

---

**Collusion, Delegation and Supervision  
with Soft Information**

**Antoine Faure-Grimaud, Jean-Jacques Laffont, and  
David Martimort**

**USC Center for Law, Economics & Organization  
Research Paper No. C02-9**



**CENTER FOR LAW, ECONOMICS  
AND ORGANIZATION  
RESEARCH PAPER SERIES**

**Sponsored by the John M. Olin Foundation**

University of Southern California Law School  
Los Angeles, CA 90089-0071

*This paper can be downloaded without charge from the Social Science Research Network  
electronic library at [http://papers.ssrn.com/abstract\\_id=279522](http://papers.ssrn.com/abstract_id=279522)*

---

# Collusion, Delegation and Supervision with Soft Information\*

Antoine Faure-Grimaud\*\*  
Jean-Jacques Laffont\*\*\*  
and David Martimort\*\*\*\*

Revised: February 12, 2002

## Abstract

This paper derives an *Equivalence Principle* between organizational forms of supervisory and productive activities. We consider an organization with an agent privately informed on his productivity and a risk averse supervisor getting signals on the agent's type. In a centralized organization, the principal can communicate and contract with both the supervisor and the agent. However, these two agents can collude against the principal. In a decentralized organization, the principal only communicates and contracts with the supervisor who in turn sub-contracts with the agent. We show that the two organizations achieve the same outcome. We discuss this equivalence and provide various comparative statics results to assess the efficiency of supervisory structures.

**Key Words:** Supervision, soft information, collusion, delegation.

**Jel Classification:** D82, G14, G32, L51

\* We thank seminar participants at Boston College, Free University Berlin, Gerzensee ESSET 2001, LSE, New-York University, University of Pennsylvania and Toulouse, the Editor Mark Armstrong and three anonymous referees for their comments. All errors are ours.

\*\* Institute of Management and FMG at the London School of Economics, and CEPR.

\*\*\* IDEI, GREMAQ, Institut Universitaire de France, Toulouse I, France.

\*\*\*\* Université de Toulouse I, GREMAQ, IDEI and CEPR.

# 1 Introduction

This paper analyzes the optimal design of an organization involving both supervisory and productive tasks when these two tasks are performed by different agents. Agents privately know different pieces of information that the principal of the organization tries to elicit to achieve the best possible resource allocation. The productive agent is privately informed of his marginal cost of production. The supervisor observes a soft information signal correlated with this cost. The principal does not observe either piece of information. We analyze the role of supervisory information when the supervisor and the agent can communicate with each other and collude against the principal. Without communication between the productive agent and the supervisor, the Revelation Principle implies that the principal can achieve the best possible outcome by directly contracting and communicating with both agents who adopt a non-cooperative behavior. A centralized organization dominates, at least weakly, any other organizational structure. Whether organizational design matters and how useful supervision remains when collusion between the supervisor and the agent takes place are still two unanswered questions.

To address these issues, we consider two settings. In the *centralized* organization, the principal contracts and communicates simultaneously with the supervisor and the agent. Before reporting to the principal, these two parties can collude. The supervisor makes a take-it-or-leave-it side-contract offer to the agent. Such a side-contract specifies some monetary transfers and a collective manipulation of their individual reports to the principal. Importantly, collusion takes place under asymmetric information. Although the supervisor is better informed than the principal about the agent's type, he does not know the latter's cost parameter. In a *decentralized* structure, the principal contracts only with the supervisor and delegates to him the right to contract in turn with the agent. The information structure is the same as in the centralized organization but the principal has no direct communication with the productive agent.

Our central result is an *Equivalence Principle* between organizational forms. The decentralized organization is *one* possible implementation of the optimal collusion-proof centralized mechanism. Contrary to what would happen absent the threat of collusion, delegating to the supervisor the design of the agent's incentive scheme achieves the best possible outcome.

To understand this result, we first need to discuss how the optimal grand-mechanism in the centralized structure is affected by the possibility of collusion between the supervisor and the agent. Collusion aims at implementing a collective manipulation of the supervisor and the agent's individual reports into the grand-mechanism. Asymmetric information within the coalition creates a trade-off between the ex post efficiency of the collective manipulation and the supervisor's desire of extracting the agent's information rent if the supervisor is only imperfectly informed on the latter. Because of this trade-off, the collective manipulation may be ex post inefficient in some states of nature. The principal benefits from these ex post inefficiencies since they facilitate collusion deterrence. As the agent's status quo utility level obtained from playing the grand-mechanism without colluding increases, the trade-off between ex post efficiency of the collective manipulation and extraction of the agent's rent within the coalition is tilted towards efficiency. The intuition is the same as in a simple principal-agent problem with type-dependent reservation utilities<sup>1</sup>. If the profile of the agent's status quo reservation utilities is sufficiently increasing, the incentive constraints at the side-contracting stage are not binding and the manipulation of reports is no longer distorted to limit the agent's information rents in the coalition. The principal is thus hurt by raising the agent's status quo utility above the minimal amount consistent with the latter's incentive and participation constraints. This instrument, only available if the principal contracts directly with the agent, is thus useless under collusion and the principal can as well delegate to the supervisor the right to design the agent's scheme.

As a by-product, this *Equivalence Principle* between organizational forms yields a number of interesting comparative statics results about the efficiency of supervision with soft information.

**Supervisor's Risk Aversion:** Keeping in mind that the decentralized structure implements the optimal collusion-proof contract, the principal would like to induce the supervisor to choose the same sub-contract as what he would offer himself if he knew the supervisor's signal. This problem has clearly a moral hazard flavor, provided that we reinterpret the moral hazard action as the non-verifiable sub-contract of the agent chosen by the supervisor. The moral hazard literature<sup>2</sup> tells us that providing incentives is costless with risk neutrality. With a risk neutral supervisor, everything happens

---

<sup>1</sup>See Jullien (2000) among others.

<sup>2</sup>See Mirrlees (1999) and Holmstrom (1979) among others. See also the literature on adverse selection and risk aversion in two-tier hierarchies (Salanié (1990) and Laffont and Rochet (1999) for such models).

as if the principal directly receives the supervisory information. Otherwise, there is a trade-off between providing insurance to the supervisor and incentivizing him to offer the sub-contract preferred by the principal.

**Information Structure:** The principal's payoff is shown to be non-monotonic in the accuracy of the supervisor's information. When the supervisor knows exactly the agent's cost, he can offer him a wage just equal to this cost and extract all his rent. The collective manipulation of the coalition is necessarily ex post efficient. The three-tier hierarchy reduces to a standard two-tier hierarchy where the supervisor and the agent are de facto merged. Supervisory information becomes useless. On the other hand, a non-informative signal is also useless. The optimal accuracy of supervisory information is thus interior.

**Bargaining Power in Collusion:** As the agent's bargaining power in side-contracting increases, the collective manipulation of reports moves towards ex post efficiency since the agent's information rent matters less at the side-contracting stage. To improve his payoff, the principal wants thus to increase inefficiency within the coalition and can do so by undoing through the grand-mechanism he offers any bargaining power left to the agent at the side-contracting stage.

This paper is linked to the recent literature on collusion under asymmetric information developed in Laffont and Martimort (1997, 2000). The present focus on supervision and organizational design is new, but we use a similar methodology. In those latter papers, colluding agents are both productive, risk neutral and asymmetrically informed on each other and the collusion is organized by a benevolent uninformed third-party.<sup>3</sup> Here instead, the asymmetry between the preferences and information of the colluding partners suggests a less symmetric treatment within the coalition. The collusive offer is made by the less informed supervisor who is risk averse.<sup>4</sup>

Our model of decentralized contracting fills also a gap between two strands of the literature on hierarchies which have evolved independently over the recent years. On the one hand, Crémer and Riordan (1987), Baron and Besanko (1992), Melumad, Mookherjee

---

<sup>3</sup>Laffont and Martimort (1997) analyze a setting with symmetric agents having independently distributed types. The binding coalition incentive compatibility constraints write as with collusion under symmetric information. Laffont and Martimort (2000) deal instead with the case of a strictly positive correlation between the agents' types. Asymmetric information within the coalition can help the principal because it increases the set of implementable output schedules when the correlation is large enough.

<sup>4</sup>See also Itoh (1993) and Holmstrom and Milgrom (1990) for models of collusion with risk averse agents in pure moral hazard contexts.

and Reichelstein (1992, 1995), Mookherjee and Reichelstein (2001) and Laffont and Martimort (1998) have developed models where decentralized contracting does not entail any welfare loss with respect to the case of centralized contracting when risk neutral agents produce and do not collude. McAfee and McMillan (1995) and Laffont and Martimort (1998) have instead shown that agency costs of delegation may appear when intermediate risk neutral agents are protected by limited liability. Our model differs from those latter works as the intermediate layer of our hierarchy does not produce but only supervises. Introducing risk aversion for the intermediate layer allows us to trace out how agency costs evolve between the two polar cases analyzed by the previous literature. More importantly, our result stresses that the relevant benchmark to assess whether decentralization involves any welfare loss for the principal is not centralized contracting with the agents adopting a non-cooperative behavior but centralized contracting with collusion.<sup>5</sup>

The collusion literature following Tirole (1986)<sup>6</sup> has preferred to view three-tier hierarchies as next of both formal grand-mechanisms offered by a principal to all members of the organization and informal collusive side-contracts linking together agents eager to promote their own goals instead of those of the organization. Following this paradigm, the informal side-contract between a risk neutral supervisor and his supervisee is illegal, purely implicit, being enforced by trust, reciprocity, or through repeated relationships.<sup>7</sup> A first weakness of this approach is that frictions in side-contracting are most often captured by stipulating exogenous transaction costs of side-contracting.<sup>8</sup> Retaining the more tractable assumption of enforceability but, as we do below, introducing asymmetric information creates endogenously some frictions in side-contracting. It shows also how the principal can play on these frictions to undermine the efficiency of collusive behavior. A second weakness of this collusion paradigm is that it takes the organizational structure as given.<sup>9</sup> This literature remains silent on whether a decentralized organization could help

---

<sup>5</sup>In the case of two risk neutral and symmetric productive agents, Laffont and Martimort (2000) show that collusion involves no profit loss for the principal with respect to the case of a non-cooperative behavior between the agents when the collusion-proof grand-mechanism is implemented in Bayesian-Nash equilibrium and the agents' types are independently distributed. Putting together this result with the fact that decentralized contracting involves no loss with respect to centralized contracting and a non-cooperative behavior implies that the decentralized structure entails no profit loss for the principal.

<sup>6</sup>See also Tirole (1992), Kofman and Lawarrée (1993) and Laffont and Tirole (1993, Chapter 11) among others.

<sup>7</sup>Tirole (1992) and Martimort (1999) present some modeling of these self-enforcing collusive behaviors.

<sup>8</sup>See nevertheless Laffont and Meleu (1997) for the role of reciprocity on the transaction costs of side-contracting and Martimort (1999) for a derivation of those transaction costs in a repeated game framework.

<sup>9</sup>Felli (1998) shows nevertheless that some form of delegation helps to deter collusion. In a model with

the principal to reduce collusion. Our Equivalence Principle shows instead that there is no gain for the principal from controlling himself the productive tasks. A decentralized organization can thus be viewed as a particular implementation of the optimal response to the threat of collusion. Notice that we also show in the sequel that the decentralized structure implements the best possible outcome as a *unique* equilibrium. This may be proposed as an argument in favor of a strict preference for decentralization. Baliga and Sjostrom (1998) address delegation issues in a moral hazard environment where risk neutral agents protected by limited liability can collude.<sup>10</sup> As in our model, information sets are nested along the hierarchy: only one agent can observe the productive effort of the other. They show that decentralization can implement the optimal collusion-proof contract.

Section 2 presents the model and discusses the two organizational forms considered in this paper. Section 3 analyzes several benchmarks and, in particular, the case where the supervisor and the agent do not collude against the principal. Section 4 describes the set of collusion-proof grand-mechanisms under centralized contracting and derives the optimal collusion-proof mechanism. Section 5 shows how a decentralized organization is a particular implementation of this outcome. We prove there the Equivalence Principle. Section 6 provides some comparative statics. All proofs are in an Appendix.

## 2 The Model

### 2.1 Players and Information

We consider a three agent organization involving a principal, a supervisor and a productive agent. Technology imposes separation between ownership, supervision and production. The agent produces a quantity  $q$  of output at a constant marginal cost  $\theta$  which is his own private information. Types are drawn from a discrete distribution on  $\Theta = \{\theta_1, \theta_2\}$  (we denote  $\Delta\theta = \theta_2 - \theta_1 > 0$ ). The supervisor is uninformed about the agent's type. Nonetheless, he receives a signal  $\tau$  on the agent's marginal cost.  $\tau$  is drawn from a discrete distribution on  $T = \{\tau_1, \tau_2\}$ . This signal is observed by both the supervisor and

---

exogenous frictions in side-contracting, Baliga (1999) describes a mechanism which allows the principal to get supervisory information even though this information is soft. This paper shares therefore with ours the idea that transaction costs help to prevent collusion.

<sup>10</sup>See also Macho-Stadler and Perez-Castrillo (1994).

the agent. Hence, informational sets are nested along the hierarchy: nature reveals to the agent both his type and the supervisor's information; only the latter is available to the supervisor while the principal observes none of these pieces of information.<sup>11</sup> The joint probabilities on  $(\theta_i, \tau_j)$  are defined as  $p_{ij} = Prob(\theta = \theta_i, \tau = \tau_j)$  with  $p_{ij} > 0$  for all  $i, j$ . From the joint distribution above, one can derive the conditional probabilities  $p(\theta_i|\tau_j)$ . There is a positive correlation between signals and types when the monotone likelihood ratio property is satisfied:  $\frac{p(\theta_1|\tau_1)}{p(\theta_2|\tau_1)} = \frac{p_{11}}{p_{21}} \geq \frac{p(\theta_1|\tau_2)}{p(\theta_2|\tau_2)} = \frac{p_{12}}{p_{22}}$ .<sup>12</sup>

## 2.2 Preferences

The agent is risk neutral<sup>13</sup> and has a utility function  $U = t - \theta q$ , where  $t$  is the monetary transfer he receives either from the principal under centralized contracting or from the supervisor under decentralized contracting. The agent accepts to produce as long as he gets his reservation utility exogenously normalized to zero.

The supervisor is risk averse and has a CARA utility function defined over his monetary payoffs:  $V = v(x) = \frac{1}{r}(1 - e^{-rx})$ <sup>14</sup>. Under centralized contracting, the supervisor's income is the wage  $s$  he receives from the principal. Under decentralized contracting, one must subtract from  $s$  the agent's transfer to get the supervisor's income.

Producing  $q$  units of output yields revenue  $R(q)$  to the principal with  $R'(\cdot) > 0$  and  $R''(\cdot) < 0$ . To ensure always positive production levels, we assume that the Inada conditions  $R'(0) = +\infty$  and  $R'(+\infty) = 0$  hold. The principal's profit is given by  $\Pi = R(q) - s - t$  under centralized contracting and  $\Pi = R(q) - s$  under decentralized contracting.

## 2.3 Organizations and Contracts

**Centralized Contracting:** In a centralized organization, the principal directly contracts and communicates with both the supervisor and the agent. A *grand-mechanism* ruling

---

<sup>11</sup>Nested information structures are standard in both the literatures on collusion and on delegation in hierarchies (see Tirole (1986, 1992) and McAfee and McMillan (1995) among others).

<sup>12</sup>An example of such an information structure is as follows: the agent has a low (resp. high) cost  $\theta_1$  (resp.  $\theta_2$ ) with probability  $\nu$  (resp.  $1 - \nu$ ) and the conditional probabilities of the signal are  $p(\tau_1|\theta_1) = p(\tau_2|\theta_2) = \epsilon$ . Then,  $p_{11} = \nu\epsilon$ ,  $p_{21} = (1 - \nu)(1 - \epsilon)$ ,  $p_{12} = \nu(1 - \epsilon)$  and  $p_{22} = (1 - \nu)\epsilon$ .  $\epsilon$  ( $\geq \frac{1}{2}$ ) can be viewed of as the signal's precision. For  $\epsilon$  close to  $\frac{1}{2}$ , the signal conveys little information about the agent's type while, if  $\epsilon = 1$ , the supervisor exactly observes the agent's type.

<sup>13</sup>Our results would be the same with a risk averse agent since his ex post participation and incentive constraints would be identical and only those constraints are relevant for the analysis.

<sup>14</sup> $r = 0$  corresponds to the limiting case where the supervisor is risk neutral.



the organization is a triplet  $GC = \{t(m_s, m_a), s(m_s, m_a), q(m_s, m_a)\}$  stipulating monetary transfers respectively for the agent and the supervisor and output targets as a function of the supervisor's and the agent's messages (denoted respectively  $m_s$  and  $m_a$ ) which belong respectively to two message spaces  $M_s$  and  $M_a$ .

A centralized organization may be subject to coalition formation between the supervisor and the agent. The supervisor has all the bargaining power at the collusion stage.<sup>15</sup> The supervisor, knowing the realization of  $\tau$ , makes a take-it-or-leave-it side-offer to the agent. This side-contract is a pair  $SC_\tau = \{\phi_\tau(\cdot), y_\tau(\cdot)\}$  where  $\phi_\tau(\cdot)$  is a collective manipulation of the messages  $(m_s, m_a)$  sent to the principal and  $y_\tau(\cdot)$  is a side-transfer from the supervisor to the agent.  $\phi_\tau(\cdot)$  maps the agent's report to the supervisor into the set  $\Delta(M_a \times M_s)$  of measures on messages sent to the principal. The Revelation Principle applies at the side-contracting stage, and there is no loss of generality in assuming that  $SC_\tau$  is a direct mechanism. As a by-product of the collusion-proofness principle shown in the Appendix, there is also no loss of generality in restricting the principal to offer direct grand-mechanisms.

Following the earlier literature on collusion, the side-contract is fully enforceable. This is of course a simplifying assumption which yields an upper bound on what can be achieved by the collusion between the supervisor and the agent in a centralized organization. This assumption also implies that the supervisor is able to commit to the collective manipulation of reports and the side-transfers proposed to the agent. As in standard principal-agent models, this commitment is crucial to make credible any distortion away from ex post efficiency in the manipulation of reports  $\phi_\tau(\cdot)$  which is needed to solve the asymmetric information problem within the coalition.

**Decentralized Contracting:** The principal contracts only with the supervisor who, in turn, sub-contracts with the agent. With respect to the case of centralized contracting, a grand-mechanism is now restricted in two ways. First, the agent receives no transfer from the principal,  $t = 0$ . Second, the agent does not communicate directly with the principal so that  $M_a = \emptyset$ . However, the supervisor can use a message space  $M_s$  to communicate with the principal which may be larger than  $T$  since he may have to communicate information on the agent's type that he will obtain from sub-contracting. A *sub-contract*<sup>16</sup> links now

---

<sup>15</sup>Section 6.3 will drop this assumption and show that our results are robust as long as the supervisor retains some bargaining power.

<sup>16</sup>The terminology *sub-contract* is used to distinguish this formal contract from the secret side-contract of the centralized organization.

the supervisor and the agent. This sub-contract writes as  $SC_\tau = \{\phi_\tau(\cdot), y_\tau(\cdot)\}$ .  $y_\tau(\cdot)$  is the agent's transfer received from the supervisor,  $\phi_\tau(\cdot)$  is the manipulation of the supervisor's report to the principal. The Revelation Principle applies at the sub-contracting stage and there is no loss of generality in restricting the analysis to sub-contracts which are direct truthful mechanisms.<sup>17</sup>

## 2.4 Timings and Comparison of the Different Contracting Games

The timings of the contractual games under both organizational forms are as follows.

- The agent learns  $\theta$  and  $\tau$ . The supervisor learns only  $\tau$ .
- The principal offers a grand-mechanism to the supervisor and the agent (resp. to the supervisor only) under centralized (resp. decentralized) contracting.<sup>18</sup>
- The supervisor and the agent both accept or refuse the grand-mechanism under centralized contracting. If any of them refuses, the game ends. Under decentralized contracting, only the supervisor is asked to accept the grand-mechanism. Again, if he refuses, the game ends.
- Under centralized (resp. decentralized) contracting, the supervisor offers a collusive side-contract (resp. a sub-contract) to the agent. The agent accepts or refuses this contract. If he refuses, the grand-mechanism is played non-cooperatively by the supervisor and the agent under centralized contracting and no production takes place under decentralized contracting.
- Production and transfers in all mechanisms take place.

With both organizations, the acceptance of the grand-mechanism by the supervisor takes place before learning the agent's type during side-contracting. Hence, the supervisor's interim participation constraints must be satisfied by the grand-mechanism. Instead,

---

<sup>17</sup>Finally, note that the agent refuses de facto the decentralized grand-mechanism if he refuses the sub-contract.

<sup>18</sup>Suppose that the principal offers the contract before the agents learn their information. If the agent is infinitely risk averse below zero wealth, his ex post participation constraints must still be satisfied. The risk averse supervisor's interim participation constraints would be replaced by an ex ante participation constraint. Similar output distortions to those we obtain below would appear but no rent would be given up to the supervisor.

since the agent is informed on his own type and on the supervisor’s signal at the time of accepting the side-contract, the agent’s ex post participation constraints must be satisfied.

The key aspect in which centralized contracting differs from decentralized contracting is that only in the former the principal directly communicates and contracts with the agent. The principal can thus affect the status quo payoffs obtained by the different types of agent when they refuse the collusive side-contract offered by the supervisor. Since the agent’s possible acceptance of the side-contract is made knowing his type and the supervisor’s signal, any side-contract which guarantees these status quo payoffs must satisfy some ex post participation constraints which are not only type-dependent but also dependent on the non-cooperative play of the grand-mechanism  $GC$ . Technically, only the participation constraints differ at the side-contracting stage between both organizational forms. The agent’s status quo payoff under decentralized contracting is always zero.

The assumption of perfect enforceability of the side-contract makes the comparison between organizational forms more meaningful as, in the latter form, sub-contracts are legally enforceable. It also allows to isolate the impact of asymmetric information on the efficiency of side-contracting.

### 3 Benchmarks

• **Direct Supervision:** Let us first consider the case where the principal directly receives the signal  $\tau$  on the agent’s private information. Using the Revelation Principle, there is no loss of generality in looking for the optimal contract within the class of direct truthful revelation mechanisms of the form  $\{t_\tau(\hat{\theta}); q_\tau(\hat{\theta})\}$  where  $\hat{\theta}$  is the agent’s report on his efficiency parameter to the principal. For ease of notations, we denote thereafter by  $t_{ij} = t_{\tau_j}(\theta_i)$ ,  $q_{ij} = q_{\tau_j}(\theta_i)$  the agent’s transfer and output when he reports  $\theta_i$  and the principal knows  $\tau_j$ . We define also by  $u_{ij} = t_{ij} - \theta_i q_{ij}$  the agent’s information rent in state  $(\theta_i, \tau_j)$ . As it is standard in two-type adverse selection models, the following constraints are of particular importance:<sup>19</sup>

• Incentive compatibility constraints for an efficient agent:

$$u_{1j} \geq u_{2j} + \Delta\theta q_{2j}, \quad \text{for } j = 1, 2, \tag{1}$$

---

<sup>19</sup>When the following constraints are binding, as it will be the case at the optimum of the principal’s problem, it is easy to show that the remaining constraints are strictly satisfied.

- Participation constraints for an inefficient agent:

$$u_{2j} \geq 0, \quad \text{for } j = 1, 2. \quad (2)$$

When he has observed a signal  $\tau_j$ , the principal updates his beliefs on the agent's type. Conditional probabilities become  $p(\theta_1|\tau_j) = \frac{p_{1j}}{p_{1j}+p_{2j}}$  for  $j = 1, 2$ . Accordingly, the optimal contract solves:

$$\begin{aligned} \max_{\{q_{1j}, q_{2j}, u_{1j}, u_{2j}\}} & p(\theta_1|\tau_j)(R(q_{1j}) - \theta_1 q_{1j} - u_{1j}) + p(\theta_2|\tau_j)(R(q_{2j}) - \theta_2 q_{2j} - u_{2j}) \\ & \text{subject to (1) and (2)}. \end{aligned}$$

Solving this problem yields the *conditionally optimal* second-best grand-mechanism  $GC^{sb}$ . This mechanism implements the first best outputs  $q_{1j}^{sb} = q_1^{fb}$  for an efficient agent and outputs  $q_{2j}^{sb}$  for an inefficient one where:

$$R'(q_1^{fb}) = \theta_1 \quad (3)$$

$$R'(q_{2j}^{sb}) = \theta_2 + \frac{p_{1j}}{p_{2j}} \Delta\theta. \quad (4)$$

To reduce the cost of the efficient agent's incentive compatibility constraint (1) and to make less valuable for an efficient agent to mimic an inefficient one, the principal reduces the output produced by an inefficient agent. A positive rent is left to the efficient agent ( $u_{1j}^{sb} = t_{1j}^{sb} - \theta_1 q_{1j}^{sb} = \Delta\theta q_{2j}^{sb}$ ) while the participation constraint (2) of an inefficient agent is binding ( $u_{2j}^{sb} = t_{2j}^{sb} - \theta_2 q_{2j}^{sb} = 0$  for all  $j$ ). The monotone likelihood ratio property implies that the inefficient agent's output is more distorted after the observation of  $\tau_1$  than after the observation of  $\tau_2$ :

$$q_{21}^{sb} < q_{22}^{sb}. \quad (5)$$

Indeed, the agent is more likely to be efficient when  $\tau_1$  is observed than when  $\tau_2$  is observed. Reducing the efficient agent's information rent calls then for a greater allocative inefficiency of the inefficient agent's output.

- **Uninformative Supervision:** Let us now consider the polar case where the supervisor never gets any signal on the agent. Again, output distortions only concern the inefficient agent. It is immediate to show that the optimal quantity for an inefficient agent is the unconditional second-best  $q_2^p$  defined as:

$$R'(q_2^p) = \theta_2 + \frac{p_{11} + p_{12}}{p_{21} + p_{22}} \Delta\theta. \quad (6)$$

Only the efficient agent gets a strictly positive rent ( $u_{1j}^p = \Delta\theta q_2^p$ ,  $u_{1j}^p = 0$ ).

• **Non-Cooperative Implementation and Centralized Contracting:** Applying the Revelation Principle, there is no loss of generality in restricting the principal to use direct truthful revelation mechanisms when the supervisor and the agent do not collude. Let us denote by  $s_{ijk}$  (resp.  $t_{ijk}$  and  $q_{ijk}$ ) the supervisor's wage (resp. the agent's transfer and the output target) when the agent reports that he has type  $\theta_i$  and that the supervisor's signal is  $\tau_k$  and when the supervisor reports he has observed  $\tau_j$ . To simplify notations, we also write  $s_{ijj} = s_{ij}$  (resp.  $t_{ijj} = t_{ij}$ ).

Because  $\tau$  is a piece of information which is commonly known by the supervisor and the agent, the logic of Nash implementation applies.<sup>20</sup> The principal can costlessly elicit this signal by building a revelation scheme such that the agent and the supervisor truthfully report  $\tau$  to the principal. The agent's incentive constraints can be reduced to the following relevant incentive constraints:

$$u_{ij} \geq t_{ij} - \theta_i q_{ij} \quad \text{for all } (i, i', j). \quad (7)$$

The optimal contracting outcome with a non-cooperative behavior is thus the conditionally optimal outcome. The risk averse supervisor is perfectly insured and gets zero wage  $s_{ij}^{sb} = 0$  for all  $(i, j)$  in this contract. If the principal can perfectly control and forbid communication between the agent and the supervisor, he can thus achieve the same outcome as with direct supervision. Importantly, this result is independent of the supervisor's degree of risk aversion when the agents do not collude.

• **Unique Implementation:** A clever design of the out-of-equilibrium wages offered to the supervisor ensures also unique Nash implementation. The supervisor can be used to break any unwanted equilibrium. Suppose indeed that the principal offers wage  $s_{ijk}^{sb}$  for  $j \neq k$  such that:

$$p(\theta_1|\tau_1)v(s_{112}^{sb}) + p(\theta_2|\tau_1)v(s_{212}^{sb}) > 0 > p(\theta_1|\tau_2)v(s_{112}^{sb}) + p(\theta_2|\tau_2)v(s_{212}^{sb}) \quad (8)$$

and

$$p(\theta_1|\tau_2)v(s_{121}^{sb}) + p(\theta_2|\tau_2)v(s_{221}^{sb}) > 0 > p(\theta_1|\tau_1)v(s_{121}^{sb}) + p(\theta_2|\tau_1)v(s_{221}^{sb}). \quad (9)$$

The first left-hand side inequality in (8) says that the supervisor prefers to report the true state of nature  $\tau_1$  when the agent reports instead  $\tau_2$ . This makes impossible to sustain a

---

<sup>20</sup>See Maskin (1999).

non-truthful equilibrium when  $\tau_1$  realizes. On the other hand, the second right-hand side inequality in (8) ensures that the supervisor does not want to lie when  $\tau_2$  realizes and the agent reports this state of nature truthfully to the principal. It is easy to check that the monotone likelihood ratio property  $\frac{p_{11}}{p_{21}} > \frac{p_{12}}{p_{22}}$  ensures that the indifference curves of the supervisor in different states of nature only cross once and that such punishments and rewards exist.<sup>21</sup> In a similar vein, (9) ensures that a non-truthful equilibrium does not exist when  $\tau_2$  realizes and that the supervisor's prefers to tell the truth when  $\tau_1$  realizes.<sup>22</sup>

• **Non-Robustness to Collusion:** The second-best mechanism  $GC^{sb}$  is not robust to collusion. Consider the following side-contract  $SC^{sb}$ :

- $\{\phi_{\tau_j}^{sb}(\theta_i) = (\theta_i, \tau_j); y_{ij}^{sb} = 0\}$  for all  $(i, j) \neq (1, 1)$ .
- $\{\phi_{\tau_1}^{sb}(\theta_1) = (\theta_1, \tau_2); y_{11}^{sb} = -\Delta\theta(q_{22}^{sb} - q_{21}^{sb})\}$  for  $(i, j) = (1, 1)$ .

The composition of the side- and the grand-mechanism is incentive compatible. The efficient agent is just indifferent between telling the truth or lying to the supervisor since the following incentive constraints hold:

$$\begin{aligned} t_{12}^{sb} - \theta_1 q_{12}^{sb} + y_{11}^{sb} &= \Delta\theta q_{22}^{sb} - \Delta\theta(q_{22}^{sb} - q_{21}^{sb}) = \Delta\theta q_{21}^{sb} \\ &\geq (=) t_{21}^{sb} - \theta_1 q_{21}^{sb} \quad \text{when } \tau_1 \text{ realizes and,} \\ t_{12}^{sb} - \theta_1 q_{12}^{sb} &= \Delta\theta q_{22}^{sb} \geq (=) t_{22}^{sb} - \theta_1 q_{22}^{sb} \quad \text{when } \tau_2 \text{ realizes.} \end{aligned}$$

Similarly, the inefficient agent strictly prefers to tell the truth to the supervisor since:

$$\begin{aligned} t_{21}^{sb} - \theta_2 q_{21}^{sb} &= 0 > t_{12}^{sb} - \theta_2 q_{12}^{sb} + y_{11} = \Delta\theta(q_{21}^{sb} - q_1^{fb}) \quad \text{when } \tau_1 \text{ realizes and,} \\ t_{22}^{sb} - \theta_2 q_{22}^{sb} &> t_{12}^{sb} - \theta_2 q_{12}^{sb} \quad \text{when } \tau_2 \text{ realizes.} \end{aligned}$$

The composition of the side- and the grand-mechanism is also individually rational for the agent who prefers this composition to the non-cooperative play of  $GC^{sb}$ . This is of course the case for the inefficient agent who gets zero whether he colludes or not. Instead,

<sup>21</sup>In other words, the decision rule is uniquely Nash implementable.

<sup>22</sup>Since the agent's utility function does not depend on  $\tau$  directly, the agent cannot be used to break unwanted equilibria. Moreover, we impose  $t_{ijk} = -q_{ijk} = -\infty$  for  $j \neq k$  for the agent's punishments so that there cannot be any non-truthful equilibrium where the agent deviates both on his reports of type and signal.

if he refuses the side-contract and the truthful Nash equilibrium of the grand-mechanism  $GC^{sb}$  is played, the efficient agent gets

$$t_{12}^{sb} - \theta_1 q_{12}^{sb} + y_{11}^{sb} = \Delta\theta q_{21}^{sb} \geq (=) t_{11}^{sb} - \theta_1 q_{11}^{sb} = \Delta\theta q_{21}^{sb}$$

when  $\tau_1$  realizes and his utility in state  $\tau_2$  is unchanged. Finally, the supervisor benefits strictly from the manipulation of reports and gets a strictly positive expected payoff when  $\tau_1$  has been observed while he gets zero without collusion:

$$p(\theta_1|\tau_1)v(\Delta\theta(q_{22}^{sb} - q_{21}^{sb})) + p(\theta_2|\tau_1)v(0) > 0.$$

In the next section, we look for the optimal centralized mechanism immune to this kind of collusion.

## 4 Centralized Contracting

We start by deriving the optimal side-contract for any given grand-mechanism  $GC$ . Then, we obtain a *Collusion-Proofness Principle* which allows us to restrict the analysis to grand-mechanisms which are robust to side-contracting. Collusion-proof mechanisms define status quo payoffs such that the null side-contract is optimal given these reservation payoffs. Then, we characterize those collusion-proof grand-mechanisms with a set of simple coalition incentive constraints. Finally, we optimize within this set to find the principal's optimal grand-mechanism.

### 4.1 Collusion-Proof Grand-Mechanisms

Let us denote by  $U_{ij}$  the agent's status quo payoff when his type is  $\theta_i$ , and when the supervisor has observed signal  $\tau_j$ , and they play non-cooperatively the truthful equilibrium of an individually incentive compatible grand-mechanism  $GC$ . If any of the agents refuses the null side-contract and the partners end up behaving non-cooperatively, they would still be punished by this grand-mechanism when their reports on  $\tau$  to the principal differ. As we discuss below, punishments can still be designed to ensure unique implementation if the agents fail to cooperate provided that the supervisor does not change his beliefs on the agent following the latter's refusal of collusion. In what follows, we will thus assume that side-contracting is sustained with those passive beliefs.<sup>23</sup>

---

<sup>23</sup>See Laffont and Martimort (1997, 2000) for a similar use of passive beliefs and below for further remarks on the case of other out-of-equilibrium beliefs.

By definition, we have thus  $U_{ij} = t_{ij} - \theta_i q_{ij}$ . The agent's information rent obtained from truthfully playing the side-contract is instead:  $u_{ij} = y_{ij} + t(\phi_{\tau_j}(\theta_i)) - \theta_i q(\phi_{\tau_j}(\theta_i))$  where  $\phi_{\tau_j}(\theta_i)$  is the manipulation of reports induced by the collusive side-contract when the agent reports having type  $\theta_i$  to the supervisor and the latter has observed  $\tau_j$ .<sup>24</sup> Acceptance of the side-contract by both types of agent imposes thus the following type-dependent ex post participation constraints:

$$u_{1j} \geq U_{1j} \quad \text{for } j = 1, 2, \quad (10)$$

$$u_{2j} \geq U_{2j} \quad \text{for } j = 1, 2. \quad (11)$$

#### 4.1.1 Optimal Side-Contracts under Asymmetric Information

We focus in what follows on side-contracts such that only the efficient agent's incentive constraint may be binding at the optimum of the supervisor's problem.<sup>25</sup> The optimal side-contract is thus solution to the following problem:

$$\begin{aligned} \max_{\{\phi_{\tau_j}(\cdot), u_{ij}\}} & p(\theta_1|\tau_j)v(s(\phi_{\tau_j}(\theta_1)) + t(\phi_{\tau_j}(\theta_1)) - \theta_1 q(\phi_{\tau_j}(\theta_1)) - u_{1j}) \\ & + p(\theta_2|\tau_j)v(s(\phi_{\tau_j}(\theta_2)) + t(\phi_{\tau_j}(\theta_2)) - \theta_2 q(\phi_{\tau_j}(\theta_2)) - u_{2j}) \\ & \text{subject to (1)-(10) and (11).} \end{aligned}$$

where, in the right-hand side of (1),  $q_{2j}$  is now replaced by  $q(\phi_{\tau_j}(\theta_2))$ .

The supervisor reduces the side-transfer  $y_{2j}$  given to the inefficient agent to the point where the ex post participation constraint (11) is binding and thus  $u_{2j} = U_{2j}$ . Consequently, the optimal manipulation function  $\phi_{\tau_j}^*(\cdot)$  and the efficient agent's rents  $u_{1j}^*$  are solutions to a reduced problem with (1) and (10) as the only relevant constraints. Multiplying the objective function by  $p_{1j} + p_{2j} > 0$ , the supervisor's problem (S) becomes:

$$\begin{aligned} (S) : \quad & \max_{\{\phi_{\tau_j}(\cdot), u_{ij}\}} p_{1j}v(s(\phi_{\tau_j}(\theta_1)) + t(\phi_{\tau_j}(\theta_1)) - \theta_1 q(\phi_{\tau_j}(\theta_1)) - u_{1j}) \\ & + p_{2j}v(s(\phi_{\tau_j}(\theta_2)) + t(\phi_{\tau_j}(\theta_2)) - \theta_2 q(\phi_{\tau_j}(\theta_2)) - U_{2j}) \\ & \text{subject to (1) and (10).} \end{aligned}$$

---

<sup>24</sup>This manipulation being possibly stochastic,  $t(\phi_{\tau_j}(\cdot))$  (resp.  $q(\phi_{\tau_j}(\cdot))$ ) should be viewed as an expectation with respect to the distribution of reports in the grand-mechanism induced by  $\phi_{\tau_j}(\cdot)$ .

<sup>25</sup>This is a standard feature of adverse selection models as long as  $q(\phi_{\tau_j}(\theta_1)) \geq q(\phi_{\tau_j}(\theta_2))$  and  $U_{1j} - U_{2j}$  not too large (no countervailing incentives coming from type-dependent participation constraint).



Let us denote by  $\lambda_j$  and  $\mu_j$  the respective positive multipliers of the two constraints of the associated Lagrangian.

- Optimizing with respect to  $u_{1j}$  yields the following relationship between the supervisor's marginal utility of income and the multipliers of (1) and (10):

$$p_{1j}v'(s(\phi_{\tau_j}^*(\theta_1)) + t(\phi_{\tau_j}^*(\theta_1)) - \theta_1 q(\phi_{\tau_j}^*(\theta_1)) - u_{1j}^*) = \lambda_j + \mu_j. \quad (12)$$

- Optimizing with respect to  $\phi_{\tau_j}^*(\theta_1)$  and  $\phi_{\tau_j}^*(\theta_2)$ , we obtain respectively:

$$\phi_{\tau_j}^*(\theta_1) \in \operatorname{argmax}_{\tilde{\phi} \in \Delta(M_a \times M_s)} s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_1 q(\tilde{\phi}),^{26} \quad (13)$$

$$\phi_{\tau_j}^*(\theta_2) \in \operatorname{argmax}_{\tilde{\phi} \in \Delta(M_a \times M_s)} p_{2j}v(s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_2 q(\tilde{\phi}) - U_{2j}) - \lambda_j \Delta \theta q(\tilde{\phi}). \quad (14)$$

- Finally, the slackness conditions tell us that  $\lambda_j = 0$  (resp.  $\mu_j = 0$ ) when (1) (resp. when (10)) is slack in  $(S)$ .

When the multiplier  $\lambda_j$  of (1) is positive,  $\phi_{\tau_j}^*(\theta_2)$  may thus differ from the ex post optimal collective manipulation. Instead,  $\phi_{\tau_j}^*(\theta_1)$  is always ex post efficient.

#### 4.1.2 The Collusion-Proofness Principle

There is no loss of generality in restricting the analysis to *collusion-proof grand-mechanisms* such that the optimal side-contract proposed by the supervisor to the agent and accepted by both types of agent entails no manipulation of reports (namely,  $\phi_{\tau_j}^*(\theta_i) = (\theta_i, \tau_j)$ , for all  $(i, j)$ ) and zero side-transfers ( $y_{ij}^* = 0$  for all  $(i, j)$ ). The principal designs now the grand-mechanism  $GC$  so that the supervisor and the agent do not collectively manipulate their reports to the principal. The logic of the argument is the same as that of the Revelation Principle and the proof is relegated to the Appendix. Note that a collusion-proof mechanism is such that both (10) and (11) are binding in  $(S)$  since  $y_{ij} = 0$  and  $\tilde{\phi}^*(\cdot) = Id$ . This observation helps to characterize the collusion-proof mechanisms.

#### 4.1.3 Coalition Incentive Constraints

Using the CARA specification makes the analysis tractable and allows us to easily rewrite equations (13) and (14). Reports being truthful for collusion-proof grand-mechanisms,

---

<sup>26</sup>The  $\phi_{\tau_j}^*(\cdot)$  are measures (we allow for stochastic side-mechanisms) and the argmax above should be viewed as saying that any manipulation in the support of  $\phi_{\tau_j}^*(\cdot)$  must maximize the right-hand side.

we obtain the following characterization of the *coalition incentive constraints* satisfied by those grand-mechanisms.

**Proposition 1** : *A grand-mechanism GC is collusion-proof if and only if,  $M_s = T$ ,  $M_a = \Theta \times T$ , (1) and (2) both hold, and there exist  $\lambda_j$  for  $j = 1, 2$  such that the following coalition incentive constraints are also satisfied:*

$$(\theta_1, \tau_j) \in \arg \max_{\tilde{\phi} \in \Theta \times T} s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_1 q(\tilde{\phi}), \quad \text{for all } j = 1, 2 \quad (15)$$

$$(\theta_2, \tau_j) \in \arg \max_{\tilde{\phi} \in \Theta \times T} - \left( \lambda_j r \Delta \theta q(\tilde{\phi}) + p_{2j} e^{-r(s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_2 q(\tilde{\phi}) - u_{2j})} \right) \quad \text{for all } j = 1, 2. \quad (16)$$

with

$$0 \leq \lambda_j \leq p_{1j} e^{-r(s_{1j} + t_{1j} - \theta_1 q_{1j} - u_{1j})}. \quad (17)$$

Let us define the coalition's aggregate payoff as  $w_{ij} = s_{ij} + t_{ij} - \theta_i q_{ij}$ . The coalition incentive constraints ensuring that this coalition does not manipulate reports to the principal can then easily be derived with these new variables. From two revealed-preference arguments using (15) respectively when  $\tau = \tau_1$  and  $\tau = \tau_2$ , we obtain immediately:

$$w_{11} = w_{12}. \quad (18)$$

The optimal collusion-proof grand-mechanism cannot screen apart two coalitions involving an efficient agent but different supervisory signals with respect to their aggregate payoffs.

For a given value of the supervisory information  $\tau_j$ , (15) implies that the coalition will not misreport  $\theta$  when the agent is efficient if:

$$w_{11} \geq w_{21} + \Delta \theta q_{21}, \quad (19)$$

and

$$w_{12} \geq w_{22} + \Delta \theta q_{22}. \quad (20)$$

Moreover, (16) also prevents a supervisor having observed signal  $\tau_1$  and learned that the agent has type  $\theta_2$  through side-contracting from reporting to the principal that the state of nature is  $(\theta_2, \tau_2)$ . The corresponding coalition incentive constraint becomes:

$$p_{21} e^{-r(w_{22} - u_{21})} + \lambda_1 r \Delta \theta q_{22} \geq p_{21} e^{-r(w_{21} - u_{21})} + \lambda_1 r \Delta \theta q_{21}. \quad (21)$$

From (17),  $\lambda_1$  is such that:

$$0 \leq \lambda_1 \leq p_{11}e^{-r(w_{11}-u_{11})}. \quad (22)$$

Finally, using again (16), we obtain also the reverse coalition incentive compatibility constraint. A supervisor having observed signal  $\tau_2$  and having learned that the agent has type  $\theta_2$  through side-contracting prefers to tell the truth rather than reporting that the state of nature is  $(\theta_2, \tau_1)$ . This constraint rewrites as:

$$p_{22}e^{-r(w_{21}-u_{22})} + \lambda_2 r \Delta\theta q_{21} \geq p_{22}e^{-r(w_{22}-u_{22})} + \lambda_2 r \Delta\theta q_{22}. \quad (23)$$

Using again (17),  $\lambda_2$  is such that:

$$0 \leq \lambda_2 \leq p_{12}e^{-r(w_{12}-u_{12})}. \quad (24)$$

Dividing (21) by  $p_{21}$  and (23) by  $p_{22}$  and summing yields

$$r\Delta\theta \left( \frac{\lambda_1}{p_{21}e^{ru_{21}}} - \frac{\lambda_2}{p_{22}e^{ru_{22}}} \right) (q_{22} - q_{21}) \geq 0. \quad (25)$$

Hence, the monotonicity constraint

$$q_{22} \geq q_{21} \quad (26)$$

must be satisfied by an implementable collusion-proof grand-mechanism<sup>27</sup> when  $r > 0$  and when the following inequality between multipliers holds:

$$\frac{\lambda_1}{p_{21}e^{ru_{21}}} > \frac{\lambda_2}{p_{22}e^{ru_{22}}}. \quad (27)$$

Feasible grand-mechanisms must also satisfy the supervisor's interim participation constraints: which write respectively as:

$$p_{11} + p_{21} \geq p_{11}e^{-r(w_{11}-u_{11})} + p_{21}e^{-r(w_{21}-u_{21})}. \quad (28)$$

and

$$p_{12} + p_{22} \geq p_{12}e^{-r(w_{12}-u_{12})} + p_{22}e^{-r(w_{22}-u_{22})}. \quad (29)$$

## 4.2 The Optimal Collusion-Proof Grand-Mechanism

We now proceed as follows. First, we take the values of  $\lambda_1$  and  $\lambda_2$  as given and we consider the whole class of collusion-proof grand-mechanisms satisfying coalition incentive compatibility constraints with those values of the multipliers. Assuming that (27)

---

<sup>27</sup>Note that this constraint is satisfied by the second-best allocation obtained without collusion.

holds, one can omit the coalition incentive constraint (23) and consider only (21) and the implementability condition (26). (26) is then checked ex post. Moreover, when (23) is slack, the principal's expected profit does not depend on  $\lambda_2$ . We can thus denote this profit as  $\Pi(\lambda_1)$ . Taking as given the results of the optimization within this class of  $\lambda_1$ -collusion-proof grand-mechanisms for a fixed  $\lambda_1$ , we finally optimize with respect to  $\lambda_1$  under the constraint that (24) remains satisfied. Finally, we check that  $\lambda_2$  can be chosen to satisfy (27) and (24). The first step of the optimization consists thus in finding the optimal  $\lambda_1$ -collusion-proof grand-mechanism under centralized contracting as a solution to the following problem

$$P(\lambda_1) : \quad \Pi(\lambda_1) = \max_{\{q_{ij}, w_{ij}, U_{ij}\}} \sum_{i,j} p_{ij} (R(q_{ij}) - \theta_i q_{ij} - w_{ij})$$

subject (1)-(2)-(18)-(19)-(20)-(21)-(28) and (29).

The second step of the optimization is to find the optimal value of  $\lambda_1$ . This is obtained by solving the following problem:

$$\begin{aligned} & \max_{\lambda_1} \Pi(\lambda_1) \\ & \text{subject to} \\ & 0 \leq \lambda_1 \leq p_{11} e^{-r(w_{11}(\lambda_1) - u_{11}(\lambda_1))} \end{aligned} \quad (30)$$

where  $w_{11}(\lambda_1)$  and  $u_{11}(\lambda_1)$  are solutions to  $(P(\lambda_1))$ . The next proposition summarizes the results of this optimization.

**Proposition 2 :** *The optimal collusion-proof grand-mechanism  $GC^c$  under centralized contracting entails:*

- *Constraints (1)-(2)-(18)-(20)-(21) and (29) are all binding. All other constraints are strictly satisfied.*
- *Denoting the difference of outputs by  $\Delta q^c = q_{22}^c - q_{21}^c$ , the agent's information rents are given by:*

$$u_{11}^c = \Delta \theta q_{21}^c, \quad (31)$$

$$u_{12}^c = \Delta \theta q_{22}^c, \quad (32)$$

$$u_{21}^c = u_{22}^c = 0. \quad (33)$$

*The supervisor's wages are given by:*

$$s_{11}^c = \Delta \theta \Delta q^c, \quad (34)$$

$$s_{12}^c = s_{22}^c = 0, \quad (35)$$

$$s_{21}^c = -\frac{1}{r} \ln \left( 1 + \frac{p_{11}}{p_{21}} r \Delta \theta \Delta q^c e^{-r \Delta \theta \Delta q^c} \right). \quad (36)$$

The supervisor's expected information rents when  $\tau_1$  and  $\tau_2$  are observed are respectively given by:

$$V_1^c = p(\theta_1 | \tau_1) v(s_{11}^c) + p(\theta_2 | \tau_1) v(s_{21}^c) = p(\theta_1 | \tau_1) \left( 1 - e^{-r \Delta \theta \Delta q^c} (1 + r \Delta \theta \Delta q^c) \right) > 0, \quad (37)$$

$$V_2^c = p(\theta_1 | \tau_2) v(s_{12}^c) + p(\theta_2 | \tau_2) v(s_{22}^c) = 0. \quad (38)$$

- The schedule of outputs is decreasing;  $q_{11}^c = q_{12}^c = q_1^{fb} > q_{22}^c > q_{21}^c$  where  $q_{2j}^c$  for  $j = 1, 2$  are implicitly defined by:

$$R'(q_{21}^c) = \theta_2 + \Delta \theta \frac{p_{11} e^{-r \Delta \theta \Delta q^c}}{p_{21} + p_{11} r \Delta \theta \Delta q^c e^{-r \Delta \theta \Delta q^c}}, \quad (39)$$

$$R'(q_{22}^c) = \theta_2 + \frac{\Delta \theta}{p_{22}} \left( p_{11} + p_{12} - \frac{p_{21} p_{11} e^{-r \Delta \theta \Delta q^c}}{p_{21} + p_{11} r \Delta \theta \Delta q^c e^{-r \Delta \theta \Delta q^c}} \right). \quad (40)$$

- In state  $\tau_1$ , the values of the multipliers are given by  $\lambda_1^c = p_{11} e^{-r \Delta \theta \Delta q^c}$  and  $\mu_1^c = 0$ . In state  $\tau_2$ ,  $\lambda_2^c = p_{12}$  and  $\mu_2^c = 0$  is one possible choice for those multipliers.

To understand this proposition, let us first come back to the conditionally optimal grand-mechanism  $GC^{sb}$  and show how it must be modified to deter the side-contract  $SC^{sb}$ . First, the wage  $s_{11}^c$  must be at least equal to the positive bribe  $-y_{11}^{sb} = \Delta \theta (q_{22}^{sb} - q_{21}^{sb})$  that the supervisor can get with a manipulation of reports. This raises the supervisor's expected utility in state  $\tau_1$ . Given that the supervisor accepts the grand-mechanism before learning the agent's type, the principal can recoup some of this extra cost by reducing the supervisor's wage  $s_{21}^c$ , i.e., the wage when he faces an inefficient agent. Of course, using such a wage lottery for the risk averse supervisor is costly. Moreover, reducing  $s_{21}^c$  may now also conflict with another incentive problem, namely ensuring that the coalition does not report  $(\theta_2, \tau_2)$  when the true state is  $(\theta_2, \tau_1)$  (as captured by (21)).<sup>28</sup>

If the supervisor is risk neutral, it is costless to design a grand-mechanism so that the supervisor has no incentive to collude when he observes  $\tau_1$ . Consider a grand-mechanism stipulating the second-best levels of outputs and transfers to the agent as before but with the following wages for the supervisor  $s_{11}^c = \Delta \theta (q_{22}^{sb} - q_{21}^{sb})$ ,  $s_{21}^c = -\frac{p_{11}}{p_{21}} \Delta \theta (q_{22}^{sb} - q_{21}^{sb})$ ,  $s_{12}^c =$

<sup>28</sup>Notice that the principal would still like to fully insure a supervisor who has observed  $\tau_2$  at the minimum level of wages, i.e.,  $s_{12}^c = s_{22}^c = 0$ .

$s_{22}^c = 0$ . The risk neutral supervisor's expected utility is still zero whether he observes  $\tau_1$  or  $\tau_2$ . Moreover,  $s_{11}^c$  is high enough to induce the truthful report of state  $(\theta_1, \tau_1)$ . The only thing left to check is that the coalition does not want to report  $(\theta_2, \tau_2)$  when  $(\theta_2, \tau_1)$  realizes. Such a manipulation increases the supervisor's wage in this state of the world by  $s_{22}^c - s_{21}^c = \frac{p_{11}}{p_{21}} \Delta\theta(q_{22}^{sb} - q_{21}^{sb})$ . However, this side-contract has to be offered before the agent has revealed his type to the supervisor and the supervisor takes thus into account that changing what he commits to announce in state  $(\theta_2, \tau_1)$  also affects the information rent paid to the agent in state  $(\theta_1, \tau_1)$ . Offering such a side-contract has a cost for the supervisor since the efficient agent's rent increases from  $\Delta\theta q_{21}^{sb}$  to  $\Delta\theta q_{22}^{sb}$  when  $\tau_1$  realizes. The expected cost of this manipulation borne by the supervisor is then  $p(\theta_1|\tau_1)\Delta\theta(q_{22}^{sb} - q_{21}^{sb})$  which just cancels out with its possible benefit. Hence, this grand-mechanism is coalition incentive compatible and implements the second-best outcome.

It is no longer costless to deter collusion when the supervisor is risk averse. Offering the previous wage lottery to the supervisor becomes costly since the supervisor must now receive a risk premium to accept the grand-mechanism. The principal faces a trade-off between providing insurance to the supervisor and inducing him to truthfully reveal his signal. The solution to this problem requires that, in comparison with the collusion-proof grand-mechanism offered under risk neutrality, the lottery faced by a supervisor knowing  $\tau_1$  is less risky. The collusion stake  $\Delta\theta(q_{22} - q_{21})$  and the difference between  $s_{22}$  and  $s_{21}$  are both reduced.

• **Collusion under Symmetric Information:** The fact that the supervisor is not fully informed on the agent's type is key for the results above. If collusion takes place under symmetric information, the trade-off faced by the supervisor between increasing his payoff in state  $(\theta_2, \tau_1)$  and reducing the information rent of the efficient agent no longer exists. The optimal collective manipulation of reports is always ex post efficient. This does not change the coalition incentive constraints (18), (19) and (20). However, (21) and (23) are written now with  $\lambda_1 = \lambda_2 = 0$  since there is no incentive constraint within the collusion. We derive from those latter two coalition incentive compatibility constraints that  $w_{21} = w_{22}$ . The principal can no longer screen with respect to  $\tau$ . He offers a pooling contract such that  $q_{21} = q_{22} = q_2^p$ . Moreover,  $w_{21} = w_{22} = 0$  and  $w_{11} = w_{12} = \Delta\theta q_2^p$ . Soft supervisory information is useless when collusion takes place under symmetric information.

• **Unique Implementation with Passive Beliefs:** Without uniqueness of the non-cooperative play of  $GC^c$ , there could be potentially other continuation equilibria where agents collude with the threat of playing another non-cooperative equilibrium of the grand-mechanism if they refuse to collude. In the Appendix, we show how to construct the out-of-equilibrium wages  $s_{ijk}^c$  for  $j \neq k$  to ensure unique Nash implementation when the supervisor holds passive beliefs following the agent's refusal of playing the null side-contract. The way these wages are constructed is similar to that when they do not collude.

Things are different when, in a continuation where collusion takes place, the agent's acceptance of a side-contract is sustained by the threat of playing non-cooperatively the grand-mechanism with *non-passive* beliefs. Indeed, the acceptance or refusal of the collusive side-contract can be viewed as a cheap talk stage which could be used to update the supervisor's beliefs on the agent's type. Let us still assume that the principal offers  $GC^c$  but that, when  $\tau_1$  realizes, the supervisor holds pessimistic beliefs following the agent's refusal of the null side-contract and assumes that the agent is efficient with probability  $p(\theta_1|\tau_2)$ . Remember that the principal has designed the out-of-equilibrium wages  $s_{ijk}^c$  to ensure unique implementation when a supervisor having observed  $\tau_2$  keeps passive beliefs following the agent's refusal of a side-contract. Hence, the unique non-cooperative equilibrium of  $GC^c$  when this mechanism is played with these pessimistic beliefs is such that both the supervisor and the agent report  $\tau_2$  to the principal. By refusing the side-contract, the efficient agent gets now a rent  $\Delta\theta q_{22}^c$  which is greater than what he gets in the continuation equilibrium sustained with passive beliefs, namely  $\Delta\theta q_{21}^c$ . With those pessimistic beliefs, the supervisor has an expected payoff  $p(\theta_1|\tau_2)v(s_{12}^c) + p(\theta_2|\tau_2)v(s_{22}^c) = 0$  which is lower than his expected payoff from playing the mechanisms with passive beliefs. Given these non-passive beliefs, the supervisor finds optimal to offer a non-truthful collusive side-contract  $SC^{nt}$  entailing the manipulations  $\phi_{\tau_j}^{nt}(\theta_1) = (\theta_1, \tau_2)$  and  $\phi_{\tau_j}^{nt}(\theta_2) = (\theta_2, \tau_j)$  for  $j = 1, 2$  and zero side-transfers. With this continuation, the supervisor gets an expected payoff (computed with prior beliefs on the equilibrium path since both types of agent accept the side-contract)  $p(\theta_1|\tau_1)v(w_{12}^c - \Delta\theta q_{22}^c) + p(\theta_2|\tau_1)v(w_{22}^c) = 0$ . The efficient agent gets instead  $\Delta\theta q_{22}^c$ . Note that this side-contract allocates all the gains from collusion to the agent and entails the same manipulation of reports as under complete information within the coalition. There are nevertheless two problems with this non-truthful continuation following the offer of  $GC^c$ . First, it yields a lower payoff to the principal who, in the spirit of standard mechanism design, could recommend to the agents the continua-

tion he prefers and select thereby the truthful continuation sustained with passive beliefs. Second, given that the supervisor also prefers this latter equilibrium and is endowed with all the bargaining power at the collusion stage, he should also recommend it.<sup>29</sup>

By contrast, as we see below, the multiplicity of continuation equilibria disappears if the principal offers a decentralized mechanism. Then, the agent's refusal from playing this mechanism has no impact on his status quo payoff which is identically equal to zero.<sup>30</sup>

## 5 The Equivalence Principle

The optimal collusion-proof grand-mechanism is clearly an upper bound on what the principal can achieve as any other mechanism can be reproduced by a centralized one. The point of this section is to show that the principal can also implement this outcome by simply delegating to the supervisor the right to contract with the agent.

Starting from the transfers and outputs of the optimal collusion-proof centralized grand-mechanism  $GC$ , let us consider the following direct grand-mechanism  $\tilde{GC}$ :

- Transfers:  $\tilde{t}_{ij} = 0, \tilde{s}_{ij} = s_{ij}^c + t_{ij}^c$  for all  $(i, j)$ .
- Outputs:  $\tilde{q}_{ij} = q_{ij}^c$  for all  $(i, j)$ .

$\tilde{GC}$  is thus obtained from  $GC^c$  by keeping the same outputs but giving to the supervisor all the transfers needed to implement those outputs.  $\tilde{GC}$  is a decentralized mechanism since the principal deals only with the supervisor who reports both his signal and the information he has learned on the agent first and, second, chooses the agent's transfer. The

---

<sup>29</sup>One may also wonder whether sensible restrictions on out-of-equilibrium beliefs eliminate the non-truthful continuation. Cho and Kreps (1987) intuitive criterion does not eliminate this equilibrium. The most suited such refinement for mechanism design problem was developed by Cramton and Palfrey (1995) who defined the notions of ratifiable and strongly ratifiable mechanisms. First, they define a credible veto belief as putting positive weights on those types who do not agree to play a collusive mechanism when the status quo mechanism is played with the corresponding beliefs. A side-contract is (resp. strongly) ratifiable when there does not exist a credible veto belief or if such a (resp. all) system (s) does not change the payoff of any type. One can show that both the truthful null side-contract sustained with passive beliefs and the non-truthful collusive side-contract defined above are strongly ratifiable.

<sup>30</sup>An alternative alley which is generally taken by the implementation literature to eliminate unwanted equilibria in non-cooperative environments would be to use extended messages in a centralized mechanism. We do not investigate this route since decentralization provides already a simple and more attractive solution to this multiplicity problem.



non-cooperative play of  $\tilde{G}\tilde{C}$  yields thus zero payoff to both the supervisor and the agent since the agent does not play the grand-mechanism if he refuses also the sub-contract.<sup>31</sup>

Given this grand-mechanism, the best sub-contract offered by  $S$  is obtained as a special case of Section 4.1.1. By definition, the agent receives now no rent in the non-cooperative play of  $\tilde{G}\tilde{C}$ ,  $U_{ij} = 0$ , and these null status quo payoffs cannot constrain ( $S$ ). Hence,  $\mu_j = 0$  for  $j = 1, 2$ , and, from (12), the multipliers  $\lambda_j$  are equal to the supervisor's marginal utility of income when facing an efficient agent times  $p_{1j}$ . These are precisely the same as under centralized contracting. Therefore, coalition incentive constraints take the same form. Since the aggregate payoffs of the coalition in all states of nature are unchanged, the optimal manipulation of reports remain truthful. Moreover, only the efficient agent's incentive constraints are binding at the sub-contracting stage and minimized by the supervisor in the same way as the principal would have done under centralized contracting. The agent's information rents from playing the composition of the optimal sub-contract and the grand-mechanism  $\tilde{G}\tilde{C}$  are thus respectively  $\tilde{u}_{1j} = \Delta\theta q_{2j}^c = \tilde{y}_{1j} - \theta_1 q_{1j}^c$  and  $\tilde{u}_{2j} = 0 = \tilde{y}_{2j} - \theta_2 q_{2j}^c$  where  $\tilde{y}_{1j}$  denote the transfers offered by the supervisor to the agent in the sub-contract. It is easy to check that these transfers are the same as with the centralized mechanism,  $\tilde{y}_{1j} = t_{1j}^c$ . This finally leads us to state our *Equivalence Principle*.

**Proposition 3 : The Equivalence Principle:** *The optimal collusion-proof centralized grand-mechanism can be implemented with a decentralized mechanism.*

Under centralization, the principal directly communicates with the agent and sets his wages. Recall that the principal is able to deter collusion only by benefitting from asymmetric information at the collusive stage. So the basic question behind the Equivalence Principle is whether the principal can use the agent's status quo payoffs to exacerbate the frictions coming from those informational asymmetries at the collusive stage. The answer to this question comes from the screening literature with type-dependent reservation utilities. Consider the supervisor-agent relationship. It is a standard principal-agent problem where the upward incentive constraint of the efficient agent implies the downward incentive constraint of the inefficient type. Suppose now that we raise the reservation utility of the efficient type. If this reservation utility is sufficiently large,<sup>32</sup> the incentive constraint is

---

<sup>31</sup>Moreover, note that  $s_{ijk}, t_{ijk}, q_{ijk}$  are not defined for  $j \neq k$ .

<sup>32</sup>But not too much in order to avoid countervailing incentives at the collusion stage. A collusion-proof

less binding at the collusion stage and the trade-off between rent extraction and efficiency in the coalition is tilted towards efficiency. Increasing the agent's status quo utility in the grand-mechanism is not going to help reducing the cost of coalition incentive constraints: the efficient agent's incentive constraint in the collusion is less likely to bind. So the principal would like to lower as much as he can the agent's utility levels in the centralized grand-mechanism. But this is exactly what is achieved by decentralizing contracting since the supervisor is willing to put the agent at the minimum level of rents compatible with the latter's participation and incentive constraints. The centralized mechanism cannot do better than the decentralized one.

The optimal decentralized grand-mechanism can thus be viewed as a particular implementation of the optimal collusion-proof grand-mechanism under centralized contracting. This implementation ensures also uniqueness of the continuation equilibrium. In a decentralized organization the principal shuts down the communication channel with the agent and kills the possibility of various continuation equilibria. In particular, even if the agent's refusal of the sub-contract changes the supervisor's beliefs, it has no impact on his payoff following this refusal since this payoff remains identically equal at zero. This gives an argument for the strict dominance of a decentralized mechanism where collusion, by design, takes place on the equilibrium path.<sup>33</sup>

## 6 Comparative Statics

With the insight of our Equivalence principle, we can now examine the impact on the optimal contract of the different parameters characterizing the economic environment. We now make the dependence of output on  $r$  explicit.

### 6.1 The Role of Risk Aversion

**Proposition 4** :  $q_{21}^c(r)$  (resp.  $q_{22}^c(r)$ ) is an increasing (resp. decreasing) function of the supervisor's degree of risk aversion  $r$ . Moreover, for all  $r > 0$ ,  $q_{21}^c(r) < q_{22}^c(r)$  and:

mechanism such that those constraints are binding is clearly dominated and, hence, we have restricted the analysis to the case where the agent's upward incentive constraint is binding.

<sup>33</sup>The reader may have noticed that the supervisor is indifferent between claiming truthfully  $(\theta_2, \tau_1)$  or  $(\theta_2, \tau_2)$  even in the decentralized organization. This indifference is, as usual, broken in favor of the principal and this can be done at  $\epsilon$  cost by perturbing transfers. Hence, this non-truthful continuation is irrelevant.

$$\lim_{r \rightarrow \infty} q_{21}^c(r) = \lim_{r \rightarrow \infty} q_{22}^c(r) = q_2^p,$$

$$\lim_{r \rightarrow 0} q_{2j}^c(r) = q_{2j}^{sb}.$$

We have already discussed the reason why, with a risk neutral supervisor, collusion can be deterred at no cost. Equivalently, delegating the sub-contract to the risk neutral supervisor is costless for the principal in the decentralized structure. The analogy with a standard moral hazard problem helps to understand Proposition 4. The principal wants to give the right incentives to the supervisor regarding the choice of a sub-contract. Extending the logic of the standard moral hazard literature, this incentive problem can be solved at no cost with risk neutrality. There exists a system of state-dependent rewards and punishments which induces a costless revelation of supervisory information. The contractual outcome is conditionally optimal.

With a risk averse supervisor, the principal must provide costly insurance and outputs are distorted to limit the risk borne by the supervisor. This risk being proportional to  $\Delta q^c$ , the more risk averse the supervisor, the greater the insurance concern and the larger the output distortions needed to fight collusion with respect to the conditionally optimal outcome.  $q_{22}^c(r)$  and  $q_{21}^c(r)$  are more and more distorted away from their values obtained with risk neutrality as  $r$  increases. In the limit, these outputs converge one towards the other and supervisory information becomes useless for the principal and bunching along  $\tau$  becomes optimal. The optimal contract becomes somewhat “incomplete” as it is now almost independent of  $\tau$ . Except in this limiting case, the information collected by the supervisor has always a positive value.

## 6.2 Precision of Supervisory Information

Taking risk aversion as given, we inspect now the effect of the accuracy of the supervisor’s information on the optimal contract. The informational structure defined in Footnote 13 (with  $\epsilon = Prob(\tau_i|\theta_i) \geq 1/2$ ) is useful for examining this question.

**Proposition 5 :** *When  $\epsilon$  converges towards  $\frac{1}{2}$  (uninformative signal) and 1 (perfectly informative signal), the inefficient agent’s expected output converges to the unconditional*

second-best  $q_2^p$ . The principal's expected welfare  $W(\epsilon)$  is non-monotonic in  $\epsilon$  with  $W\left(\frac{1}{2}\right) = W(1)$ ,  $W'\left(\frac{1}{2}\right) = 0$  and  $W'(1) = -\infty$ , i.e., it has an interior maximum for  $\epsilon^{sb} \in \left(\frac{1}{2}, 1\right)$ .

The Equivalence Principle is here also useful to understand the impact of the accuracy of supervisory information. Consider the decentralized structure, the reason why the second-best is produced when  $\epsilon = \frac{1}{2}$  is trivial: the signal is uninformative, so de facto the three-tier hierarchy boils down to a standard principal-agent pair without supervision. But similarly the three-tier hierarchy reduces to a standard principal-agent pair when the supervisory information is almost perfect. If  $\epsilon = 1$ , there is indeed no asymmetry of information between the supervisor and the agent. Then, everything happens as if the supervisor is endowed with the production technology himself: he can produce any quantity  $q$  by simply paying  $t = \theta_i q$  to the agent. In the decentralized setting, the principal ends up facing a merger between the supervisor and the agent.

The most preferred supervisory technology from the principal's point of view is an interior one which trades off the direct benefit of a more precise supervisory signal against the increase in the agency cost of decentralized contracting (or equivalently in the cost of deterring collusion). A more accurate supervisory information makes easier the control of the agent. However, it also makes easier collusion. When  $\epsilon$  increases, this latter effect dominates. A more precise information may now hurt the principal. Therefore, the optimal information structure trades-off a better control of the agent by the supervisor against a worse control of the supervisor by the principal himself.<sup>34</sup>

### 6.3 Bargaining Power and Side-Contracting

Let us now consider the case where the agent has some bargaining power at the time of designing the collusive agreement. To model such a setting, we assume that there exists a third party which offers the side-contract to maximize a weighted sum of the agent and the supervisor's utility with a weight  $\alpha$  on the agent's utility.<sup>35</sup>  $\tau_j$  is known to the third party but the agent still has to be provided with the right incentive to report his type. The

---

<sup>34</sup>This result on the optimality of an intermediate accuracy for supervisory information belongs to an emerging literature endogenizing information structures in principal-agent models. See Crémer (1995), Dewatripont and Maskin (1995), Schmidt (1996) and Faure-Grimaud (1998). A similar result is obtained in Laffont and Meleu (1997) for the case of hard information and exogenous transaction costs of collusion.

<sup>35</sup>See Laffont and Martimort (1997, 2000).

optimal side-contract is now solution to the following problem:

$$\begin{aligned} & \max_{\{\phi_{\tau_j}(\cdot), u_{ij}\}} p(\theta_1|\tau_j) \left( v(s(\phi_{\tau_j}(\theta_1)) + t(\phi_{\tau_j}(\theta_1)) - \theta_1 q(\phi_{\tau_j}(\theta_1)) - u_{1j}) + \alpha u_{1j} \right) \\ & + p(\theta_2|\tau_j) \left( v(s(\phi_{\tau_j}(\theta_2)) + t(\phi_{\tau_j}(\theta_2)) - \theta_2 q(\phi_{\tau_j}(\theta_2)) - u_{2j}) + \alpha u_{2j} \right) \\ & \text{subject to (1)-(10)-(11)} \end{aligned}$$

and the supervisor's participation constraints at the side-contracting stage:

$$\begin{aligned} & p(\theta_1|\tau_j) v(s(\phi_{\tau_j}(\theta_1)) + t(\phi_{\tau_j}(\theta_1)) - \theta_1 q(\phi_{\tau_j}(\theta_1)) - u_{1j}) \tag{41} \\ & + p(\theta_2|\tau_j) v(s(\phi_{\tau_j}(\theta_2)) + t(\phi_{\tau_j}(\theta_2)) - \theta_2 q(\phi_{\tau_j}^d(\theta_2)) - u_{2j}) \geq V_j \end{aligned}$$

where  $V_j$  denotes the supervisor's status quo payoff from playing non-cooperatively the grand-mechanism.

The solution to this problem can be derived as before. Denoting now the respective positive multipliers of (1), (2), (41) by  $\lambda_j(\alpha)$ ,  $\mu_j(\alpha)$  and  $\nu_j(\alpha)$ , we find the following characterization of the optimum:

- Optimizing with respect to  $u_{1j}$  yields:

$$p_{1j} \left( (1 + \nu_j(\alpha)) v'(s(\phi_{\tau_j}^*(\theta_1)) + t(\phi_{\tau_j}^*(\theta_1)) - \theta_1 q(\phi_{\tau_j}^*(\theta_1)) - u_{1j}^* - \alpha) \right) = \lambda_j(\alpha) + \mu_j(\alpha). \tag{42}$$

- Optimizing with respect to  $\phi_{\tau_j}^*(\theta_1)$  and  $\phi_{\tau_j}^*(\theta_2)$ , we obtain respectively:

$$\phi_{\tau_j}^*(\theta_1) \in \arg \max_{\tilde{\phi} \in \Delta(M_a \times M_s)} s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_1 q(\tilde{\phi}), \tag{43}$$

$$\phi_{\tau_j}^*(\theta_2) \in \arg \max_{\tilde{\phi} \in \Delta(M_a \times M_s)} p_{2j} (1 + \nu_j(\alpha)) v(s(\tilde{\phi}) + t(\tilde{\phi}) - \theta_2 q(\tilde{\phi}) - u_{2j}^*) - \lambda_j(\alpha) \Delta \theta q(\tilde{\phi}). \tag{44}$$

- Finally, the slackness conditions tell us that  $\lambda_j(\alpha) = 0$  (resp.  $\mu_j(\alpha) = 0$  and  $\nu_j(\alpha) = 0$ ) when (1) (resp. when (10) and when (41)) is slack.

The characterization of the collusion-proof mechanisms still obey (12) and (14). Those constraints remain in fact the same as before except for the fact that  $\tilde{\lambda}_j(\alpha) = \frac{\lambda_j(\alpha)}{1 + \nu_j(\alpha)}$  and  $\tilde{\mu}_j(\alpha) = \frac{\mu_j(\alpha) + \alpha p_{1j}}{1 + \nu_j(\alpha)}$  are now replacing respectively  $\lambda_j$  and  $\mu_j$ .

For fixed values of the multipliers  $\lambda_j(\alpha)$  and  $\mu_j(\alpha)$ , we can, as in Section 4, define various classes of collusion-proof grand-mechanisms. As before, the principal can also play on the values of these multipliers to improve contracting. By setting  $\nu_j(\alpha)$  very

large but keeping  $\frac{\lambda_i(\alpha)}{1+\nu_j(\alpha)}$  constant and equal to  $\lambda_j$ , the principal can come as close as he wants to the outcome obtained with the supervisor having all the bargaining power in the collusion. The principal can thus replicate the same outcome as if the agent has no bargaining power.

To get further insight on this issue, let us consider now the polar case where the agent has all the bargaining power in designing the collusive agreement under centralized contracting. To keep some symmetry with the previous analysis, we assume that the collusive offer is made by the agent before he learns  $\theta$  but after  $\tau$  is learned by both the agent and the supervisor. We already know that the optimal collusion-proof mechanism  $GC^c$  found in Section 4 may not be robust to this kind of collusion if the principal cannot recommend which out-of-equilibrium beliefs are held by the supervisor following the agent's refusal from colluding. Indeed, the side-contract  $SC^{mt}$  sustained with pessimistic beliefs gives all the bargaining power to the agent and clearly maximizes his expected payoff. The principal can nevertheless avoid this outcome by recommending to the supervisor that he holds passive beliefs following the agent's refusal of a side-contract which amounts to ensuring that the supervisor gets  $V_1^c$  in the non-cooperative truthful play of  $GC^c$  when he has observed  $\tau_1$ . Even if the agent has all the bargaining power, no other side-contract than the null one can improve his own payoff and be accepted by the supervisor since  $V_1^c$  maximizes the latter's payoff when he has all the bargaining power.

## 7 Conclusion

Our first main result is that soft supervisory information helps the principal even though the supervisor and the supervised agent collude if this collusion takes place under asymmetric information. A second insight of this paper is that the optimal collusion-proof contract can easily be implemented by delegating to the less informed agent in the organization, namely the supervisor, the task to contract with the more informed productive agent.

Beyond these results, it is worth stressing some conditions under which either organizational form strictly dominates. As already noticed in the text, decentralization may be strictly preferred if one insists on unique implementation. The idea, new to our knowledge, that decentralization is a way to get rid of multiple equilibria when fighting collusion

could extend to other settings of interest. Other elements could be considered to pursue the comparison between the two structures. First, limiting the enforceability of collusive contracts increases the principal's profit in the centralized organization and favors this organizational form. The principal is then better off integrating the productive activities rather than decentralizing when frictions on the enforceability of the side-contract exist. In particular, the supervisor may not be always corruptible. If the probability of honesty is large enough in a centralized organization, the principal allows for some (almost costless) equilibrium collusion between a dishonest supervisor and the agent. In this case, a decentralized organization is dominated. Reciprocally, if informational problems within the coalition are less acute under centralized contracting, collusion may then be more costly than decentralized contracting. Finally, one reading of our result is that delegating contracting to the less informed agent within the coalition helps the principal. It would be interesting to investigate how this result extends or changes when the agent is also asymmetrically informed on the supervisor's signal.

Under decentralized contracting, we have assumed so far that the supervisor keeps all bargaining power at the sub-contracting stage. A more even distribution of this bargaining would make decentralization costly for the principal since he can no longer restore his most preferred allocation by shifting up the status quo payoff of the supervisor. This would invalidate the Equivalence Principle and restore the benefits of centralized contracting. More generally, this last point clearly raises the issue of the exact relationship between the allocation of formal authority in the hierarchy and the kind of social relationships which establish among its different layers. This remains a fascinating issue which should deserve more work.

## References

- Baliga, S.**, 1999, "Collusion with Soft Information", *Journal of Law, Economics and Organization*, 15: 434-440.
- Baliga, S. and T. Sjoström**, 1998, "Decentralization and Collusion," *Journal of Economic Theory*, 83: 196-232.
- Baron, D. and D. Besanko**, 1992, "Information, Control and Organizational Structure," *Journal of Economics and Management Strategy*, 1: 237-275.

- Cho, I-K, and D. Kreps.**, 1987, "Signalling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102: 179-221.
- Cramton, P. and T. Palfrey**, 1995, "Ratifiable Mechanisms: Learning from Disagreement," *Games and Economic Behavior*, 10: 255-283.
- Crémer, J.**, 1995, "Arm's Length Relationships," *Quarterly Journal of Economics*, 110: 275-295.
- Crémer, J. and M. Riordan**, 1987, "On Governing Multilateral Transactions with Bilateral Contracts," *Rand Journal of Economics*, 18: 436-451.
- Dewatripont, M. and E. Maskin**, 1995, "Contractual Contingencies and Renegotiation," *Rand Journal of Economics*, 26: 704-719.
- Faure-Grimaud, A.**, 1998, "Using Stock Price Information to Regulate Firms," Mimeo London School of Economics, London.
- Felli, L.**, 1997, "Preventing Collusion through Delegation," Mimeo STICERD, London.
- Holmstrom, B.**, 1979, "Moral Hazard and Observability," *Bell Journal of Economics*, 10: 74-91.
- Holmstrom, B. and P. Milgrom**, 1990, "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics*, 146: 85-105.
- Itoh, H.**, 1993, "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory*, 60: 410-427.
- Jullien, B.**, 2000, "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93: 1-47.
- Kofman, F. and J. Lawarrée**, 1993, "Collusion in Hierarchical Agency," *Econometrica*, 61: 629-656.
- Laffont, J.J. and D. Martimort**, 1997, "Collusion under Asymmetric Information," *Econometrica*, 65: 875-911.
- Laffont, J.J. and D. Martimort**, 1998, "Collusion and Delegation," *Rand Journal of Economics*, 29: 280-305.



- Laffont, J.J. and D. Martimort**, 2000, "Mechanism Design under Collusion and Correlation," *Econometrica*, 68: 309-342.
- Laffont, J.J. and M. Meleu**, 1997, "Reciprocal Supervision, Collusion and Organizational Design," *Scandinavian Journal of Economics*, 99: 519-540.
- Laffont, J.J. and J.C Rochet**, 1999, "Regulation of a Risk-Averse Firm," in C. D'Aspremont ed. *Social Organizations and Mechanism Design*, De Boeck Université.
- Laffont, J.J. and J. Tirole**, 1993, *A Theory of Incentives in Procurement and Regulation*, MIT Press, Cambridge.
- Macho-Stadler, I. and J. Perez-Castrillo**, 1994, "Centralized and Decentralized Contracts in a Moral Hazard Environment," mimeo Universitat Autònoma Barcelona.
- Martimort, D.**, 1999, "The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs," *Review of Economic Studies*, 229: 929-948.
- Maskin, E.**, 1999, "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 226: 23-38.
- McAfee, P. and J. McMillan**, 1995, "Organizational Diseconomies of Scope," *Journal of Economics and Management Strategy* 4: 399-426.
- Melumad, N., D. Mookherjee, and S. Reichelstein**, 1992, "A Theory of Responsibility Centers," *Journal of Accounting Economics*, 15: 445-484.
- Melumad, N., D. Mookherjee, and S. Reichelstein**, 1995, "Hierarchical Decentralization of Incentive Contracts," *Rand Journal of Economics*, 26: 654-692.
- Mirrlees, J.**, 1999, "The Theory of Moral Hazard and Unobservable Behavior: Part 1," *The Review of Economic Studies*, 66: 3-23.
- Mookherjee, D. and S. Reichelstein**, 2001, "Incentives and Coordination in Hierarchies," *Bepress Quality Rated Journal of Theoretical Economics*.
- Salanié, B.**, 1990, "Selection Adverse et Aversion pour le Risque," *Annales d'Economie et de Statistiques*, 18: 131-150.
- Schmidt, K.**, 1996, "The Costs and Benefits of Privatization: an Incomplete Contract Approach," *Journal of Law, Economics and Organization*, : 1-24.

**Tirole, J.**, 1986, "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economic and Organization*, 2: 181-214.

**Tirole, J.**, 1992, "Collusion and the Theory of Organizations," in *Advances in Economic Theory*, vol. 2, ed. J.J. Laffont, 151-206, Cambridge University Press, Cambridge.

## Appendix

**Proof of the Collusion-Proofness-Principle:** We denote by  $SC^*$  the pair of optimal side-contracts  $SC_{\tau_j}^*$  proposed by the supervisor to the agent according to his supervisory signal  $\tau_j$ . These side-contracts are accepted by both types along an equilibrium path where non-trivial side-contracts are enforced. These side-contracts  $SC_{\tau_j}^*$  are incentive compatible for both types (in particular, they satisfy (1)), stipulate a manipulation of reports  $\phi^*(\cdot)$  and side-transfers  $y_{ij}^*$  according to the supervisor's signal such that (10) and (11) are satisfied for both signals  $\tau_1$  and  $\tau_2$ . Let us denote by  $u_{ij}^*$  the information rents of a  $\theta_i$  agent when accepting the optimal side-contracts  $SC_{\tau_j}^*$ , reporting truthfully to the supervisor and obeying the latter's recommendations of reports to the principal. The reservation utility of the agent is then  $U_{ij}$ , i.e., his payoff in one non-cooperative equilibrium of the grand-mechanism. Then, starting from any grand-mechanism  $GC$  (possibly an indirect one) let us construct a new grand-mechanism  $\tilde{GC} = GC \circ SC^*$  with message spaces  $\tilde{M}_s = T$  and  $\tilde{M}_a = \Theta \times T$ . This direct grand-mechanism is such that, once it is offered by the principal, the optimal side-contracts  $\tilde{SC}^*$  which are offered by the supervisor at the side-contracting stage are the null-side-contracts entailing no further manipulation of reports,  $\tilde{\phi}^*(\cdot) = Id$  and no side-transfers,  $\tilde{y}_{ij}^* = 0 \quad \forall (i, j)$ , i.e., the grand-mechanism is also truthful. Suppose it is not the case, then, there would exist a non-null pair of side-contracts  $\tilde{SC}_{\tau_j}^*$  for  $j = 1, 2$  which would satisfy (1) and would leave both types with enough information rent so that they prefer these new side-contracts  $\tilde{SC}_{\tau_j}^*$ , i.e,  $\tilde{u}_{ij}^* \geq u_{ij}^* \geq U_{ij}$ . The agent's acceptance of playing the new side-contract  $\tilde{SC}_{\tau_j}^*$  is sustained by the threat of playing side-contract  $SC_{\tau_j}^*$  with passive beliefs so that the corresponding play of the game gives him  $u_{ij}^*$ . Moreover, at least one of these side-contracts  $\tilde{SC}_{\tau_j}^*$  for  $j = 1, 2$  would yield a strictly higher profit to the supervisor than the null side-contract. Then, the supervisor would have been strictly better off offering the side-contract  $SC^{c*} \circ \tilde{SC}^*$  in the first place when  $GC$  was offered by the principal. Hence, a contradiction with the definition of  $SC^*$ .

**Proof of Proposition 1:** First, since (13) holds for all  $\tilde{\phi}(\cdot) \in \Delta(M_s \times M_a)$  it holds also in particular for  $\phi_{\tau, j'}^*(\theta_{i'})$  for all  $j' \neq j$  and  $i' \neq i$ .  $GC \circ SC^*$  is thus a collusion-proof grand-mechanism satisfying (15). One obtains similarly (16) from (14). Second,  $\mu_j^c \geq 0$  and (12) imply that  $\lambda_j^c$  satisfies (17). In a collusion-proof grand-mechanism, it must be that  $u_{1j}^* = U_{1j}$  and this equality has been used to express the right-hand side in (17). Lastly, (1) being satisfied by the information rents obtained by the agent from side-contracting and those rents being those committed to by the principal offering a collusion-proof grand-mechanism, (1) must also be part of the description of these collusion-proof grand-mechanisms.

**Proof of Propositions 2 and 4:**

- It is useful to make the following change of variables  $z_{ij} = e^{-rw_{ij}}$ . For a fixed schedule of outputs  $q_{ij}$ , the principal's objective function becomes then strictly concave in  $z_{ij}$  (for  $r > 0$ ) and constraints (18)-(19)-(20)-(21)-(23)-(28)-(29) form now a system of linear constraints in  $z_{ij}$ . We will first assume that the only relevant constraints are (18)-(20)-(21)-(29). We will check ex post that the other constraints are satisfied at the optimum as well as the monotonicity conditions  $q_{1j} \geq q_{2j}$  for  $j = 1, 2$  and  $q_{22} \geq q_{21}$ . Rewriting constraints (18)-(19)-(20) and (21) with our new variables yields respectively:

(18) becomes

$$z_{11} = z_{12}, \tag{45}$$

(19) becomes

$$z_{11} \leq z_{21} e^{-r\Delta\theta q_{21}}, \tag{46}$$

(20) when becomes

$$z_{12} \leq z_{22} e^{-r\Delta\theta q_{22}}, \tag{47}$$

(21) becomes

$$p_{21}(z_{22} - z_{21}) \geq -\lambda_1 r \Delta\theta (q_{22} - q_{21}) e^{-ru_{21}}. \tag{48}$$

With the new variables, (28) and (29) rewrite respectively as:

$$p_{11} + p_{21} \geq p_{11} e^{ru_{11}} z_{11} + p_{12} e^{ru_{21}} z_{21}, \tag{49}$$

$$p_{11} + p_{22} \geq p_{12}e^{ru_{12}}z_{12} + p_{22}e^{ru_{22}}z_{22}. \quad (50)$$

Moreover, the efficient agent's incentive compatibility constraint rewrite respectively as:

$$u_{11} \geq u_{21} + \Delta\theta q_{21}, \quad (51)$$

$$u_{12} \geq u_{22} + \Delta\theta q_{22}. \quad (52)$$

Making the agent's incentive constraints (51) and (52) both binding relaxes the supervisor's participation constraints (49) and (50). Hence, (49) and (50) rewrite respectively as:

$$(p_{11}z_{11}e^{r\Delta\theta q_{21}} + p_{21}z_{21})e^{ru_{21}} \leq p_{11} + p_{21}. \quad (53)$$

$$(p_{12}z_{11}e^{r\Delta\theta q_{22}} + p_{22}z_{22})e^{ru_{22}} \leq p_{12} + p_{22}. \quad (54)$$

• For a given schedule of outputs  $q_{ij}$ , the principal wants first to minimize the cost of implementing this schedule:

$$\max_{\{z_{ij}, u_{2j}\}} \sum_{i,j} \frac{p_{ij}}{r} \ln(z_{ij})$$

subject to (2)-(45)-(46)-(47)-(48)-(53) and (54).

First, we fix all the  $z_{ij}$  and optimize with respect to the remaining variables, i.e., (once one has noticed that (1) is binding for all  $j$ )  $u_{2j}$  only.

Since  $\lambda_1 \geq 0$  and  $q_{22} \geq q_{21}$ , (48) is relaxed when  $e^{-ru_{21}}$  is as large as possible. Similarly, (53) is also relaxed when  $e^{-ru_{21}}$  is as large as possible as long as (2) for  $j = 1$  remains satisfied. The constrained set is finally larger when  $u_{21} = 0$ . Similarly, (54) is relaxed when  $u_{22}$  is as small as possible as long as (2) for  $j = 2$  remains satisfied. The principal's objective function is thus also maximized when  $u_{22} = 0$ .

Inserting these values of  $u_{21}$  and  $u_{22}$  into the constrained of the principal's problem, these constraints become linear in  $z_{ij}$  and we can continue the optimization.

• The maximization of the principal's problem yields that (45), (47), (48) and (54) are all binding. We check ex post that other constraints are slack when those four constraints are binding. We get immediately:

$$z_{22}^c = 1, \quad (55)$$

$$z_{12}^c = z_{11}^c = e^{-r\Delta\theta q_{22}} \quad (56)$$

$$z_{21}^c = 1 + \frac{\lambda_1 r}{p_{21}} \Delta\theta(q_{22} - q_{21}). \quad (57)$$

We thus obtain  $w_{11}^c = w_{12}^c = \Delta\theta q_{22}$ ,  $w_{22}^c = 0$  and  $w_{21}^c = -\frac{1}{r} \ln \left( 1 + \frac{\lambda_1 r}{p_{21}} \Delta\theta(q_{22} - q_{21}) \right) < 0$ .

• Inserting the corresponding values of  $z_{ij}^c$  as functions of outputs into the principal's objective function and optimizing with respect to  $q_{ij}$  yields no distortion for the  $\theta_1$  type with respect to the first-best and distortions for the  $\theta_2$  type which are given by:

$$R'(q_{21}^c(\lambda_1)) = \theta_2 + \Delta\theta \frac{\lambda_1}{p_{21} + \lambda_1 r \Delta\theta(q_{22}^c(\lambda_1) - q_{21}^c(\lambda_1))}, \quad (58)$$

$$R'(q_{22}^c(\lambda_1)) = \theta_2 + \frac{\Delta\theta}{p_{22}} \left( p_{11} + p_{12} - \frac{p_{21} \lambda_1}{p_{21} + \lambda_1 r \Delta\theta(q_{22}^c(\lambda_1) - q_{21}^c(\lambda_1))} \right) \quad (59)$$

where we make explicit the dependence of those outputs on  $\lambda_1$ . We observe that the implementability condition  $q_{22}(\lambda_1) \geq q_{21}(\lambda_1)$  is satisfied by the solutions to (58) and (59) for  $\lambda_1 \geq \lambda_m = \frac{p_{21}(p_{11} + p_{21})}{p_{21} + p_{22}}$ . For  $\lambda_1 = \lambda_m$ , the two outputs are equal and when  $\lambda_1 < \lambda_m$ , the solutions to (58) and (59) do not satisfy the monotonicity condition. Hence, there is some bunching yielding  $q_{22}(\lambda_1) = q_{21}(\lambda_1) = q_2^p$ . Note that the principal's profit does not depend on  $\lambda_1$  for  $\lambda_1 \leq \lambda_m$ .

• Let us now optimize  $\Pi(\lambda_1)$  with respect to  $\lambda_1$ . The constraint on  $\lambda_1$  rewrites as:

$$0 \leq \lambda_1 \leq p_{11} e^{-r\Delta\theta\Delta q^c(\lambda_1)}. \quad (60)$$

First note that  $\Pi(\lambda_1)$  is an increasing function of  $\lambda_1$  since a direct use of the envelope theorem yields  $\Pi'(\lambda_1) = \frac{\Delta\theta\Delta q^c(\lambda_1)}{p_{21} + \lambda_1 \Delta\theta\Delta q^c(\lambda_1)} \geq 0$  with a strict inequality for  $\lambda_1 > \lambda_m$ . Hence, increasing  $\lambda_1$  improves the principal's payoff. Moreover, the function  $\phi(\lambda_1) = p_{11} e^{-r\Delta\theta\Delta q^c(\lambda_1)} - \lambda_1$  is positive for  $\lambda_1 = \lambda_m$  since  $\frac{p_{11}}{p_{21}} > \frac{p_{11} + p_{21}}{p_{21} + p_{22}}$  holds when  $\tau$  is positively correlated with  $\theta$ . Moreover,  $\phi(\lambda_1)$  is negative for  $\lambda_1$  large enough. Hence, there exists a maximal solution to the equation  $\phi(\lambda_1) = 0$  and it maximizes  $\Pi(\lambda_1)$ . (60) is thus binding at the optimum. Inserting the corresponding value of  $\lambda_1$  obtained from (60) into (58) and (59) yields output distortions given by (39) and (40) for the  $\theta_2$  type. We also obtain immediately the aggregate payoff's of the coalition as:

$$w_{11}^c = w_{12}^c = \Delta\theta q_{22}^c, \quad (61)$$

$$w_{22}^c = 0, \quad (62)$$

$$w_{21}^c = -\frac{1}{r} \ln \left( 1 + \frac{p_{11}}{p_{21}} r \Delta\theta \Delta q^c e^{-r\Delta\theta\Delta q^c} \right). \quad (63)$$

- We prove now that  $q_{22}^c(r) > q_{21}^c(r)$  for all  $r$  and that both outputs converge to  $q_2^p$  as  $r$  goes to  $\infty$ . Consider the solutions to equations (39) and (40). For  $r = 0$ ,  $q_{22}^c(0) = q_{22}^{sb}$  and  $q_{21}^c(0) = q_{21}^{sb}$ . Moreover, differentiating (39) and (40) w.r.t.  $r$  yields respectively:

$$R''(q_{21}^c(r)) \frac{dq_{21}^c}{dr} = -\frac{p_{11}r\Delta\theta^2(p_{21}e^{-r\Delta\theta\Delta q^c(r)} + p_{11}e^{-2r\Delta\theta\Delta q^c(r)})}{(p_{21} + p_{11}r\Delta\theta\Delta q^c(r)e^{-r\Delta\theta\Delta q^c(r)})^2} \left(1 + r \left(\frac{dq_{22}^c}{dr} - \frac{dq_{21}^c}{dr}\right)\right), \quad (64)$$

and

$$R''(q_{22}^c(r)) \frac{dq_{22}^c}{dr} = -\frac{p_{21}}{p_{22}} R''(q_{21}^c(r)) \frac{dq_{21}^c}{dr}. \quad (65)$$

Hence  $\frac{dq_{21}^c}{dr}$  and  $\frac{dq_{22}^c}{dr}$  have opposite signs and are never equal to zero. Moreover, if  $\frac{dq_{21}^c}{dr} < 0 < \frac{dq_{22}^c}{dr}$  we would have a contradiction with (64). Hence,  $\frac{dq_{22}^c}{dr} < 0$  and  $\frac{dq_{21}^c}{dr} > 0$ .

Suppose that there exists  $r_1$  such that  $\infty > r_1 > 0$  and such that  $q_{22}^c(r_1) = q_{21}^c(r_1)$ . Inserting into (39) and (40), we obtain  $q_{22}^c(r_1) = q_{22}^{sb}$  and  $q_{21}^c(r_1) = q_{21}^{sb} < q_{22}^{sb}$ . A contradiction. Hence,  $q_{22}^c(r) > q_{21}^c(r)$  for all  $r$  and there is never any bunching along  $\tau$ .

- Since  $q_{21}^c(r) < q_{22}^c(r) < q_{22}^{sb} < q_{11}^c(r) = q_{12}^c(r) = q_1^{fb}$ , other monotonicity conditions on outputs are satisfied.
- As  $q_{22}^c(r) \geq q_{21}^c(r)$ ,  $q_{22}^c(r)$  is bounded below by  $q_{21}^{sb}$  and  $q_{21}^c(r)$  is bounded above by  $q_{22}^{sb}$ . As  $q_{22}^c(r)$  (resp.  $q_{21}^c(r)$ ) is decreasing (resp. increasing) it converges towards a limit  $q_{22}^c(\infty)$  (resp.  $q_{21}^c(\infty)$ ) as  $r$  goes to infinity. Moreover,  $q_{22}^c(\infty) \geq q_{21}^c(\infty)$ . Suppose that  $q_{22}^c(\infty) > q_{21}^c(\infty)$ . Then, taking the limit in (39) and (40) we obtain

$$R'(q_{21}^c(\infty)) = \theta_2 < R'(q_{22}^c(\infty)) = \theta_2 + \left(\frac{p_{11} + p_{12}}{p_{22}}\right) \Delta\theta.$$

A contradiction since  $R''(\cdot) \leq 0$ . Hence, necessarily  $q_{21}^c(\infty) = q_{22}^c(\infty) = q_2(\infty)$ . But, using (39) and (40), we obtain:

$$p_{21}(R'(q_{21}^c(r)) - \theta_2) + p_{22}(R'(q_{22}^c(r)) - \theta_2) = \Delta\theta(p_{11} + p_{12}).$$

Making  $r = \infty$  yields  $q_2(\infty) = q_2^p$ .

- We check all other neglected incentive and participation constraints.
  - Note that the monotonicity constraints  $q_{1j}^c > q_{2j}^c$  are satisfied and they imply that the  $\theta_2$ 's incentive constraint is strictly satisfied for all  $j$ .
  - (1) is binding for all  $j$  and thus, since  $u_{2j} = 0$ ,  $u_{1j} > 0$  and the  $\theta_1$ 's agent's participation constraints are satisfied for all  $j$ .

- Inserting the value of  $z_{11}$  obtained in (56), the supervisor's participation constraint (53) is strictly satisfied when:

$$p_{11}(1 - e^{-r\Delta\theta\Delta q^c(r)}) > \lambda_1^c r \Delta\theta\Delta q^c(r). \quad (66)$$

Using that  $\lambda_1^c = p_{11}e^{-r\Delta\theta\Delta q^c(r)}$ , this amounts to checking that  $e^X > 1 + X$  for  $X = r\Delta\theta\Delta q^c(r) > 0$  which obviously holds.

- (46) is strictly satisfied when

$$e^{-r\Delta\theta\Delta q^c(r)} < 1 + \lambda_1^c r \Delta\theta\Delta q^c(r) \quad (67)$$

which holds since  $\lambda_1^c > 0$  and  $\Delta q^c(r) > 0$ .

- The coalition incentive constraint (23) and the condition  $\frac{\lambda_1^c}{p_{21}} - \frac{\lambda_2^c}{p_{22}} > 0$  hold if we take  $\lambda_2^c = p_{12}e^{-r(w_{22}^c - \Delta\theta q_{22}^c(r))} = p_{12}$ , i.e., a multiplier in state  $\tau_2$  equal to the marginal utility of income of the supervisor when he faces an efficient agent times  $p_{12}$ . This amounts to taking also  $\mu_2^c = 0$  in state  $\tau_2$ . Moreover, note that all values  $\tilde{\lambda}_2^c \in [0, p_{12}]$  could alternatively be used. The reason is that telling the truth is a strict optimum for the coalition when  $(\theta_2, \tau_2)$  so that the coalition incentive constraint (23) is not binding and the value of  $\lambda_2^c$  is not pinned down by the optimization of the principal's problem.

Taking into account that  $\lambda_1^c = p_{11}e^{-r(w_{21}^c - \Delta\theta q_{21}^c(r))}$ , we have:

$$\frac{\lambda_1^c}{p_{21}} - \frac{\lambda_2^c}{p_{22}} = \frac{p_{11}}{p_{21}}e^{-r\Delta\theta\Delta q^c(r)} - \frac{p_{12}}{p_{22}}. \quad (68)$$

Comparing the right-hand sides of (39) and (40) and using the fact that  $X = r\Delta\theta\Delta q^c(r) > 0$ , we get:

$$\frac{p_{11}e^{-X}}{p_{21} + p_{11}Xe^{-X}} > \frac{1}{p_{22}} \left( p_{11} + p_{12} - \frac{p_{21}p_{11}e^{-X}}{p_{21} + p_{11}Xe^{-X}} \right)$$

which amounts to

$$\frac{p_{11}}{p_{21}} \left( \frac{p_{21}e^{-X}}{p_{21} + p_{11}Xe^{-X}} \left( 1 + \frac{p_{21}}{p_{22}} \right) - \frac{p_{21}}{p_{22}} \right) > \frac{p_{12}}{p_{22}}.$$

Therefore, a sufficient condition to ensure that the right-hand side of (68) will be strictly positive is

$$e^{-X} > \frac{p_{21}e^{-X}}{p_{21} + p_{11}Xe^{-X}} \left( 1 + \frac{p_{21}}{p_{22}} \right) - \frac{p_{21}}{p_{22}}$$

or

$$\frac{p_{21}}{p_{22}}e^X + 1 > \frac{p_{21}}{p_{21} + p_{11}Xe^{-X}} \left(1 + \frac{p_{21}}{p_{22}}\right)$$

which is true since, using  $e^X > 1$  (for  $r > 0$ ) the left-hand side above can thus be bounded below by  $\frac{p_{21}}{p_{22}} + 1$  which is greater than the right-hand side since  $\frac{p_{21}}{p_{21} + p_{11}Xe^{-X}} < 1$  (for  $r > 0$ ).

- A coalition  $(\theta_2, \tau_2)$  must not be willing to claim that it is  $(\theta_1, \tau_2)$ . This requires:

$$0 = p_{12}v(w_{12}^c - \Delta\theta q_{22}^c(r)) + p_{22}v(w_{22}^c) > p_{12}v(w_{12}^c - \Delta\theta q_{12}^c(r)) + p_{22}v(w_{12}^c - \Delta\theta q_{12}^c(r)).$$

This holds since  $w_{12}^c - \Delta\theta q_{12}^c(r) = w_{11}^c - \Delta\theta q_{12}^c(r) = \Delta\theta(q_{22}^c(r) - q_1^{fb}) < 0$ .

- A coalition  $(\theta_2, \tau_1)$  must not be willing to claim that it is  $(\theta_1, \tau_1)$ . This requires:

$$p_{11}v(w_{11}^c - \Delta\theta q_{21}^c(r)) + p_{21}v(w_{21}^c) > p_{11}v(w_{11}^c - \Delta\theta q_{11}^c(r)) + p_{21}v(w_{11}^c - \Delta\theta q_{11}^c(r)).$$

The left-hand side is strictly positive and the right-hand side is negative since again  $w_{11}^c - \Delta\theta q_{11}^c(r) = \Delta\theta(q_{22}^c(r) - q_1^{fb}) < 0$ .

- A coalition  $(\theta_2, \tau_2)$  must not be willing to claim that it is  $(\theta_1, \tau_1)$ . This requires:

$$0 = p_{12}v(w_{12}^c - \Delta\theta q_{22}^c(r)) + p_{22}v(w_{22}^c) > p_{12}v(w_{12}^c - \Delta\theta q_{11}^c(r)) + p_{22}v(w_{11}^c - \Delta\theta q_{11}^c(r)).$$

This holds since  $w_{12}^c - \Delta\theta q_{11}^c(r) = w_{11}^c - \Delta\theta q_{11}^c(r) = \Delta\theta(q_{22}^c(r) - q_{11}^c(r)) < 0$ .

- A coalition  $(\theta_2, \tau_1)$  must not be willing to claim that it is  $(\theta_1, \tau_2)$ . This requires:

$$p_{11}v(w_{11}^c - \Delta\theta q_{21}^c(r)) + p_{21}v(w_{21}^c) > p_{11}v(w_{11}^c - \Delta\theta q_{11}^c(r)) + p_{21}v(w_{12}^c - \Delta\theta q_{12}^c(r)).$$

The left-hand side is strictly positive and the right-hand side is negative since again  $w_{12}^c - \Delta\theta q_{12}^c(r) = w_{11}^c - \Delta\theta q_{11}^c(r) = \Delta\theta(q_{22}^c(r) - q_1^{fb}) < 0$ .

• **Unique Implementation:** Let us now describe how the principal can design out of equilibrium punishments to ensure unique Nash implementation if the agents fail to collude. Consider wages  $s_{ik}^c$  such that:

$$p(\theta_1|\tau_1)v(s_{112}^c) + p(\theta_2|\tau_1)v(s_{212}^c) > V_1^c \tag{69}$$

$$V_2^c = 0 > p(\theta_1|\tau_2)v(s_{112}^c) + p(\theta_2|\tau_2)v(s_{212}^c). \tag{70}$$



(69) makes impossible to sustain a non-truthful equilibrium when  $\tau_1$  realizes. (70) ensures that the supervisor does not want to lie when  $\tau_2$  realizes and the agent reports truthfully to the principal. Again, the monotone likelihood ratio property  $\frac{p_{11}}{p_{21}} > \frac{p_{12}}{p_{22}}$  ensures that such punishments exist. Similarly, wages  $s_{i2k}^c$  such that:

$$p(\theta_1|\tau_1)v(s_{121}^c) + p(\theta_2|\tau_1)v(s_{221}^c) < V_1^c \quad (71)$$

$$V_2^c = 0 < p(\theta_1|\tau_2)v(s_{121}^c) + p(\theta_2|\tau_2)v(s_{221}^c). \quad (72)$$

(71) ensures that the supervisor tells the truth when  $\tau_1$  realizes and the agent reports truthfully. (72) ensures that a non-truthful equilibrium does not exist when  $\tau_2$  realizes. Moreover, we impose  $t_{ijk} = -q_{ijk} = -\infty$  for  $j \neq k$  for the agent's punishments so that there cannot be any non-truthful equilibrium where the agent deviates both on his reports of type and signal.

### Proof of Proposition 5:

- Inserting the expressions for the  $p_{ij}$  obtained with the information structure characterized by  $Prob(\tau_i|\theta_i) = \epsilon$  gives the following expression of outputs (where we now make explicit the dependence of outputs on  $\epsilon$ ):

$$R'(q_{21}^c(\epsilon)) = \theta_2 + \Delta\theta \frac{\nu\epsilon e^{-r\Delta\theta\Delta q^c(\epsilon)}}{(1-\nu)(1-\epsilon) + \nu\epsilon r\Delta\theta\Delta q^c(\epsilon)e^{-r\Delta\theta\Delta q^c(\epsilon)}} \quad (73)$$

$$R'(q_{22}^c(\epsilon)) = \theta_2 + \frac{\Delta\theta}{(1-\nu)\epsilon} \left( \nu - \frac{\nu\epsilon(1-\nu)(1-\epsilon)e^{-r\Delta\theta\Delta q^c(\epsilon)}}{(1-\nu)(1-\epsilon) + \nu\epsilon r\Delta\theta\Delta q^c(\epsilon)e^{-r\Delta\theta\Delta q^c(\epsilon)}} \right). \quad (74)$$

When  $\epsilon = \frac{1}{2}$ , these equations reduce to:

$$R' \left( q_{21}^c \left( \frac{1}{2} \right) \right) = \theta_2 + \Delta\theta \frac{\nu e^{-r\Delta\theta\Delta q^c(\frac{1}{2})}}{1-\nu + \nu r\Delta\theta\Delta q^c(\frac{1}{2})e^{-r\Delta\theta\Delta q^c(\frac{1}{2})}}$$

$$R' \left( q_{22}^c \left( \frac{1}{2} \right) \right) = \theta_2 + \frac{2\Delta\theta}{1-\nu} \left( \nu - \frac{\nu(1-\nu)e^{-r\Delta\theta\Delta q^c(\frac{1}{2})}}{2(1-\nu + \nu r\Delta\theta\Delta q^c(\frac{1}{2})e^{-r\Delta\theta\Delta q^c(\frac{1}{2})})} \right),$$

and thus:

$$R' \left( q_{21}^c \left( \frac{1}{2} \right) \right) - R' \left( q_{22}^c \left( \frac{1}{2} \right) \right) = 2\Delta\theta \left( \frac{\nu e^{-r\Delta\theta\Delta q^c(\frac{1}{2})}}{1-\nu + \nu r\Delta\theta\Delta q^c(\frac{1}{2})e^{-r\Delta\theta\Delta q^c(\frac{1}{2})}} - \frac{\nu}{1-\nu} \right). \quad (75)$$

Assume first that  $q_{21}^c(\frac{1}{2}) < q_{22}^c(\frac{1}{2})$ , i.e.,  $\Delta q^c(\frac{1}{2}) > 0$ . Since  $\frac{\nu}{(1-\nu)e^X + \nu X}$  is monotonically decreasing in  $X$ , the right-hand side of (75) is thus negative and  $R'(q_{21}^c(\frac{1}{2})) - R'(q_{22}^c(\frac{1}{2})) < 0$ . Then, from  $R''(\cdot) < 0$   $q_{21}^c(\frac{1}{2}) > q_{22}^c(\frac{1}{2})$ , a contradiction. Starting from  $q_{21}^c(\frac{1}{2}) > q_{22}^c(\frac{1}{2})$ ,

we would get similarly  $q_{21}^c(\frac{1}{2}) > q_{22}^c(\frac{1}{2})$ , another contradiction. Thus, we necessarily have  $q_{21}^c(\frac{1}{2}) = q_{22}^c(\frac{1}{2}) = q_2^p$  such that  $R'(q_2^p) = \theta_2 + \frac{\nu}{1-\nu}\Delta\theta$ .

As  $\Delta q^c(\epsilon)$  is bounded above by  $(q_{22}^{sb} - q_{21}^{sb})$ , we can pass to the limit in (74) when  $\epsilon$  is close to one. We immediately obtain that  $q_{22}^c(\epsilon)$  converges to  $q_2^p$ .  $q_{21}^c(\epsilon)$  converges to a fixed value  $q_{21}^c(1)$  but the agent is almost never asked to produce this quantity.

• Let us now write the principal's expected payoff as:

$$W(\epsilon) = \sum_{i,j} p_{ij} (R(q_{ij}(\epsilon)) - \theta_i q_{ij}(\epsilon) - w_{ij}^c(\epsilon, q^c(\epsilon)))$$

where  $p_{ij}$  are functions of  $\epsilon$  as in Footnote 12. Note that, only

$$w_{21}^c(\epsilon, q^c(\epsilon)) = -\frac{1}{r} \ln \left( 1 + \frac{\nu\epsilon}{(1-\nu)(1-\epsilon)} r \Delta\theta \Delta q^c(\epsilon) e^{-r\Delta\theta\Delta q^c(\epsilon)} \right)$$

is both directly a function of  $\epsilon$  and indirectly through the value of  $\Delta\Delta q^c(\epsilon)$ . Others  $w_{ij}^c(\epsilon, q^c(\epsilon))$  only depend on  $\epsilon$  indirectly through the vector of outputs  $q^c(\epsilon)$ .

From the fact that  $p_{21}(\epsilon)w_{21}^c(\epsilon, q^c(\epsilon))$  goes to zero as  $\epsilon$  goes to 1, the fact that  $q_{22}^c(\epsilon)$  converges to  $q_2^p$  and  $q_{21}^c(\epsilon)$  converges to a fixed value  $q_{21}^c(1)$  which is almost never asked to be produced as  $\epsilon$  goes to 1, we have immediately that  $W(\frac{1}{2}) = W(1)$ .

Using the Envelope Theorem to differentiate  $W(\cdot)$  with respect to  $\epsilon$ , and the observation that  $q_{11}^c(\epsilon) = q_{12}^c(\epsilon)$  and  $w_{11}^c = w_{12}^c$  we find that:

$$W'(\epsilon) = (1-\nu) \left( R(q_{22}^c(\epsilon)) - \bar{\theta}q_{22}^c(\epsilon) - (R(q_{21}^c(\epsilon)) - \bar{\theta}q_{21}^c(\epsilon) - w_{21}^c(\epsilon, q^c(\epsilon))) - (1-\epsilon) \frac{\partial w_{21}^c}{\partial \epsilon}(\epsilon, q^c(\epsilon)) \right).$$

Using the expression of  $w_{21}^c$  given above, we thus obtain that:

$$W'(\epsilon) = (1-\nu) \left( R(q_{22}^c(\epsilon)) - \bar{\theta}q_{22}^c(\epsilon) - (R(q_{21}^c(\epsilon)) - \bar{\theta}q_{21}^c(\epsilon)) \right) - \frac{1-\nu}{r} \ln \left( 1 + \frac{\nu\epsilon}{(1-\nu)(1-\epsilon)} r \Delta\theta \Delta q^c(\epsilon) e^{-r\Delta\theta\Delta q^c(\epsilon)} \right) + \frac{1}{r} \frac{\nu r \Delta\theta \Delta q^c(\epsilon)}{(1-\epsilon) e^{r\Delta\theta\Delta q^c(\epsilon)} + \nu \epsilon r \Delta\theta \Delta q^c(\epsilon)}.$$

Since  $\Delta q^c(1) \neq 0$  is bounded, we have  $W'(1) = -\infty$ . Moreover,  $W'(\frac{1}{2}) = 0$ . The non-monotonicity of the expected welfare immediately follows from the previous observations. Hence, there exists an optimal information structure which is interior.