$, to the event $\theta^1 = \theta_m$ given that his own type is n ($m, n \in \{L, H\}$); $\bar{p}(\theta_m|\theta_n)$ may differ from the prior belief $p(\theta_m|\theta_n)$. Let $BNE(M_{\bar{p},p})$ represent the set of BNE of M when A^1 has prior beliefs about θ^2 and A^2's beliefs about θ^1 are given by $\bar{p}$. Clearly, M is incentive compatible if and only if $\sigma^\bullet \in BNE(M_{p,p})$. Finally, let A^i_m denote A^i having type $m \in \{L, H\}$, $U^i_m(\sigma)$ denote the payoff of A^i_m when σ is played in M, computed with prior beliefs; we have already introduced $U^\bullet_j \equiv U^1_j(\sigma^\bullet) = U^2_j(\sigma^\bullet)$ as the payoff in the truthful BNE.

Definition 4 *Given an incentive compatible M, a belief system $\bar{p}$ is a credible veto system of $\sigma^\bullet$ if there exists $\sigma \in BNE(M_{\bar{p},p})$ and refusal probabilities r_m ($m \in \{L, H\}$) for A^1 such that $r_L + r_H > 0$ and*
(i) $\bar{p}(\theta_m|\theta_n) = \frac{p(\theta_m|\theta_n)r_m}{p(\theta_L|\theta_n)r_L + p(\theta_H|\theta_n)r_H}$, *for* $m, n \in \{L, H\}$
(ii) $r_m = 1$ *for any* m *such that* $U^1_m(\sigma) > U^\bullet_m$ *and* $r_m = 0$ *for any* m *such that* $U^1_m(\sigma) < U^\bullet_m$.

Thus, A^2's beliefs following a rejection of A^1 are required to satisfy a consistency condition similar to the one underlying the definition of perfect sequential equilibrium in Grossman and Perry (1986). In words, the non-deviant A^2 rationalizes a veto of A^1 by finding beliefs about θ^1 which are consistent with A^1's incentive to veto. In our context, if A^1 vetoes S^0 then he is actually vetoing $\sigma^\bullet$. Hence, A^2 should infer that A^1's type is m such that A^1_m will improve his own payoff with respect to $U^\bullet_m$ by not playing $\sigma^\bullet$ in M; if it is common knowledge that σ is played in M after a veto of A^1, that means $U^1_m(\sigma) > U^\bullet_m$.

On the other hand, types m of A^1 who are going to lose if σ is played $[U_m^1(\sigma) < U_j^\bullet]$ should not have vetoed $\sigma^\bullet$; thus they receive zero probability in the revised beliefs of A^2. The belief system $\bar{p}$ of A^2 about θ^1 is consistent with Bayesian updating given the prior beliefs and the above argument, as embodied in refusal probabilities. Observe that σ is required to be a BNE of M when A^2 updated his beliefs about θ^1 as indicated above. The set of types m satisfying $r_m > 0$ is called *credible veto set.*

Definition 5 *An incentive compatible M is strongly ratifiable if no credible veto system exists or if, for any given credible veto system $\bar{p}$ and associated credible veto set, there exists $\hat{\sigma} \in BNE(M_{\bar{p},p})$ such that $U_m^1(\hat{\sigma}) = U_m^\bullet$ for any m in the credible veto set.*

The next proposition establishes that M' when $\rho \leq \rho^*$ and $M''(\varepsilon)$ when $\rho > \rho^*$ are strongly ratifiable, although for different reasons.

Proposition 8 *(i) When $\rho \leq \rho^*$, M' is strongly ratifiable because no credible veto system exists.*
(ii) When $\rho > \rho^*$, there exists a unique credible veto systems for $M''(\varepsilon)$: the credible veto set is $\{\theta_L\}$, $(r_L, r_H) = (1, 0)$, $\bar{p}(\theta_L|\theta_L) = \bar{p}(\theta_L|\theta_H) = 1$, $\bar{p}(\theta_H|\theta_L) = \bar{p}(\theta_H|\theta_H) = 0$ and $\sigma = (HH, HH)$. However, $\sigma^\bullet$ is a BNE of $M''_{\bar{p},p}(\varepsilon)$, $U_L^1(\sigma^\bullet) = U_L^\bullet$ and thus $M''(\varepsilon)$ is strongly ratifiable because $\sigma^\bullet$ plays the role of $\hat{\sigma}$ in definition 5.

The proof of Proposition 8(i) shows that, given M' and $\rho \leq \rho^*$, there exists no behavior of A^2 (let alone an equilibrium) which allows A_m^1 to earn more than $U_m^\bullet$ for some $m \in \{L, H\}$. Thus, there is no credible veto system. In the case of $\rho > \rho^*$ and $M''(\varepsilon)$, there exists a credible veto set and it separates L-type from H-type; the associated BNE of $M''_{\bar{p},p}(\varepsilon)$ is $\sigma \equiv (HH, HH)$ such that all types report θ_H. However, truth-telling is still a BNE in $M''_{\bar{p},p}(\varepsilon)$ and $U_L^1(\sigma^\bullet) = U_L^\bullet$ because A_L^1 has prior beliefs on θ^2. This makes $M''(\varepsilon)$ strongly ratifiable by definition 5.

Although Proposition 8 establishes that M' and $M''(\varepsilon)$ are strongly ratifiable, $\bar{\sigma} \equiv (HH, HH)$ is a BNE of $M''_{p,p}(\varepsilon)$ for $\rho > \rho^*$ and thus, even though each agent accepts S^0, the profile $\bar{\sigma}$ is an equilibrium when S^0 is played. It turns out that there is no Pareto dominance between $\bar{\sigma}$ and $\sigma^\bullet$,[21] but $\bar{\sigma}$ is the unique profile that survives iterated deletion of weakly dominated strategies in $M''(\varepsilon)$.[22] We notice that the optimal collusion proof

[21] An L type's payoff is larger in $\bar{\sigma}$ than in $\sigma^\bullet$, but the opposite is true for an H type.

[22] Truthtelling is strictly dominant for an H type. If A_L^1 and A_L^2 have prior beliefs, as they do after unanimous acceptance of S^0, then A_L^1 is indifferent between reporting L and H *if* A_L^2 plays L (given that A_H^2 chooses H). If instead A_L^2 plays H with positive (even small) probability, then A_L^1 strictly prefers reporting H to reporting L and a similar argument applies to A_L^2.

mechanism in Laffont and Martimort (2000) under strong and positive correlation has the same multiplicity problem as $M''(\varepsilon)$, even though they do not analyze ratifiability and multiplicity when correlation is strong and positive correlation.[23]

5 Conclusion

We considered the regulation of firms producing complementary inputs and studied how correlation of the firms' cost parameters and collusion (under VS) affects the organizational choice between VI and VS when the firms are protected by ex post break-even constraints. First, we found that VI dominates VS for a negative or weakly positive correlation since VI internalizes negative bidding externalities. On the other hand, VS dominates VI for strong positive correlation because of the loss of control effect, i.e. the principal has an incentive to implement a non-monotonic quantity schedule, but can only do so under VS. Second, when the agents collude under complete information, we found that VS is equivalent to VI. Finally, when the agents collude under asymmetric information, collusion does not affect the principal's payoff under VS and therefore does not affect comparison between VI and VS. In particular, when there is strong positive correlation, the principal can implement the non-monotonic optimal quantity schedule in a collusion-proof way, which is impossible if collusion occurs under complete information.

References

Armstrong, M., S. Cowan, and J. Vickers. 1994. "Regulatory Reform: Economic Analysis and British Experience." Cambridge (MA), MIT Press.

Armstrong, M. and J.-C. Rochet. 1999. "Multidimensional Screening: A User's guide." *European Economic Review*, 43: 959-979.

Baron, D. and D. Besanko. 1992. "Information, Control, and Organizational Structure." *Journal of Economics and Management Strategy*, 1 (2): 237-275.

Baron, D. and D. Besanko. 1999. "Informational Alliances." *Review of Economic Studies*, 66 (4): 743-768.

Brocas, I., K. Chan, and I. Perrigne. 2006. "Regulation under Asymmetric Information in Water Utilities." *American Economic Review,* 96 (2): 62-66.

[23] Che and Kim (2006a) consider neither ratifiability nor multiplicity in their analysis.

Che, Y.-K. and J. Kim. 2006a. “Robustly Collusion-Proof Implementation.” *Econometrica*, 74 (4): 1063-1107.

Che, Y.-K. and J. Kim. 2006b. “Optimal Collusion-Proof Auctions.”, mimeo, Columbia University.

Cramton, P., and T.R. Palfrey. 1995. “Ratifiable Mechanisms: Learning from Disagreement.” *Games and Economic Behavior*, 10: 255-283.

Crémer, J. and R. McLean. 1988. “Full Extraction of surplus in Bayesian and Dominant Strategy Auctions.” *Econometrica*, 56: 1247-1258.

Dana, J. D. 1993, “The Organization and Scope of Agents: Regulating Multiproduct Industries.” *Journal of Economic Theory*, 59: 288-310.

Dequiedt, V. 2004. “Optimal Collusion and Optimal Auctions.” Working Paper, INRA-GAEL University of Grenoble.

Dequiedt, V. and D. Martimort. 2004. “Delegated Monitoring Versus Arm’s-Length Contracting.” *International Journal of Industrial Organization*, 22: 951-981.

Garcia, S., M. Moreaux, and A. Reynaud. 2006. “Measuring Economies of Vertical Integration in Network Industries: An Application to the Water Sector.” *International Journal of Industrial Organization*, forthcoming.

Gilbert, R. and M. Riordan. 1995. “Regulating Complementary Products: A Comparative Institutional Analysis.” *Rand Journal of Economics,* 26 (2): 243-256.

Grossman, S., and M. Perry. 1986. “Perfect Sequential Equilibrium,” *Journal of Economic Theory*, 39: 97-119.

Jeon, D.-S. 2005. “Mechanism Design under Collusion and Uniform Transfers.” *Journal of Public Economic Theory*, 7 (4): 641-667.

Jeon, D.-S. and D. Menicucci. 2005. “Optimal Second-degree Price Discrimination and Arbitrage: On the Role of Asymmetric Information among Buyers.” *Rand Journal of Economics,* 36(2).

Laffont, J.-J. and D. Martimort. 1997. “Collusion under Asymmetric Information.” *Econometrica*, 65: 875-911.

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

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

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

Mookherjee, D. and M. Tsumagari. 2004. "The Organization of Supplier Networks: Effects of Delegation and Intermediation." *Econometrica*, 72 (4): 1179-1219.

Pavlov, G. 2004. "Colluding on Participation Decisions." Working Paper, Northwestern University.

Pouyet, J. 2002. "Collusion Under Asymmetric Information: The Role of the Correlation." *Journal of Public Economic Theory*, 4(4): 543-572.

Quesada, L. 2005. "Collusion as an Informed Principal Problem." Working Paper, University of Wisconsin-Madison.

Sappington, D. 1983. "Limited Liability Contracts between Principal and Agent." *Journal of Economic Theory,* 29: 1-21.

Severinov, S. 2005. "The Value of Information and Optimal Organization." Working Paper, Fuqua School of Business, Duke University.

Wolak, F.A. 1994. "An Econometric Analysis of the Asymmetric Information, Regulator-Utility Interaction." *Annales d'Économie et de Statistique,* 34: 13-69.

Appendix

The proofs of proposition 2 and proposition 4 are missing because they are either well known or straightforward.

Proof of proposition 1

In order to maximize (3) subject to (1) and (2) for $m, n \in \{L, H\}$, we consider the relaxed problem in which (BE_{LL}), (BE_{LH}) and (BIC_H) are neglected:

$$\max \ (3) \quad \text{s.t.} \quad (BE_{HL}), \ (BE_{HH}), \ (BIC_L) \tag{17}$$

We verify ex post that (BE_{LL}), (BE_{LH}) and (BIC_H) are satisfied by the solution of (17).

It is simple to see that (BE_{HL}), (BE_{HH}) and (BIC_L) all bind in the solution to (17). For instance, if (BE_{HL}) is slack then it is feasible and profitable to reduce slightly t_{HL}.

Likewise, if (BIC_L) is slack then it pays to reduce slightly t_{LL} and/or t_{LH}. Substitution of the transfers obtained from these binding constraints into (3) gives the program:

$$\begin{aligned} & \max_q \quad p_{LL}[S(q_{LL}) - 2\theta_L q_{LL}] + 2p_{LH}[S(q_{LH}) - (\theta_L + \theta_H + \frac{p_{LL}}{p_{LH}}\Delta)q_{LH}] \\ & \qquad + p_{HH}[S(q_{HH}) - 2(\theta_H + \frac{p_{LH}}{p_{HH}}\Delta)q_{HH}] \\ = & \max_q \quad \sum_{m\in\{L,H\}} \sum_{n\in\{L,H\}} p_{mn}[S(q_{mn}) - \theta^*_{mn} q_{mn}] \end{aligned}$$

The first-order necessary and sufficient conditions are: $S'(q^*_{mn}) = \theta^*_{mn}$ for $m, n \in \{L, H\}$. If $\rho \le \rho^*$, the inequality $\theta^*_{LH} \le \theta^*_{HH}$ holds, which implies $q^*_{LH} \ge q^*_{HH}$. In this case, the transfers mentioned in Proposition 1 (ii) for the case of $\rho \le \rho^*$ make (BE_{HL}), (BE_{HH}) and (BIC_L) binding and it is easy to verify that they also satisfy (BE_{LL}), (BE_{LH}) and (BIC_H). If $\rho > \rho^*$, the inequality $\theta^*_{LH} > \theta^*_{HH}$ holds, which implies $q^*_{LH} < q^*_{HH}$. The transfers mentioned in Proposition 1 (ii) for the case of $\rho > \rho^*$ make (BE_{HL}), (BE_{HH}) and (BIC_L) binding and satisfy (BE_{LL}) and (BE_{LH}). Furthermore, (BIC_H) is satisfied because $t^*_{LH} - \theta_H q_{LH} = 0$ and $t^*_{LL} - \theta_H q_{LL} < 0$. We prove below $t^*_{LL} - \theta_H q_{LL} < 0$. Notice that $\rho > \rho^*$ is equivalent to $p_{HH}p_{LL} > p_{LH}p_{HH} + 2p^2_{LH}$, which implies $p_{LH} < p_{LL}$; then, from $q^*_{LL} > q^*_{HH} > q^*_{LH}$ follows $t^*_{LL} - \theta_H q_{LL} < \theta_L q^*_{LL} + \Delta q^*_{LH} + \Delta(q^*_{HH} - q^*_{LH}) - \theta_H q_{LL} = \Delta(q^*_{HH} - q^*_{LL}) < 0$.

Proof of proposition 3

(i) First, if $\rho < \rho^*$, then the optimal quantity schedule $(q^*_{LL}, q^*_{LH}, q^*_{HH})$ under VS is implementable under VI by using the menu of contracts M^I with $q^I_{LL} = q^*_{LL}$, $q^I_{LH} = q^*_{LH}$, $q^I_{HH} = q^*_{HH}$ and $t^I_{LL} = 2\theta_L q^*_{LL} + \Delta(q^*_{LH} + q^*_{HH})$, $t^I_{LH} = (\theta_H + \theta_L)q^*_{LH} + \Delta q^*_{HH}$, $t^I_{HH} = 2\theta_H q^*_{HH}$. Moreover, the right hand side of (9) is non-negative as $q^*_{LH} \ge q^*_{HH}$. Thus the expected rents the principal needs to pay to implement $(q^*_{LL}, q^*_{LH}, q^*_{HH})$ with M^I under VI are weakly smaller than with M^* of Proposition 1 under VS. This yields:

$$\begin{aligned} W^S(t^*, q^*) &= \sum_{m,n\in\{L,H\}} p_{mn}[S(q^*_{mn}) - (t^*_{mn} + t^*_{nm})] \\ &\le \sum_{m,n\in\{L,H\}} p_{mn}[S(q^*_{mn}) - t^I_{mn}] < \sum_{m,n\in\{L,H\}} p_{mn}[S(q^{*I}_{mn}) - t^{*I}_{mn}] = W^I(t^{*I}, q^{*I}). \end{aligned}$$

where the last inequality follows from the fact that $(q^*_{LL}, q^*_{LH}, q^*_{HH})$ is different from $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$. Therefore, VI dominates VS when $\rho \le \rho^*$.

Second, if $\rho \ge \rho^{**}$, then the optimal quantity schedule $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$ under VI is implementable under VS by using mechanism M with $q_{LL} = q^{*I}_{LL}$, $q_{LH} = q^{*I}_{LH}$, $q_{HH} = q^{*I}_{HH}$

and $t_{HH} = \theta_H q^{*I}_{HH}$, $t_{HL} = \theta_H q^{*I}_{LH}$, $t_{LH} = \theta_L q^*_{LH} + \Delta q^*_{LH}$, $t_{LL} = \theta_L q^*_{LL} + \Delta q^*_{LH} + \frac{p_{LH}}{p_{LL}} \Delta(q^*_{HH} - q^*_{LH})$. In this case the right hand side of (9) is zero because $q^{*I}_{LH} = q^{*I}_{HH}$. This implies that the expected rent payments for implementation of $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$ under VI and VS are equal. However, the principal can strictly increase her payoff under VS with respect to $(q^{*I}_{LL}, q^{*I}_{LH}, q^{*I}_{HH})$ by choosing the quantity schedule $(q^*_{LL}, q^*_{LH}, q^*_{HH})$ with $q^*_{LH} < q^*_{HH}$, as shown in proposition 1. In other words:

$$
\begin{aligned}
W^I(t^{*I}, q^{*I}) &= \sum_{m,n \in \{L,H\}} p_{mn} \left[S(q^{*I}_{mn}) - t^{*I}_{mn} \right] = \sum_{m,n \in \{L,H\}} p_{mn} \left[S(q^{*I}_{mn}) - (t_{mn} + t_{nm}) \right] \\
&< \sum_{m,n \in \{L,H\}} p_{mn} \left[S(q^*_{mn}) - (t^*_{mn} + t^*_{nm}) \right] = W^S(t^*, q^*).
\end{aligned}
$$

in which t^*_{mn} denotes the transfer described by Proposition 1(ii) for the case of $\rho \geq \rho^{**}$. Thus, VS dominates VI when $\rho \geq \rho^{**}$.

(ii) Consider the case in which $S(q) = aq - \frac{b}{2} q^2$ with $a > 0$, $b > 0$ and, in order to avoid corner solutions, suppose that $a = S'(0)$ is larger than the inputs' virtual cost in every state of nature. Notice that we need to consider only the case of $\rho \in (\rho^*, \rho^{**})$, since otherwise the part (i) of the proposition applies. Given that $S'(q) = a - bq$, we find $q^*_{mn} = \frac{1}{b}(a - \theta^*_{mn})$, and $q^{*I}_{mn} = \frac{1}{b}(a - \theta^I_{mn})$ for any $m, n \in \{L, H\}$. Furthermore, $S(\frac{1}{b}(a - x)) - x\frac{1}{b}(a - x) = \frac{1}{2b}(a - x)^2$. Thus, the expected payoff difference between production under VS and under VI equals:

$$
W^S - W^I = \frac{1}{2b} \sum_{m,n \in \{L,H\}} p_{mn} \left[(a - \theta^*_{mn})^2 - \left(a - \theta^I_{mn} \right)^2 \right].
$$

Simple but lengthy manipulations show that $W^S - W^I$ has the same sign as $\rho - \frac{p_{LH}}{3}(4 - 3p_{LH} - 2p_{LL})$. This motivates the definition $\tilde{\rho} \equiv \frac{p_{LH}}{3}(4 - 3p_{LH} - 2p_{LL})$ and it is easy to verify that $\rho^* < \tilde{\rho} < \rho^{**}$ holds.

Proof of proposition 5

Let $\hat{M}$ be an initial grand-mechanism offered by the principal, which satisfies the incentive compatibility, acceptance and ex post break-even constraints. Let $S^\bullet$ be a coalition-interim-efficient side-contract with regard to the reservation utilities given by $\{U^\bullet_L, U^\bullet_L\}$, the payoffs when agents play non-cooperatively $\hat{M}$. Suppose that $S^\bullet$ contingent on the offer of the grand-mechanism $\hat{M}$ gives a payoff $U^i(\theta^i)$ to each agent i with type θ^i. Then, from the interim-efficiency of the side-contract $S^\bullet$, we know that it satisfies the incentive compatibility, acceptance, budget balance and ex post break-even constraints.

Define now a new grand-mechanism $\hat{M}^{\bullet}$ by $\hat{M} \circ S^{\bullet}$. We note that from the interim-efficiency of $S^{\bullet}$, the new grand-mechanism satisfies the incentive compatibility and ex post break-even constraints. We can show that $\hat{M}^{\bullet}$ is weakly collusion-proof. Equivalently, it is optimal for the third-party to offer the null side-contract.

Suppose not, then there exists an interim-efficient side-contract $\bar{S}$ different from the null side-contract, which gives a sum of expected utilities strictly higher than the one achieved by the null side-contract. But this contradicts the coalition-interim-efficiency of $S^{\bullet}$. Suppose $\bar{S}$ contingent on the offer of the grand-mechanism $\hat{M}^{\bullet}$ gives a payoff $\bar{U}^i(\theta^i)$ for each agent. Then, from the interim-efficiency of $\bar{S}$, it satisfies the incentive compatibility, acceptance, budget balance and ex post break-even constraints. In particular, the following inequality holds for each i and θ^i: $\bar{U}^i(\theta^i) \geq U^i(\theta^i)$. Consider now the side-contract $S^{\bullet} \circ \bar{S}$ contingent on the offer of the grand-mechanism $\hat{M}$. Since we have that $\bar{U}^i(\theta^i) \geq U^i(\theta^i) \geq U^{\bullet}_{\theta^i}$, $S^{\bullet} \circ \bar{S}$ can be implemented by the third-party. Moreover, it guarantees strictly higher utility at least for one agent without reducing the other's utility with respect to $\hat{M}$. This contradicts the interim-efficiency of $S^{\bullet}$.

Proof of proposition 6

We show that M' is weakly collusion proof when $\rho \leq \rho^*$ and $M''(\varepsilon)$ (with $\varepsilon > 0$ and small) is weakly collusion proof when $\rho > \rho^*$. In order to prove these results, we examine the third-party's problem in which (BIC^S_H) is neglected:

$$\max_{\phi, y^1, y^2} \sum_{m \in \{L,H\}} \sum_{n \in \{L,H\}} p_{mn}[t^1(\phi_{mn}) + t^2(\phi_{mn}) - (\theta_m + \theta_n) q(\phi_{mn})]$$

subject to

• Budget balance constraints:

$$y^1_{mn} + y^2_{mn} = 0, \quad \text{for any } m, n \in \{L, H\} \tag{18}$$

• Bayesian incentive constraint for L-type agent A^1:

$$\sum_{n \in \{L,H\}} p_{Ln}[t^1(\phi_{Ln}) - y^1_{Ln} - \theta_L q(\phi_{Ln})] \geq \sum_{n \in \{L,H\}} p_{Ln}[t^1(\phi_{Hn}) - y^1_{Hn} - \theta_L q(\phi_{Hn})] \tag{19}$$

• Bayesian incentive constraint for L-type agent A^2:

$$\sum_{n \in \{L,H\}} p_{Ln}[t^2(\phi_{nL}) - y^2_{nL} - \theta_L q(\phi_{nL})] \geq \sum_{n \in \{L,H\}} p_{Ln}[t^2(\phi_{nH}) - y^2_{nH} - \theta_L q(\phi_{nH})] \tag{20}$$

• Acceptance constraint for m-type agent A^1, with $m \in \{L, H\}$:

$$\sum_{n\in\{L,H\}} p_{mn}[t^1(\phi_{mn}) - y^1_{mn} - \theta_m q(\phi_{mn})] \geq (p_{mL} + p_{mH})U^\bullet_m, \tag{21}$$

• Acceptance constraint for m-type agent A^2, with $m \in \{L, H\}$:

$$\sum_{n\in\{L,H\}} p_{nm}[t^2(\phi_{nm}) - y^2_{nm} - \theta_m q(\phi_{nm})] \geq (p_{mL} + p_{mH})U^\bullet_m, \tag{22}$$

• Ex post break-even constraints for an m-type agent A^1 when A^2 has an n-type, with $m, n \in \{L, H\}$:

$$t^1(\phi_{mn}) - y^1_{mn} - \theta_m q(\phi_{mn}) \geq 0, \tag{23}$$

• Ex post break-even constraints for m-type agent A^2 while A^1 has an n-type, with $m, n \in \{L, H\}$:

$$t^2(\phi_{nm}) - y^2_{nm} - \theta_m q(\phi_{nm}) \geq 0, \tag{24}$$

In order to define the Lagrangian function we introduce the following multipliers; we restrict to symmetric multipliers without loss of generality:

• γ_{mn} for the budget-balance constraint if $(\theta^1, \theta^2) = (\theta_m, \theta_n)$ with $m, n \in \{L, H\}$,

• δ for the L-type's Bayesian incentive constraint,

• v_m for the m-type's acceptance constraint with $m \in \{L, H\}$,

• λ_{mn} for the ex post break-even constraint for agent i if $\theta^i = \theta_m$ and $\theta^j = \theta_n$ with $m, n \in \{L, H\}$.

The Lagrangian function is given by:

$$\begin{aligned}\mathcal{L} &= E(U^1 + U^2) + \sum_{i\in\{1,2\}} \delta(BIC^S)^i(\theta_L) + \sum_{m\in\{L,H\}} \sum_{i\in\{1,2\}} v_m(BIR^S)^i(\theta_m) \\ &\quad + \sum_{m,n\in\{L,H\}} \gamma_{mn}(BB)(\theta_m, \theta_n) + \sum_{m,n\in\{L,H\}} \sum_{i\in\{1,2\}} \lambda_{mn}(ExPostBE)^i(\theta_m, \theta_n).\end{aligned}$$

Note that the slackness conditions obtained from the Lagrangian optimization do not give any information on the multipliers associated with the binding constraints in the third party's problem. In particular, (21) and (22) bind for a weakly collusion-proof mechanism for any $m \in \{L, H\}$, and therefore v_L and v_H can take on any non-negative values. The same remark applies to δ, λ_{HH} and λ_{HL} if we consider any grand-mechanism which is optimal under no collusion and weakly collusion-proof. This explains why the principal has some flexibility in choosing μ_{LH} and μ_{HH} introduced below.

First order conditions for side transfers The first order conditions for y^1_{mn} and y^2_{mn} are

$$\begin{aligned} \gamma_{LL} - p_{LL}(\delta + v_L) - \lambda_{LL} &= 0 \quad \gamma_{HH} + \delta p_{LH} - p_{HH}v_H - \lambda_{HH} = 0 \\ \gamma_{LH} + p_{LL}\delta - p_{LH}v_H - \lambda_{HL} &= 0 \quad \gamma_{LH} - p_{LH}(\delta + v_L) - \lambda_{LH} = 0 \quad \gamma_{HL} = \gamma_{LH} \end{aligned} \tag{25}$$

By combining the equations in (25) we obtain

$$v_H = \delta + v_L + \frac{\lambda_{LH} - \lambda_{HL} + p_{LL}\delta}{p_{LH}} \tag{26}$$

The conditions under which $\phi^\bullet(.) = Id$

• *Coalition LL.* Since ϕ_{LL} affects $\mathcal{L}$ through $[p_{LL}(1+\delta+v_L)+\lambda_{LL}][t^1(\phi_{LL})+t^2(\phi_{LL})-2\theta_L q(\phi_{LL})]$, we infer that the optimal ϕ_{LL}, denoted by $\phi^\bullet_{LL}$, is equal to LL if and only if

$$LL \in \arg\max_{\phi_{LL}} [t^1(\phi_{LL}) + t^2(\phi_{LL}) - 2\theta_L q(\phi_{LL})]. \tag{27}$$

• *Coalition LH* (or HL). We find that $\mathcal{L}$ is affected by ϕ_{LH} through

$$\begin{aligned} &(p_{LH} + \lambda_{LH})[t^1(\phi_{LH}) + t^2(\phi_{LH}) - (\theta_L + \theta_H)q(\phi_{LH})] \\ &+[p_{LH}(\delta + v_L) + \lambda_{LH}][t^1(\phi_{LH}) - \theta_L q(\phi_{LH})] \\ &+(\lambda_{HL} - p_{LL}\delta + p_{LH}v_H)[t^2(\phi_{LH}) - \theta_H q(\phi_{LH})] \\ &\quad = [p_{LH}(1 + \delta + v_L) + \lambda_{LH}][t^1(\phi_{LH}) + t^2(\phi_{LH}) - (\theta_L + \theta_H)q(\phi_{LH})] - \delta p_{LL}\Delta q(\phi_{LH}) \end{aligned}$$

where the equality is a consequence of (25). Thus, $\phi^\bullet_{LH} = LH$ is equivalent to

$$LH \in \arg\max_{\phi_{LH}} [t^1(\phi_{LH}) + t^2(\phi_{LH}) - (\theta_L + \theta_H + \mu_{LH})q(\phi_{LH})], \tag{28}$$

with $0 \le \mu_{LH} \equiv \frac{\delta p_{LL}\Delta}{p_{LH}(1+\delta+v_L)+\lambda_{LH}} < \frac{p_{LL}}{p_{LH}}\Delta$.

• *Coalition HH.* Using (26), we see that ϕ_{HH} affects $\mathcal{L}$ through

$$p_{HH}[1+(1+\frac{\rho}{p_{LH}p_{HH}})\delta+v_L+\frac{\lambda_{LH} - \lambda_{HL}}{p_{LH}}+\frac{\lambda_{HH}}{p_{HH}}][t^1(\phi_{HH})+t^2(\phi_{HH})-2\theta_H q(\phi_{HH})]-2p_{LH}\delta\Delta q(\phi_{HH})]$$

Hence, as long as $1 + (1 + \frac{\rho}{p_{LH}p_{HH}})\delta + v_L + \frac{\lambda_{LH}-\lambda_{HL}}{p_{LH}} + \frac{\lambda_{HH}}{p_{HH}} > 0$, $\phi^\bullet_{HH} = HH$ holds if and only if

$$HH \in \arg\max_{\phi_{HH}} [t^1(\phi_{HH}) + t^2(\phi_{HH}) - 2(\theta_H + \mu_{HH}\Delta)q(\phi_{HH})], \tag{29}$$

with $0 \le \mu_{HH} \equiv \frac{\delta p_{LH}\Delta}{p_{HH}[1+(1+\frac{\rho}{p_{LH}p_{HH}})\delta+v_L+\frac{\lambda_{LH}-\lambda_{HL}}{p_{LH}}+\frac{\lambda_{HH}}{p_{HH}}]}$.

Proof that M' is weakly collusion proof when $\rho \leq \rho^*$ We set all the multipliers equal to zero, so that (27)-(29) reduce to (11). We already know that M' satisfies (11), and thus it is weakly collusion proof.

Proof that $M''(\varepsilon)$ is weakly collusion proof when $\rho > \rho^*$ We suppose that $\varepsilon > 0$ is small. Set $\lambda_{LL} = \lambda_{LH} = \lambda_{HL} = \lambda_{HH} = 0$, $v_L = 0$, $v_H = \frac{p_{LH}+p_{LL}}{p_{LH}}\delta$, $\gamma_{LL} = p_{LL}\delta$, $\gamma_{LH} = \gamma_{HL} = p_{LH}\delta$, $\gamma_{HH} = (\frac{\rho}{p_{LH}} + p_{HH})\delta$; notice that these values are consistent with the fact that (BE_{HL}), (BE_{LH}) and (BE_{LL}) are slack in $M''(\varepsilon)$. Finally, we take $\delta > 0$ and very large. This implies that μ_{LH} is close to $\frac{p_{LL}}{p_{LH}}\Delta$ and μ_{HH} is close to $\frac{p_{LH}^2}{p_{HH}p_{LH}+\rho}\Delta$. Then, (27) reduces

to (13)-(14) and (28), (29) are approximately equivalent to

$$t_{HL} + t_{LH} - (\theta_L + \theta_H + \frac{p_{LL}}{p_{LH}}\Delta)q_{LH} \geq 2t_{LL} - (\theta_L + \theta_H + \frac{p_{LL}}{p_{LH}}\Delta)q_{LL} \tag{30}$$

$$t_{HL} + t_{LH} - (\theta_L + \theta_H + \frac{p_{LL}}{p_{LH}}\Delta)q_{LH} \geq 2t_{HH} - (\theta_L + \theta_H + \frac{p_{LL}}{p_{LH}}\Delta)q_{HH} \tag{31}$$

$$2t_{HH} - 2(\theta_H + \frac{p_{LH}^2}{p_{HH}p_{LH} + \rho}\Delta)q_{HH} \geq 2t_{LL} - 2(\theta_H + \frac{p_{LH}^2}{p_{HH}p_{LH} + \rho}\Delta)q_{LL} \tag{32}$$

$$2t_{HH} - 2(\theta_H + \frac{p_{LH}^2}{p_{HH}p_{LH} + \rho}\Delta)q_{HH} \geq t_{HL} + t_{LH} - 2(\theta_H + \frac{p_{LH}^2}{p_{HH}p_{LH} + \rho}\Delta)q_{LH} \tag{33}$$

We find that (13)-(14) are satisfied by $M''(\varepsilon)$ and that (30)-(33) hold strictly. In particular, (33) is equivalent to $-2\frac{p_{LH}^2}{p_{HH}p_{LH}+\rho}\Delta q_{HH}^* \geq \theta_L q_{LH}^* + \Delta q_{HH}^* - \frac{p_{LL}}{p_{LH}}\Delta(q_{HH}^* - q_{LH}^*) - \theta_H q_{LH}^* - 2\frac{p_{LH}^2}{p_{HH}p_{LH}+\rho}\Delta q_{LH}^*$, which reduces to $0 \geq 1 - \frac{p_{LL}}{p_{LH}} + 2\frac{p_{LH}^2}{p_{HH}p_{LH}+\rho}$, or $(p_{LL} - p_{LH})(p_{HH}p_{LH} + \rho) \geq 2p_{LH}^3$. Since $\rho > \rho^*$ is equivalent to $p_{HH}(p_{LL} - p_{LH}) > 2p_{HL}^2$, we infer that $(p_{LL} - p_{LH})(p_{HH}p_{LH} + \rho) > 2p_{LH}^2(p_{LH} + \frac{\rho}{p_{HH}})$, which is larger than $2p_{LH}^3$.

By continuity, therefore, (28) and (29) are satisfied also if $\mu_{LH} - \frac{p_{LL}}{p_{LH}}\Delta$ and $\mu_{HH} - \frac{p_{LH}^2}{p_{HH}p_{LH}+\rho}\Delta$ are different from zero but close to 0.

Proof of proposition 8

It is useful to consider the following payoff matrices for A^1, in which A^1 chooses a row and A^2 chooses a column. In (34), A^1 has an H type; in (35) he has an L type

H-type : (34)

$A^1 \backslash A^2$	L	H
L	$u_{LL}^* - \Delta q_{LL}^*$	$u_{LH}^* - \Delta q_{LH}^*$
H	u_{HL}^*	u_{HH}^*

L-type :

$A^1 \backslash A^2$	L	H
L	u^*_{LL}	u^*_{LH}
H	$u^*_{HL} + \Delta q^*_{LH}$	$u^*_{HH} + \Delta q^*_{HH}$

(35)

Proof that M' **is strongly ratifiable when** $\rho \leq \rho^*$. In M', we know that $u^*_{HL} = u^*_{HH} = 0$, while $u^*_{LL} - \Delta q^*_{LL} = \Delta \frac{(p_{LL}+p_{LH})(q^*_{LH}-q^*_{LL})+p_{LH}(q^*_{HH}-q^*_{LL})}{p_{LL}+2p_{LH}} < 0$ and $u^*_{LH} - \Delta q^*_{LH} = \Delta \frac{2p_{LH}(q^*_{HH}-q^*_{LH})}{p_{LL}+2p_{LH}} < 0$. Furthermore, we have $u^*_{LL} < \Delta q^*_{LH}$, $u^*_{LH} > \Delta q^*_{HH}$ and $u^*_{LL} > u^*_{LH}$.

Therefore, for A^1_H (as well as for A^2_H) it is strictly dominant to report truthfully. This implies that A^1_H cannot gain from vetoing S^0. About A^1_L, we know that A^2_H will play H; A^2_L plays L with probability α and H with probability $1-\alpha$. If $\alpha = 1$, then A^1_L is indifferent between L and H; his payoff is $U^\bullet_L$. If $\alpha < 1$, then A^1_L prefers L because $u^*_{LH} > \Delta q^*_{HH}$ and his payoff is smaller than $U^\bullet_L$ because $u^*_{LL} > u^*_{LH}$. Thus, there is no BNE of M' (let alone beliefs which support it) in which A^1_L earns more than $U^\bullet_L$.

Proof that $M''(\varepsilon)$ **is strongly ratifiable when** $\rho > \rho^*$. In this case we still have $u^*_{LL} - \Delta q^*_{LL} < 0$ and $u^*_{LH} - \Delta q^*_{LH} < u^*_{HH} = 0$, but $u_{HL} = \varepsilon$. Hence, A^2_H certainly plays H and A^1_H's payoff cannot be higher than $U^\bullet_H = \frac{p_{LH}}{p_{LH}+p_{HH}}\varepsilon$. Furthermore, $u^*_{LL} > u^*_{HL} + \Delta q^*_{LH}$ and $u^*_{LH} < \Delta q^*_{HH}$. As above, we denote with α the probability that A^2_L plays L. In order for A^1_L to gain more than U^*_L, it is necessary that $\alpha < 1$ and this makes A^1_L prefer report H over L because $u^*_{LH} < \Delta q^*_{HH}$. In this case, A^1_L plays H and the best reply for A^2_L is to report H. Thus, the unique credible veto system is such that A^1_L rejects S^0 (his payoff is higher after rejection) while A^1_H does not reject because he earns zero by rejecting, which is smaller than $U^\bullet_H$. The updated beliefs are as described in the proposition. However, we can prove that $\sigma^\bullet$ is a BNE of $M''(\varepsilon)$ also with these beliefs: $\sigma^\bullet \in BNE(M''_{\bar{p},p}(\varepsilon))$. Indeed, it is clear that truth-telling is a best reply for A^1_L and A^1_H, since A^1 has prior beliefs on θ^2. Furthermore, since $u^*_{LL} > \varepsilon + \Delta q^*_{LH}$, truth-telling is a best reply also for A^2_L and A^2_H. Hence, $\sigma^\bullet$ is a BNE of $M''_{\bar{p},p}(\varepsilon)$. Since A^1_L has prior beliefs on θ^2, we have that $U^1_L(\sigma^\bullet) = U^\bullet_L$ and we conclude that $M''(\varepsilon)$ is strongly ratifiable.