Information Transmission between Principals[1]

Giacomo Calzolari[2] Alessandro Pavan[3]

[1]We would like to thank Jean Tirole for his constant encouragement and supervision. We are also grateful to Vincenzo Denicolò, Roberta Dessì, Guido Fribel, Doh-Shin Jeon, Jean Jacques Laffont, David Martimort, Massimo Motta, Marek Pycia, Patrick Rey, Carlo Scarpa and Lars Stole, as well as seminar participants at the Universities of Bologna, CORE Louvaine-la-Neuve and Toulouse for discussions and useful suggestions. All remaining errors remain our own.

[2]Gremaq, University of Toulouse and University of Bologna. E-mail: calzolar@spbo.unibo.it

[3]Gremaq, University of Toulouse, Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France. Tel: 0033-5-61128569. Fax: 0033-5-61225563. E-mail: alessandro.pavan@univ-tlse1.fr.

Abstract

This paper examines a contracting environment where two principals sequentially interact with a common agent. The agent's private information includes his original "type" and all payoff-relevant decisions that result from upstream contracting.

We endogenize the information flow between the two contractual relationships and show how the upstream principal may benefit from offering a contract that strategically discloses information to the downstream principal. Information transmission is motivated by rent extraction: By providing the receiver with information that is correlated with the agent's type and his past contractual experience, the upstream principal succeeds in appropriating part of the agent's surplus from downstream contracting. We characterize the equilibrium contracts and propose a simple implementation in which the agent receives private or public allocations.

Finally, we show that information transmission between principals, although strategic, may well result in an increase in welfare since it reduces the (expected) distortions associated to the asymmetry of information.

Keywords: *asymmetry of information, sequential common agency, information disclosure, mechanism design, multiprincipals, sequential contracting.*

1 Introduction

Several contracting environments can be described as dynamic games in which different principals sequentially interact with an informed agent (we adopt the convention of using female pronouns for the principals and masculine pronouns for the agent). A venture capitalist (upstream principal) who offers a financing contract to a start-up (agent) usually anticipates that the firm will also contract with suppliers, retailers and perhaps regulatory agencies (downstream principals). In organizations, a division manager who hires a worker typically expects that the latter will be asked to change division after a while and pass under the control of other managers. Moreover, workers rarely stay within the same organization for all of their life; rather, they enter into a sequence of job relationships with different employers. In politics, the ruling administration that signs a procurement contract with a contractor, or that offers a trade policy to a lobby, knows that its counterpart will also interact with the next appointed administration. Similarly, in e-commerce, a vendor who sets up a menu of contract-offers, usually expects her customers will need to procure complementary products and services also from other vendors.

Under sequential contracting, a downstream principal who offers a contract to an agent usually makes the best possible use of any information that derives from the agent's upstream contractual experiences. For, the value the agent attaches to a contract may well depend on the decisions that have been taken through other contracts: For example, in a labor relationship the ability for a worker to perform a certain task is typically influenced by the type of jobs exerted in the past. Similarly, in a trade contract, the willingness to pay for a certain product or service usually depends on the complementarity with products and services acquired from other suppliers. In such cases, the more a principal knows of her agent's previous contractual activity, the better she can design her contract.

The observation of the elements of past contractual relationships, whenever possible, is also a useful signal of the agent's private information. By observing a consumer's shopping activity a vendor can infer the preferences of a customer. Similarly, the covenants of a financial contract between a venture capitalist and an entrepreneur convey information on the profitability of the borrower's project and may affect the result of subsequent financing stages.

The dynamic interaction between different principals suggests the possibility of strategic behavior: an upstream principal will typically try to take advantage of her Stackelberg position with respect to downstream principals by designing the contractual relationship in such a way that optimally controls for the influence it has on downstream contracting. There are two ways a contract can affect another one. First directly, though the decisions that are stipulated. Second, through the information that the contract discloses to outsiders. This is the main focus of this paper. We investigate how a principal should design a mechanism that simultaneously screens the private information of her agent and strategically transmit it to another principal.

Information transmission to an outside party does occur in several bilateral contractual relationships. Consider, for example, e-commerce. An e-consumer (agent) who buys a product on line reveals (at least partially) his preferences to the e-seller (principal). For example, the path the agent follows during the visit of a website is a signal of his interests, as well as the final choice of a product, or the request of a certain service, represents valuable information of the agent's preferences. When shopping on line consumers are often notified that other sellers might have access to some of the information consumers are revealing to the current counterpart (this can be read, for example, in the privacy policy webpages of the main e-vendors)[1].

Other examples of information transmission between two contractual relationships can be found in labor and insurance contracts. When a worker leaves a job, he usually receives a letter of recommendation to give to a future employer; this letter signals the information that the principal has learned from the interaction with the agent, like for example his talent, fairness and ability to cope with colleagues. Similarly, insurees who change company are notified that part of the information collected by the insurance company on the insuree's characteristics, like his propensity towards risk, will be transmitted to the new company.

What motivates a principal to design a contract that discloses the agent's information? How does it occur? And is this necessary harmful for the agent?

This paper develops a general model of sequential contracting with multiple principals that endogenizes the informational linkage between two contracts. We suggest a possible rationale for information disclosure and derive some important welfare implications.

In our theory information transmission between the two contractual relationships is motivated by rent extraction. Through an optimal disclosure policy, an upstream principal can in fact increase the agent's surplus from downstream contracting by reducing the distortions that are associated to the asymmetry of information. The upstream principal has an interest in favoring the agent since she can make the latter pay for the increase in his expected utility.

To illustrate the point, consider the following example. Suppose there are two differentiated e-sellers (principals) that sequentially contract with a common buyer (the agent). The two contracts consist of a price the buyer has to pay for the quality (or quantity) of a certain product, or service. Suppose that the buyer is either unsophisticated, or sophisticated, in which case he has a higher marginal valuation for quality. As it is well known from the theory of price discrimination, if the probability of facing a sophisticated consumer is high, then the profit-maximizing price is so much distorted that it discourages the purchase from the unsophisticated consumer and leaves no surplus to the buyer. In a sequential contracting environment the beliefs of a seller can be influenced by the information provided by another vendor who previously traded with the same agent. When such information induces the current seller to perceive the probability of an unsophisticated buyer as being sufficiently high, then the best she can do is to make a less distorted offer that is always accepted and that

[1]E-consumers often receive "cookies" during their e-shopping. These cookies are simple devices to record the agent's information so that subsequent interactions either with the same vendor, or with trustworthy business partners, will be modelled in a more personalized way.

necessarily leaves a rent to the consumer.

This idea is consistent with what suggested in several privacy-policy webpages: Consumers who accept to let vendors share information with trustworthy third parties may enjoy a better future shopping experience in terms of personalized price-discounts, and in general more favorable deals. Not surprising, those vendors who offer to consumers an optimal information disclosure policy face higher expected profits: indeed, the price consumers pay for a product, or a service, incorporates also the value consumers expect from future contractual experiences.

Our analysis suggests that information transmission is profitable when there is some complementarity between the two principals' decisions. This prediction is also in the spirit of several e-vendors' webpages in which is noted that consumers' private information should be shared only with "*carefully selected business partners that offer complementary services and products*" (see, for example, www.drugstore.com, www.yahoo.com, www.networksolutions.com).

The flow of information occurs even if the upstream principal is not directly interested in the decisions that are stipulated with the downstream principal, nor can she make the latter pay for the information she receives. We acknowledge that both information trade and direct payoff-relevant externalities represent good explanations for which a principal can be interested in disclosing the information she learns from her agent. Nevertheless, by leaving these effects out of the analysis we succeed in isolating the rent-extraction behavior that characterize the design of contracts under asymmetric information.

Our analysis finally suggests that privacy-protecting laws that prevent information disclosure on consumers' preferences are not necessarily welfare increasing. We compare the equilibrium contracts when principals cannot release the agent's private information with the contracts that emerge when some information is strategically disclosed. When information transmission is allowed a downstream principal obtains a finer information structure and reduces the distortions that are due to the asymmetry of information. At the same time, since information transmission enables an upstream principal to reduce the informational rent she leaves to the agent, the trade-off between efficiency and rent-extraction in her optimal contract also changes in favor of efficiency. It follows that information transmission may well result in an increase of welfare.

Related literature.

- *Common Agency.*

The literature on common agency has received much attention in the last few years. Bernheim and Whinston (1986) extend the principal-agent model with moral hazard (see, for example, Grossman and Hart (1983)) to the case of multiple principals. Similarly, Martimort (1991, 1996), Martimort and Stole (1999a,b), Mezzetti (1997) and Stole (1991) extend the principal-agent model under adverse selection to a contracting environment with multiple mechanism designers. Biglaiser and Mezzetti (1993) provide an extensive analysis of common agency in a framework with both adverse selection and moral hazard. Dixit et Al. (1997) develope a model of common agency without information asymmetries but they allow for general preferences and in particular nontransferable utilities.

A standard assumption in the common agency literature is that competition between principals is simultaneous, in that each principal simultaneously offers to an agent a contract and then the agent makes his decisions; a decision can be a (nonverifiable) action, the choice of the principal to deal with, or an option specified in the contract.

In the real world, common agency relationships need not take place simultaneously; rather principals sequentially contract with the agent. For example, a manufacturer that is going to sell its products to a retailer will typically find that the latter already contracted with other manufacturers in the past.

To our knowledge, Baron (1985), Martimort (1999) and Prat and Rustichini (1998) are the only papers that examine common agency models with a sequential timing. Baron (1985) makes the strong assumption that all the information elicited by one principal is made public also to the other, so that the leader can free-ride the incentives problem of the follower. Martimort (1999) drastically simplifies the incentives problem of the common agency game by assuming the two principals have perfectly correlated decisions. Prat and Rustichini (1998) focus on a moral hazard model in which the two principals sequentially and independently offer contracts to influence the action of a common agent. The sequentiality in these models is partial since they assume principals sequentially offer their contracts but the agent has to decide only after receiving all proposals. In this paper, we allow the agent first to contract with one principal and then with another one: It is indeed in such a situation that the information flow between the two contractual relationships is an issue. With this respect, although we deal with common agency, our paper is also related to the dynamic single-principal-agent literature (see, for example, Baron and Besanko (1984) and Laffont and Tirole (1987, 1988)).

- *Information Disclosure Policies.*

The possibility of information sharing between firms has been examined in the literature of oligopolistic competition (see, for example Raith (1998) for a survey) and in the financial intermediation literature (Padilla and Pagano (1998) and Pagano and Jappelli (1993)). In these models the informed parties can decide to strategically share information with rivals before competing. In our model principals are uninformed and learn from the agent through the screening contract they design. Information transmission between different contractual relationships is motivated by a rent extraction behavior, rather than by the desire of coordination. Upstream principals commit to disclose information that is correlated with the agent's type and his past contractual experience in order to appropriate the agent's surplus from an external contractual relationship. In this respect our analysis is similar to Lizzeri (1999) who discusses the role of intermediaries who search out the information of privately informed parties and then decide what to disclose to uninformed parties. The main difference with respect to his paper is that we explicitly model information acquisition through the design of a screening contract, whereas he considers an exogenously specified technology through which an intermediary can test the quality of a product of a seller.

None of these papers considers the issue which is our main focus here: "Does a principal want to design a mechanism that simultaneously screens the private information of her agent and strategically communicates it to another principal"?

Outline.

The rest of the paper is organized as follows. Section 2 introduces a general model for sequential contracting with multiple principals. Section 3 characterizes the pure-strategy equilibrium in case principals do not share any information; it solves for general payoffs and provides an example that well describes a trading game between two sellers and a common buyer. Section 4 is dedicated to the analysis of information transmission between principals. Section 5 concludes. All proofs are in the Appendix.

2 The Model

A single agent, A, sequentially interacts with two principals[2], P_i for $i = 1, 2$. Principal *one* will be referred to as the *leader* and principal *two* as the *follower*. The set of players in our model is thus represented by $I = \{P_1, P_2, A\}$.

Allocations.

Each principal can contract with the agent over an allocation $y_i \in Y_i$, where Y_i stands for the set of feasible allocations for principal P_i. For simplicity, we consider allocations $y_i = \{x_i, t_i\}$ that consist of a *decision* $x_i \in X_i \subset \Re_+$ and a *monetary transfer* $t_i \in \Re$ that is paid by the agent to principal P_i. The decision x_i might represent the quantity (or quality) of principal i's product, or the (verifiable) task principal i asks the agent to perform. The transfer may be positive or negative depending on the particular setting under consideration. In case of two sellers that compete in nonlinear prices, t_i are positive transfers and represent the price that each principal charges for the quantity (or quality) x_i that is sold to the agent. If the two principals stand for two divisions of an organization that sign contracts with an external consultant, t_i are negative transfers and represent the payment the two divisions make to the common consultant in exchange of his advice.

Let $y = (y_1, y_2) \in Y_1 \times Y_2$ be the vector representation of a pair of feasible allocations. For simplicity, we assume that the set of feasible allocations for principal i does not depend on the allocation chosen by the other principal[3].

Payoffs functions.

The agent's preferences are represented by the payoff function $U_A(y_1, y_2, \theta)$.

The two principals have respectively payoff functions $U_i(y_1, y_2, \theta)$ for $i = 1, 2$.

We make the following assumptions on players' payoffs.

[2] Our analysis could be extended to cover the possibility that several principals sequentially interact with several agents. For simplicity, we prefer to focus on a simple, yet general, model in which there are only two principals and a single (common) agent.

[3] More generally, the two principals might face a joint feasibility constraint represented, for example, by a frontier $g(y_1, y_2) = 0$.

A1: (quasi-linearity):

$$U_i(y_1, y_2, \theta) = v_i(x_1, x_2, \theta) + t_i, \text{ for } i = 1, 2$$

$$U_A(y_1, y_2, \theta) = v_A(x_1, x_2, \theta) - t_1 - t_2,$$

where $v_i(x_1, x_2, \theta)$ and $v_A(x_1, x_2, \theta)$ are twice differentiable and strictly concave functions in $x = (x_1, x_2) \in \Re^2_+$.

A2: (two types): $\Theta = \left\{\overline{\theta}, \underline{\theta}\right\}$ and $\Pr(\overline{\theta}) = p = 1 - \Pr(\underline{\theta})$.

A3: (Spence-Mirrlees): $\frac{\partial v_A(x,\bar{\theta})}{\partial x_i} \geq \frac{\partial v_A(x,\underline{\theta})}{\partial x_i}$, for any $x \in \Re^2_+$ and for any $i = 1, 2$.

A4: $v_A(x, \bar{\theta}) \geq v_A(x, \underline{\theta})$, for any $x \in \Re^2_+$.

A5: (weak complementarity) $\frac{\partial^2 v_A(x,\theta)}{\partial x_1 \partial x_2} \geq 0$, for any $x \in \Re^2_+$ and for any θ.

Assumptions A1, A3 and A4 are standard in adverse selection models. Assumption A2 is not actually needed; it just simplifies the analysis of the information flow between the two principals.

Assumption A5 limits attention to decisions x_1 and x_2 that are either complements or independent in the agent's payoff. Our decision to rule out in this paper the substitutes case is motivated by two reasons. First, as suggested in Mezzetti (1997), common agency is more likely to occur when the agent faces complementarities between the two decisions, like in case of a worker who performs related tasks on behalf of multiple employers, a buyer who procures complementary products or services from multiple vendors, or an entrepreneur who is jointly-financed by two banks. Second, the substitutes case would require a separated analysis that cannot be presented, in a simple way, within the same homogenous framework developed for the complements case.

Note that we allow each principal's payoff to depend directly on the other principal's allocation and on the agent's type. For example, a venture capitalist and a monitor (investment banker) might be interested in the ability of an entrepreneur (here represented by θ) and in the impact of each other's activity. Similarly, two duopolists that procure the same input from a common manufacturer typically exert direct externalities in that their final profits directly depend on the production capacity of the rival firm which is determined by the amount of input that is acquired from the common supplier.

We model incomplete information by assuming that the agent has a type $\theta \in \Theta$. Only the agent knows the exact realization of the random variable $\widetilde{\theta}$ (his private information), whose probability distribution $p(\cdot)$ is common knowledge across players. We assume that Θ is a finite set and we will often limit ourselves to the case in which it contains only two elements, say $\overline{\theta}$ and $\underline{\theta}$ respectively with probability p and $1 - p$.

Contracts.

The two principals offer to the agent two mechanisms (also referred to as contracts). A mechanism is a mapping from a message space to the set of feasible allocations. We denote with $\mathcal{M}_i$ the message space of principal i and with $m_i \in \mathcal{M}_i$ a single element of such a space. A mechanism for principal i is therefore represented by $\pi_i : \mathcal{M}_i \rightarrow Y_i$ where $y_i = \pi_i(m_i)$ is the allocation that P_i assigns to the agent when the latter reports m_i. Although we focus here on *deterministic* mechanisms, the analysis can be extended to *stochastic contracts* in which $\pi_i(y_i/m_i)$ is the probability of allocation y_i contingent on message m_i. Let Π_i be the set of all possible feasible mechanisms for principal i.

Information Disclosure Policy.

Information transmission between the two principals is formally represented by a signal $s \in S$ that P_1 sends to P_2 at the beginning of stage 2. The signal can be either *soft* or *hard*. It is soft in case of a message that is sent from P_1 to P_2. It is hard, for example, in case it contains y_1 and this allocation is verifiable. The use of this abstract information technology enables us to study the optimal contracts in a general framework. The implementation of the optimal mechanism will shed some light on the possible interpretations of s.

To accommodate signalling from P_1 to P_2 we need to add to principal one's contract a *disclosure policy* δ that maps from messages $\mathcal{M}_1$ to the set of probability measures upon a (finite) set of signals S. Formally, $\delta : \mathcal{M}_1 \rightarrow \triangle(S)$. For each message m_1 selected by A, P_1 sends a signal s to P_2 with probability $\delta(s/m_1)$. P_2's mechanism will then be contingent on the signal s, i.e. $\pi_2 = \pi_2(s)$.

The agent's strategy $\pi_A = (\pi_A^1, \pi_A^2)$ specifies the reports to each principal as a function of the agent's information set, i.e. $m_1 = \pi_A^1(\theta, \pi_1)$, and $m_2 = \pi_A^2(\theta, \pi_{1,} y_1, s, \pi_2)$.

Timing: a sequential contracting game.

- At $t = 0$, the agent discovers his type which is a realization of the random variable $\widetilde{\theta}$ with support Θ and distribution $p(\cdot)$.
- At $t = 1$, principal P_1 offers to A a mechanism $\pi_1 \in \Pi_1$ and a disclosure policy δ. If the agent rejects π_1 the game ends and all players get their reservation payoffs that are normalized to zero[4]. If π_1 is accepted, the agent *secretly* reports a message m_1 to P_1, he receives the allocation $y_1(m_1)$ and the game evolves to stage two.
- At $t = 2$, P_2 observes a signal s and offers to the agent her own mechanism, $\pi_2 \in \Pi_2$. The agent can accept or reject it. If he rejects it, then the game ends and P_1, P_2 and A are left with the payoffs that derive from the first stage interaction. If the agent accepts π_2, he chooses a message $m_2 \in \mathcal{M}_2$ that he *secretly* reports to P_2.

The choice of this timing is motivated by three reasons. First, we believe that several contractual relationships do take place sequentially. Second, se-

[4]This assumption is much stronger than needed. For example, we could assume that the agent can reject π_1 and directly deal with P_2. Alternatively, we could assume that A cannot reject π_1, yet P_1 is constrained to give him the possibility to pick up a message m_1^0 which is associated with the null contract. By a null contract we mean a contract for which there is no trade between A and P_1. Yet, the null contract may involve some information transmission to P_2. For example, P_1 can always commit to inform P_2 about the agent's decision not to trade with her. For simplicity, we assume here that the game ends if the agent refuses to deal with P_1 so that we do not need to examine the out of equilibrium continuation game between A and P_2.

quential contracting games can be characterized by a strategic information flow between the two principals. In particular, the leader, anticipating that another principal will contract with the very same agent in the continuation subgame, may find it profitable to design a contract that discloses payoff-relevant information in order to affect the contract that P_2 will offer to A. Finally, as we show at the end of this section, under sequential contracting optimal mechanism can be derived using an intuitive extension of the Revelation Principle.

Commitment.

We assume that in each bilateral relation both principals can credibly commit to the contract π_i.

We also assume that P_1 can commit not to release more information than allowed by the contract she offers to A. For example, an e-seller may want to establish a reputation not to disseminate information that she is not allowed to disclose. At the end of Section 4 we will further discuss such assumption.

Observability.

We assume that P_1 is not exogenously constrained to let P_2 observe y_1. In some situations this assumption might not be adequate. For example, in case of government procurement contracts, the procurer may be obliged to disclose the final terms of the agreement with the contractor, like the number of units supplied and the price. In this case, the first principal might be forced to make the agent randomize on her contract choice in order not to perfectly inform the second principal (this possibility is studied in a common-agency framework by Calzolari and Pavan (2000a) and in the dynamic single-principal literature, for example, by Bester and Strausz (2000) and Laffont and Tirole (1988, 1990)).

Equilibrium.

Let $\pi = (\pi_1, \pi_2, \pi_A)$. A (pure-strategy) profile (π^*, δ^*) is a *Perfect Bayesian Equilibrium* for the sequential common agency game $\Gamma_{\mathcal{M}}$ with communication spaces $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ if and only if:

1. each principal selects a mechanism that is optimal given the agent's and the other principal's strategies;
2. P_1 selects an optimal disclosure policy;
3. for each signal s, P_2 updates her beliefs using Bayes' rule;
4. the agent announces payoff-maximizing messages.

Direct Revelation Mechanisms.

The Revelation Principle (hereafter RP) has proved particularly useful in contract theory since it offers a simple way to characterize the set of possible feasible allocations that can be implemented when agents have private information. Unfortunately, when multiple principals interact with a common agent the standard version of the RP has been proved to be usually invalid (see Martimort and Stole (1999)). As a result, in the last few years many attempts have been done to provide extensions of the standard RP for contracting environments with multiple mechanism designers. Epstein (1999) and Epstein and Peters (1999) have suggested a RP for simultaneous common agency. Calzolari and Pavan (2000b) have introduced a very simple version of the RP for sequential common agency games. In that paper they show that at each stage contracting is Markovian in the sense that there is no loss of generality in assuming that each principal limits herself to simple direct revelation mechanisms in which the message space is the set of von Newmann and Morgenstern equivalence classes. We recall here the main definitions and results. For any further discussion we refer the reader to that paper.

Definition 1 *(Extended type)*
The agent's extended type θ_i^E in a contractual relationship with principal i is the profile of all payoff-relevant information available to the agent at $t = i$ for $i = 1, 2$.

Let Θ_i^E denote the *extended* type space at $t = i$. Clearly, for $t = 1$ we have that $\Theta_1^E = \Theta$. At t_2, the extended type space is $\Theta_2^E = \Theta \times Y_1$, since the only payoff-relevant information regards the agent's original type and the allocation y_1 selected through the mechanism π_1.

A *direct revelation mechanism* for principal i is a mechanism in which the message space coincides with the (extended) type space, $\mathcal{M}_i = \Theta_i^E$ and in which the agent truthfully reports his (extended) type to principal P_i.

Proposition 1 *(The Revelation Principle)*
In a sequential common agency game there is no loss of generality in assuming that principals use direct revelation mechanisms. For any equilibrium with arbitrary message spaces, $\mathcal{M}$, there always exists an equilibrium in direct revelation mechanisms that induces the same final probability distribution over Y, for any $\theta \in \Theta$.

This RP (Calzolari and Pavan 2000b) is general enough to include the possibility of information transmission from one principal to another one: Any information flow that can be generated through an indirect mechanism, can always be replicated with a direct revelation mechanism.

Proposition 2 *Without loss of generality, P_1 makes her mechanism public. For any equilibrium strategy profile in which the mechanism π_1 is not announced to P_2, there always exists another equilibrium strategy profile such that:*

i) π_1 is publicly announced,

ii) it induces the same distribution over Y.

In the next sections we recursively apply Propositions 1 and 2 to characterize the optimal contracts offered by the two principals.

3 Contracts without Information Transmission

In this section we assume that P_1 does not disclose the agent's private information to P_2. This scenario can represent, for example, a situation in which P_1 does not have the possibility to signal the agent's type to P_2, she lacks of any commitment to contract with A on the information transmission to P_2, or she is prevented by law from disclosing any information on the agent's characteristics. Strategic information transmission between the two principals will be addressed in Section 4.

From the previous section we know that the RP applies so that there is no loss of generality in looking at equilibria in which the two principals use direct revelation mechanisms. Formally, we are looking for a pair of contracts

$$\pi^* = \{\pi_1^*, \pi_2^*\},$$

with $\pi_1^* : \Theta_1^E \to Y_1$ and $\pi_2^* : \Theta_2^E \to Y_2$ that map from the agent (extended) type space to the feasible allocation set for each principal. We recall that the extended type space for the leader coincides with the agent's physical type space, Θ. For P_2 the extended type space is enriched by the interaction between A and P_1, so that all the payoff-relevant information for P_2 is captured by $\Theta_2^E \equiv \Theta \times Y_1$.

In case of pure-strategy equilibria[5], there is perfect correlation between the agent's physical type, θ, and the allocation $y_1(\theta)$ the agent receives through the mechanism π_1. It follows that P_2 faces an agent with only two extended types $\overline{\theta}_2^E = (\overline{\theta}, \overline{x}_1, \overline{t}_1)$, $\underline{\theta}_2^E = (\underline{\theta}, \underline{x}_1, \underline{t}_1)$, respectively with probability p and $1-p$.

Both mechanisms must induce the agent to truthfully report his (extended) type to each principal, i.e. $\pi_A^{1*}(\theta/\pi_1^*) = \theta$ and $\pi_A^{2*}(\theta_2^E/\pi_1^*, \pi_2^*) = \theta_2^E$ for all $\theta \in \Theta$ and $y_1 \in Y_1$.

Finally, optimality conditions must be satisfied so that each principal's mechanism is a best response to the other principal's mechanism.

The equilibrium of the game can be found by backward induction.

Let $\overline{U}_A = v_A(\bar{x}_1, \bar{x}_2, \bar{\theta}) - \bar{t}_1 - \bar{t}_2$ and $\underline{U}_A = v_A(\underline{x}_1, \underline{x}_2, \underline{\theta}) - \underline{t}_1 - \underline{t}_2$ be the equilibrium payoff respectively for an agent with physical type $\bar{\theta}$ (also referred to as the "good", "high" or "efficient" type) and $\underline{\theta}$ ("bad", "low", "inefficient" type).

At $t = 2$ the follower offers to A a contract $\pi_2 = \{(\underline{x}_2, \underline{t}_2), (\bar{x}_2, \bar{t}_2)\}$ that

[5] Throughout this paper we limit our analysis to pure-strategy sequential equilibria in which none of the two principals randomizes on the contract offer to A and such that the agent follows a pure-strategy with both principals. In some situations, there might also exist mixed-strategy equilibria in which, for example, P_1 induces the agent to randomize over the choice of y_1 so that he will receive an extra informational rent with P_2. This possibility is examined in Calzolari and Pavan (2000a).

solves the following program:

$$\mathcal{P}_2 : \begin{cases} \underset{\pi_2}{Max}\; p\left[v_2(\bar{x}_1,\bar{x}_2,\bar{\theta})+\bar{t}_2\right]+(1-p)\left[v_2(\underline{x}_1,\underline{x}_2,\underline{\theta})+\underline{t}_2\right] & \\ \text{s.t.} & \\ \overline{U}_A \geq v_A\left(\bar{x}_1,0,\bar{\theta}\right)-\bar{t}_1, & (\overline{IR}_2) \\ \underline{U}_A \geq v_A\left(\underline{x}_1,0,\underline{\theta}\right)-\underline{t}_1, & (\underline{IR}_2) \\ \overline{U}_A \geq v_A\left(\bar{x}_1,\underline{x}_2,\bar{\theta}\right)-\bar{t}_1-\underline{t}_2, & (\overline{IC}_2) \\ \underline{U}_A \geq v_A\left(\underline{x}_1,\bar{x}_2,\underline{\theta}\right)-\underline{t}_1-\bar{t}_2, & (\underline{IC}_2) \end{cases}$$

The individual rationality constraints $(\overline{IR}_2)$ and $(\underline{IR}_2)$ ensure that the agent accepts the contract offered by P_2. Note that the (type-dependent) reservation utility is the payoff that the agent already obtained at $t=1$ from the contract with P_1. The incentive compatibility constraints $(\overline{IC}_2)$ and $(\underline{IC}_2)$ guarantee that the agent has the correct incentives to announce his extended type to P_2.

At $t=1$, the leader anticipates the reaction $\pi_2(\pi_1)$ which is the solution of $\mathcal{P}_2$ and commits to a contract $\pi_1=\{(\underline{x}_1,\underline{t}_1),(\bar{x}_1,\bar{t}_1)\}$ that solves

$$\mathcal{P}_1 : \begin{cases} \underset{\pi_1}{Max}\; p\left[v_1(\bar{x}_1,\bar{x}_2(\pi_1),\bar{\theta})+\bar{t}_1\right]+(1-p)\left[v_1(\underline{x}_1,\underline{x}_2(\pi_1),\underline{\theta})+\underline{t}_1\right] & \\ \text{s.t.} & \\ \overline{U}_A \geq 0, & (\overline{IR}_1) \\ \underline{U}_A \geq 0, & (\underline{IR}_1) \\ \overline{U}_A \geq v_A\left(\underline{x}_1,x_2(\bar{\theta},\underline{x}_1),\bar{\theta}\right)-\underline{t}_1-t_2(\bar{\theta},\underline{x}_1), & (\overline{IC}_1) \\ \underline{U}_A \geq v_A\left(\overline{x}_1,x_2(\underline{\theta},\overline{x}_1),\underline{\theta}\right)-\overline{t}_1-t_2(\underline{\theta},\overline{x}_1), & (\underline{IC}_1) \end{cases}$$

where $(x_2(\bar{\theta},\underline{x}_1),t_2(\bar{\theta},\underline{x}_1))$ and $(x_2(\underline{\theta},\bar{x}_1),t_2(\underline{\theta},\bar{x}_1))$ represent the allocations that respectively the high and the low type receive with P_2, when they misreport to P_1. The two individual rationality constraints guarantee that the agent accepts π_1, whereas the two IC constraints that he truthfully reports his private information.

To solve the two programs, we make use of the following lemma on monotonicity of principals' decisions.

Lemma 1 *(Monotonicity)*
Under Assumptions A1-A5, both principals's decisions are monotonic, i.e. $\bar{x}_1 \geq \underline{x}_1$ *and* $\bar{x}_2 \geq \underline{x}_2$.

Proof. See Appendix.
As in the single-principal case, when the two decisions, x_1 and x_2, are complements in the agent's payoff function, a high type must receive more of each decision than the low type in order to have the correct incentives to truthfully report his private information.

In Proposition 3 we use Lemma 1 to characterize P_2's optimal contract.

Proposition 3 *(The follower's optimal contract)*

P_2 offers to A a contract π_2 that is characterized by two decisions, $\bar{x}_2, \underline{x}_2$, with $\bar{x}_2 \geq \underline{x}_2 \geq 0$, such that

$$\bar{x}_2 = \arg\max_{x_2 \geq 0} v_2(\bar{x}_1, x_2, \bar{\theta}) + v_A\left(\bar{x}_1, x_2, \bar{\theta}\right) - v_A\left(\bar{x}_1, 0, \bar{\theta}\right),$$

$$\underline{x}_2 = \arg\max_{x_2 \geq 0} (1-p)\left[v_2(\underline{x}_1, x_2, \underline{\theta}) + v_A\left(\underline{x}_1, x_2, \underline{\theta}\right) - v_A\left(\underline{x}_1, 0, \underline{\theta}\right)\right] - pR_2(x_2),$$

where

$$R_2(x_2) = v_A\left(\bar{x}_1, x_2, \bar{\theta}\right) - v_A\left(\bar{x}_1, 0, \bar{\theta}\right) - v_A\left(\underline{x}_1, x_2, \underline{\theta}\right) + v_A\left(\underline{x}_1, 0, \underline{\theta}\right)$$

is the informational rent that P_2 must leave to the efficient agent.

Proof. See Appendix.

Proposition 3 shows how P_2's decisions are influenced by P_1's decisions. As in standard principal-agent models under adverse selection, P_2 must leave a rent to the efficient type in order to induce him to truthfully report his private information. This informational rent can be rewritten in a more familiar way as

$$R_2(\underline{x}_2) = \widehat{v}_A\left(\underline{x}_2, \bar{\theta}\right) - \widehat{v}_A\left(\underline{x}_2, \underline{\theta}\right),$$

where

$$\widehat{v}^A\left(\underline{x}_2, \bar{\theta}\right) = v_A\left(\bar{x}_1, \underline{x}_2, \bar{\theta}\right) - v_A\left(\bar{x}_1, 0, \bar{\theta}\right)$$

and

$$\widehat{v}_A\left(\underline{x}_2, \underline{\theta}\right) = v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) - v_A\left(\underline{x}_1, 0, \underline{\theta}\right),$$

are respectively the additional utility the high and the low type agent obtain from $\underline{x}_2$, conditional on having participated (and truthfully reported) to P_1's mechanism.

Note that the interaction between A and P_1 transforms the agent's payoff vis à vis P_2 from $v_A(0, x_2, \theta)$ to $\widehat{v}_A\left(x_2, \theta\right)$, where $\widehat{v}_A\left(x_2, \theta\right) \geq v_A(0, x_2, \theta)$ since the two decisions are complements in the agent's utility function.

The decision for the high type is the same that would emerge under full information.

The decision $\underline{x}_2$ is downward distorted in order to limit the informational rent P_2 leaves to the high type. It is worth mentioning that if π_1 is characterized by a strongly separating allocation, i.e. if $\bar{x}_1 \gg \underline{x}_1$, then P_2 may be better off when she "shuts down" the low type and sets $\underline{x}_2 = 0$. Hence, in a sequential common agency game "shutdown" can be endogenously determined by the leader through the choice of her contract. In Section 4 we show that this can be done also by manipulating P_2's posterior beliefs.

We are now in a position to analyze P_1's optimal contract.

The following lemma characterizes the two incentives constraints in $\mathcal{P}_1$.

Lemma 2 *When the agent lies to P_1, at $t = 2$ he receives with P_2 the allocation designed for the low (extended) type, i.e. $x_2(\bar{\theta}, \underline{x}_1) = x_2(\underline{\theta}, \bar{x}_1) = \underline{x}_2$.*

Proof. See Appendix.

Proposition 4 *(The leader's optimal contract)*

At $t = 1$, the leader offers to A a contract π_1^* that is characterized by two decisions, $\bar{x}_1^*$, $\underline{x}_1^*$ with $\bar{x}_1^* \geq \underline{x}_1^* \geq 0$, that maximize

$$\begin{aligned} U_1(\bar{x}_1, \underline{x}_1) &= p\left[v_1(\bar{x}_1, \bar{x}_2(\pi_1^*), \bar{\theta}) + v_A(\bar{x}_1, 0, \bar{\theta})\right] - pR_1(\bar{x}_1, \underline{x}_1)\right] + \\ &+ (1-p)\left[v_1(\underline{x}_1, \underline{x}_2(\pi_1^*), \underline{\theta}) + v_A(\underline{x}_1, 0, \underline{\theta})\right] \end{aligned}$$

where $R_1(\bar{x}_1, \underline{x}_1)$ is the informational rent that P_1 must leave to the efficient agent. This is given by

$$R_1(\bar{x}_1, \underline{x}_1) = R(\bar{x}_1, \underline{x}_1) - R_2(\bar{x}_1, \underline{x}_1),$$

where

$$R(\bar{x}_1, \underline{x}_1) = v_A\left(\underline{x}_1, \underline{x}_2(\pi_1^*), \bar{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2(\pi_1^*), \underline{\theta}\right)$$

is the total rent that the efficient agent obtains from the two principals and $R_2(\bar{x}_1, \underline{x}_1)$ is the informational rent P_2 leaves to the high type, as described in Proposition 3.

Proof. See Appendix.

To better understand Proposition 4, let Δ_θ denote the utility differential function between the high and the low type, so that

$$\Delta_\theta v_A\left(\underline{x}_1, \underline{x}_2\right) = v_A\left(\underline{x}_1, \underline{x}_2, \bar{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) > 0.$$

To induce the agent to reveal his private information, the leader must reduce the transfer she charges to a high type by

$$R(\bar{x}_1, \underline{x}_1) = \Delta_\theta v_A\left(\underline{x}_1, \underline{x}_2\right).$$

If principal two did not exist, the informational rent P_1 should leave to A would simply be $\Delta_\theta v_A\left(\underline{x}_1, 0\right)$. The presence of a subsequent interaction with P_2 obliges P_1 to increase this informational rent by $R(\bar{x}_1, \underline{x}_1) - \Delta_\theta v_A\left(\underline{x}_1, 0\right) > 0$.

At the same time, the contractual relationship between P_1 and A increases the informational rent the agent can obtain from P_2 so that P_1 can make the agent pay for this service by increasing his transfer of R_2.

It follows that the "net informational rent" that P_1 must leave to the high type is

$$\begin{aligned} R_1(\underline{x}_1, \bar{x}_1) &= R(\bar{x}_1, \underline{x}_1) - R_2(\bar{x}_1, \underline{x}_1) \\ &= \Delta_\theta v_A(\underline{x}_1, \underline{x}_2) - \Delta_\theta \widehat{v}_A(\underline{x}_2). \end{aligned}$$

Clearly, for the low type there is no informational rent with either principal.

Note that asymmetric information results in an indirect transfer $R_2(\bar{x}_1, \underline{x}_1)$ from the follower to the leader. On her part, the leader leaves $R(\bar{x}_1, \underline{x}_1)$ to the efficient agent.

Summarizing, the three players' expected payoffs under asymmetric information are

$$\begin{aligned}
U_A^* &= pR(\bar{x}_1^*, \underline{x}_1^*) \\
U_1^* &= p\left[v_1(\bar{x}_1^*, \bar{x}_2^*, \bar{\theta}) + v_A\left(\bar{x}_1^*, 0, \bar{\theta}\right) - R_1^*\right] + \\
&\quad +(1-p)\left[v_1(\underline{x}_1^*, \underline{x}_2^*, \underline{\theta}) + v_A\left(\underline{x}_1^*, 0, \underline{\theta}\right)\right] \\
U_2^* &= p[v_2(\bar{x}_1^*, \bar{x}_2^*, \bar{\theta}) + v_A\left(\bar{x}_1^*, \bar{x}_2^*, \bar{\theta}\right) - v_A\left(\overline{x}_1^*, 0, \overline{\theta}\right) - R_2^*] + \\
&\quad +(1-p)\left[v_2(\underline{x}_1^*, \underline{x}_2^*, \underline{\theta}) + v_A\left(\underline{x}_1^*, \underline{x}_2^*, \underline{\theta}\right) - v_A\left(\underline{x}_1^*, 0, \underline{\theta}\right)\right]
\end{aligned}$$

To derive explicit solutions for the equilibrium decisions, x_i^*, we introduce a trade example where we completely specify players' payoff functions and fully characterize the equilibrium contracts. This simple linear-quadratic model can easily cover standard screening settings, like regulation of a multinational firm in an international context, labor contracts offered by two employers, federal and local taxation of a firm.

Example: trade contracts without information disclosure on consumer's preferences.

Consider the case of two differentiated sellers (principals) that sequentially sign contracts with a common buyer (the agent) for the provision of two complementary products, x_i for $i = 1, 2$. The two contracts consist of a price $t_i(x_i)$ that the buyer has to pay for x_i units of output i[6].

Assume that the two principals have the same quadratic cost function, $C_i(x_i) = \frac{1}{2}x_i^2$. The common buyer derives an utility $\theta\left(x_1 + x_2\right) + x_1 x_2$ from the acquisition of x_1 units of product one and x_2 units of product two.

Hence, in this simple model the three players have respectively payoffs functions

$$U_1 = v_1(x_1, x_2, \theta) + t_1 = -\tfrac{1}{2}x_1^2 + t_1,$$

$$U_2 = v_2(x_1, x_2, \theta) + t_2 = -\tfrac{1}{2}x_2^2 + t_2,$$

$$U_A = v_A(x_1, x_2, \theta) - t_1 - t_2 = \theta\left(x_1 + x_2\right) + x_1 x_2 - t_1 - t_2.$$

The buyer has private information about his preferences: he is the only player who knows the exact value of θ. The two sellers simply know that the buyer's preferences may be characterized by a marginal value $\theta \in \left\{\overline{\theta}, \underline{\theta}\right\}$, with probability $\Pr(\overline{\theta}) = p = 1 - \Pr(\underline{\theta})$[7].

To make the analysis as simple as possible, let $\overline{\theta} = 1 + \frac{\Delta\theta}{2}$ and $\underline{\theta} = 1 - \frac{\Delta\theta}{2}$, with $\Delta\theta \equiv \bar{\theta} - \underline{\theta} \in [0, 2]$. In this simple model, $\Delta\theta$ represents the difference (in marginal value) that a high and a low type attach to each product.

From Proposition 3, the informational rent the follower must leave to the high type reduces to

$$R_2(\bar{x}_1, \underline{x}_1) = \left(\Delta\theta + \bar{x}_1 - \underline{x}_1\right)\underline{x}_2.$$

[6] Alternatively, x_i may well represent the quality of product i.

[7] Note that one could reinterpret the model assuming that the two sellers face a continuum of consumers with preferences $\theta\left(x_1 + x_2\right) + x_1 x_2$ and that a fraction p of consumers has marginal valuation $\overline{\theta}$, whereas the complementary fraction, $1 - p$, has valuation $\underline{\theta}$

It follows that for any contract π_1, the follower's optimal contract is characterized by the decisions[8]:

$$\begin{aligned}
\bar{x}_2 &= \arg\max_{x_2 \geq 0} -\tfrac{1}{2}x_2^2 + (\overline{\theta} + \overline{x}_1)x_2 = \overline{\theta} + \overline{x}_1, \\
\underline{x}_2 &= \arg\max_{x_2 \geq 0} (1-p)\left[-\tfrac{1}{2}x_2^2 + (\underline{\theta} + \underline{x}_1)x_2\right] - p\left(\Delta\theta + \bar{x}_1 - \underline{x}_1\right)x_2 \\
&= \max\left\{\underline{\theta} + \underline{x}_1 - \tfrac{p}{1-p}\left(\Delta\theta + \bar{x}_1 - \underline{x}_1\right), 0\right\}.
\end{aligned}$$

On her part, P_1 must leave to the high type a net informational rent equal to (see Proposition 4):

$$\begin{aligned}
R_1(\bar{x}_1, \underline{x}_1) &= R(\bar{x}_1, \underline{x}_1) - R_2(\bar{x}_1, \underline{x}_1) \\
&= \Delta\theta\left(\underline{x}_1 + \underline{x}_2\right) - \left(\Delta\theta + \bar{x}_1 - \underline{x}_1\right)\underline{x}_2 \\
&= \Delta\theta\underline{x}_1 - \left(\bar{x}_1 - \underline{x}_1\right)\underline{x}_2.
\end{aligned}$$

Notice that no matter what the high type did with P_1, he can always guarantee himself a rent with P_2 equal to $\Delta\theta\underline{x}_2$, simply because he has a higher marginal value for product two than a low type. Similarly, when contracting with P_1 the high type can always obtain $\Delta\theta\underline{x}_1$ by choosing the contract P_1 designed for the low type. If the two decisions were independent the high type would therefore obtain a total informational rent equal to $\overline{U}_A = \Delta\theta\left(\underline{x}_1 + \underline{x}_2\right)$.

When the two decisions are complements, this is not yet the end of the story. The interaction between A and P_1 transforms the marginal value of the agent for P_2's product from θ to $\theta + x_1$. If P_1's optimal contract is such that $\Delta x_1 = \bar{x}_1 - \underline{x}_1 > 0$, then the comparative advantage of the high type with respect to the low type is increased by the contractual relationship with P_1 and enables the high type to obtain from P_2 an extra payoff, in terms of price discount, equal to $\left(\bar{x}_1 - \underline{x}_1\right)\underline{x}_2$.

Nevertheless, it is not A, but P_1, who appropriates this extra surplus. In fact, the complementarity between the two products enables P_1 to increase the price $\bar{t}_1$ by $\left(\bar{x}_1 - \underline{x}_1\right)\underline{x}_2$ and make the agent pay for the better deal he obtains with P_2 when he purchases also from P_1.

The next proposition shows how the possibility to extract from the agent part of his informational rent vis à vis the second principal may induce P_1 to distort her mechanism. It turns out to exist two different parameters' regions. In the first, P_1 distorts the high type's allocation but gives to the low type the same allocation as under full information[9]. In the second, P_1 simply offers to A the standard single-principal screening contract in which only the decision for the low type is distorted.

Proposition 5 *In the absence of information disclosure, the (pure-strategy) equilibrium is characterized by the following contracts.*

[8] The program is concave so that FOCS are also sufficient.

[9] Under full information, the leader would sell $x_1^{FI} = \theta$ and the follower $x_2^{FI} = 2\theta$.

The leader's decisions are

$$\begin{cases} \bar{x}_1 = \bar{x}_1^{FI} + \frac{2(1-p)-\Delta\theta(1+2p)}{1+p} \gtrless \bar{x}_1^{FI} \\ \underline{x}_1 = \underline{x}_1^{FI} \end{cases} \quad \textit{in region } A,$$

$$\begin{cases} \bar{x}_1 = \bar{x}_1^{FI} \\ \underline{x}_1 = \max\left\{\underline{x}_1^{FI} - \frac{p}{1-p}\Delta\theta, 0\right\} < \underline{x}_1^{FI} \end{cases} \quad \textit{in region } B.$$

The follower's decisions are

$$\begin{cases} \bar{x}_2 = \bar{x}_2^{FI} + \frac{2(1-p)-\Delta\theta(1+2p)}{1+p} \gtrless \bar{x}_2^{FI} \\ \underline{x}_2 = \underline{x}_2^{FI} - \frac{p[2(1-p)+\Delta\theta]}{1-p^2} < \underline{x}_2^{FI} \end{cases} \quad \textit{in region } A,$$

$$\begin{cases} \bar{x}_2 = \bar{x}_2^{FI} \\ \underline{x}_2 = 0 < \underline{x}_2^{FI} \end{cases} \quad \textit{in region } B.$$

Region A is such that $0 \le \Delta\theta \le \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)}$.
Region B is such that $\frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)} \le \Delta\theta \le 2$.

Proof. See Appendix.

Figure 1 illustrates the optimal contracts of Proposition 5 as a function of $\Delta\theta$.

Put Figure 1 right here

A few comments are in order.

1. Distortions. Under asymmetric information there can be distortions both for $\bar{x}_1$ and $\underline{x}_1$; in this example, there can be either upward or downward distortions for the high type, depending on whether $0 \le \Delta\theta \le \frac{2(1-p)}{1+2p}$, or $\frac{2(1-p)}{1+2p} \le \Delta\theta \le \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)}$. The low type is never upward distorted. The presence of distortions for *both* types is to be contrasted with monopolistic screening models and simultaneous common agency games where all equilibria are characterized by the "*absence of distortions at the top*".

Let us briefly examine the comparison with these models.

- In the standard single-principal setting, P_1 faces a trade-off between efficiency and rent-extraction; this trade-off is best struck by distorting downward the decision for the low type and by offering the efficient outcome to the high type.

- In simultaneous common agency games with complementary decisions, there is a double downward distortion for the decision for the low type; principal i's reduction in $\underline{x}_i$ makes a reduction in $\underline{x}_j$ more desirable for principal j. For both principals it is pointless to distort the decision for the high type since this has no impact on rent extraction (see Martimort and Stole (1999a,b)).

- In dynamic common agency games, the two principals are not in a symmetric position. As already suggested in Proposition 3, the follower can make both the high and the low type pay for the extra surplus generated by the complementarity between the two products. On the other hand, under asymmetric information, the leader can affect the reaction function of the follower and make the agent pay for the extra informational rent (R_2) that P_2 must leave to the high type when his comparative advantage with respect to the low type is increased by the contractual relationship with P_1. Since the extra informational rent, $R_2(\overline{x}_1,\underline{x}_1)$, is a function of $\underline{x}_1$ and $\overline{x}_1$, depending on the elasticity of the reaction function of P_2's decisions with respect to $\underline{x}_1$ and $\overline{x}_1$, P_1's optimal contract may exhibit distortions (in both directions) for either decisions.

2. Shut-down. In the absence of any contractual relationship between A an P_1 (or, equivalently, in case of separable preferences) a single seller finds profitable to trade with both types if and only if $p\bar{\theta} \leq \underline{\theta}$, or, equivalently, if and only if $\Delta\theta \leq \frac{2(1-p)}{1+p}$. This condition is exogenous to the model and simply says that designing a mechanism that attracts both types is profitable if and only if the percentage of high types is high, or if the difference in marginal valuations is low. In a dynamic contracting environment, the "shut-down" condition for principal P_2 is endogenous and it depends on the contract that A signs with P_1. The interaction between A and the leader transforms the agent's marginal valuation for x_2 from θ to $\theta+x_1$. In equilibrium, the leader may have an interest in designing a mechanism that forces P_2 to exclude low-valuation agents from trade. This occurs if $\pi_1$is such that $\underline{\theta}+\underline{x}_1 < p(\overline{\theta}+\bar{x}_1)$. In this case, the leader's optimal mechanism is simply the standard single-principal screening contract with $\bar{x}_1 = \bar{\theta} = \bar{x}_1^{FI}$ and $\underline{x}_1 = \max\left\{\underline{x}_1^{FI} - \frac{p}{1-p}\Delta\theta, 0\right\}$. Conversely, if P_1 offers to A a contract for which $\underline{\theta} + \underline{x}_1 > p(\overline{\theta} + \bar{x}_1)$, then P_2 reacts by designing a mechanism such that $\underline{x}_2(\bar{x}_1, \underline{x}_1) > 0$; in this case P_1's optimal contract is characterized by $\bar{x}_1 = \bar{x}_1^{FI} + \frac{2(1-p)-\Delta\theta(1+2p)}{1+p}$ and $\underline{x}_1 = \underline{x}_1^{FI}$.

Proposition 5 shows that the optimal contract induces $\underline{x}_2(\bar{x}_1, \underline{x}_1) > 0$ if the difference between the two types is not too high (region A), and $\underline{x}_2(\bar{x}_1, \underline{x}_1) = 0$ if such a difference is significant (region B).

In the next section we show that the leader can improve upon this contract by sharing part of the information she learns from the agent with the follower. In order to minimize the informational rent she must leave to A, the leader benefits from manipulating P_2's beliefs by disclosing a signal that is correlated with the agent's private information.

4 Optimal disclosure policy

In this section we examine the possibility that P_1 commits to a disclosure policy δ which transmits information to the downstream principal.

At the beginning of the second stage, P_2 receives a signal $s \in S$ that is a realization of a random variable $\widetilde{s}$ with support S and distribution $\delta(s/\theta)$. Let $(\overline{\delta}(s), \underline{\delta}(s)) \in \triangle(S)^2$ represent respectively the two probability distributions over S for the high and the low type. As usual, s has no precise meaning: The signal may represent any information that is correlated with the agent's type.

At $t = 2$, after receiving signal s, P_2 updates her beliefs using Bayes' rule and δ. Let

$$\mu(\overline{\theta}_2^E/s) = \frac{\overline{\delta}(s)p}{\overline{\delta}(s)p + \underline{\delta}(s)(1-p)}$$

be the posterior probability of facing a high (extended) type $\overline{\theta}_2^E = (\bar{\theta}, \bar{x}_1, \bar{t}_1)$, conditional on observing signal s. The follower offers to A a contract $\pi_2(s) = (\bar{x}_2, \underline{x}_2, \bar{t}_2, \underline{t}_2)$ which is a solution of $\mathcal{P}_2$ with posterior beliefs $\mu(./s)$ instead of p.

At the first stage, P_1 anticipates the contract $\pi_2(s)$ and offers to A a contract $\pi_1 = (\underline{x}_1, \underline{t}_1, \bar{x}_1, \bar{t}_1)$ and a disclosure policy[10] $\delta = (\underline{\delta}(s), \overline{\delta}(s))$ which solve

$$\mathcal{P}_1 : \begin{cases} \underset{\pi_1}{Max}\ pE_{\overline{\delta}(s)}\left[v_1(\bar{x}_1, \bar{x}_2(s), \bar{\theta}) + \bar{t}_1\right] + (1-p)E_{\underline{\delta}(s)}\left[v_1(\underline{x}_1, \underline{x}_2(s), \underline{\theta}) + \underline{t}_1\right] & \\ \text{s.t.} & \\ \overline{U}_A = E_{\overline{\delta}(s)}\left\{v_A\left(\bar{x}_1, \bar{x}_2(s), \bar{\theta}\right) - \bar{t}_2(s)\right\} - \bar{t}_1 \geq 0 & (\overline{IR}_1) \\ \underline{U}_A = E_{\underline{\delta}(s)}\left\{v_A\left(\underline{x}_1, \underline{x}_2(s), \underline{\theta}\right) - \underline{t}_2(s)\right\} - \underline{t}_1 \geq 0 & (\underline{IR}_1) \\ \overline{U}_A \geq E_{\overline{\delta}(s)}\left\{v_A\left(\underline{x}_1, x_2(\bar{\theta}, \underline{x}_1/s), \bar{\theta}\right) - t_2(\bar{\theta}, \underline{x}_1/s))\right\} - \underline{t}_1 & (\overline{IC}_1) \\ \underline{U}_A \geq E_{\overline{\delta}(s)}\left\{v_A\left(\bar{x}_1, x_2(\underline{\theta}, \bar{x}_1/s)), \underline{\theta}\right) - t_2(\underline{\theta}, \bar{x}_1/s))\right\} - \bar{t}_1 & (\underline{IC}_1) \end{cases}$$

where for any $s \in S$,

$$(x_2(\bar{\theta}, \underline{x}_1/s), t_2(\bar{\theta}, \underline{x}_1/s)) \text{ and } (x_2(\underline{\theta}, \bar{x}_1/s), t_2(\underline{\theta}, \bar{x}_1/s))$$

represent the agent's allocation with P_2, when he misreports to P_1, which is conditional on the signal s.

The leader must provide incentives against any possible deviation path that the agent may follow in the continuation game. This explains the presence of the two incentive compatibility constraints in P_1's program. The first two individual rationality constraints ensure that the agent accepts π_1.

The optimal direct revelation mechanism for P_2, conditional on s, is simply the one we obtained in Proposition 3 with $\mu(./s)$ instead of p. In particular, only the decision for the low type is (downward) distorted and $\bar{x}_2^*(s)$ is constant over s, since it does not depend on the follower's beliefs. Using Lemma 2 the out-of-equilibrium continuation game is characterized by $x_2(\bar{\theta}, \underline{x}_1/s) = x_2(\underline{\theta}, \bar{x}_1/s) = \underline{x}_2(s)$, for any $s \in S$.

The following proposition characterizes P_1's reduced program.

[10] The disclosure policy, δ, is such that for any $\theta \in \Theta$:
(a) $\delta(s/\theta) \in [0, 1]$;
(b) $\sum_{s \in S} \delta(s/\theta) = 1$.

Proposition 6 *P_1 offers a contract, π_1^*, and a disclosure policy, δ^*, that solve*

$$\begin{cases} \underset{\pi_1,\delta}{Max}\ pE_{\overline{\delta}(s)}\left[v_1(\bar{x}_1,\bar{x}_2(s),\bar{\theta})+v_A\left(\bar{x}_1,0,\bar{\theta}\right)\right]+(1-p)E_{\underline{\delta}(s)}\left[v_1(\underline{x}_1,\underline{x}_2(s),\underline{\theta})+v_A\left(\underline{x}_1,0,\underline{\theta}\right)\right]+ \\ -p\left[E_{\underline{\delta}(s)}R(s)-E_{\overline{\delta}(s)}R_2(s)\right] \\ s.t. \\ \\ E_{\underline{\delta}(s)}R(s)\leq E_{\overline{\delta}(s)}\left[\Delta_\theta v_A(\bar{x}_1,\underline{x}_2(s))\right], \end{cases}$$

where $R(s)$ and $R_2(s)$ are as in Proposition 4.

Proof. See appendix.

A few comments on P_1's reduced program.

First, P_1's objective function can be decomposed in three terms. The first two represent the surplus generated by the interaction between A and P_1. The last term

$$ER_1(\underline{x}_1,\underline{\delta}(s),\bar{x}_1,\overline{\delta}(s))=\left[E_{\underline{\delta}(s)}R(s)-E_{\overline{\delta}(s)}R_2(s)\right]$$

is the expected informational rent that P_1 must leave to the efficient type in order to make him reveal his private information. From Lemma 2 we know that an efficient type that lies to $P_{1,}$ always receives $\underline{x}_2(s)$ with P_2. Hence, in order to create the correct incentive scheme, P_1 is obliged to reduce $\bar{t}_1$ by $E_{\underline{\delta}(s)}R(s)$, where

$$R(s)=\Delta_\theta v_A(\underline{x}_1,\underline{x}_2(s))=v_A\left(\underline{x}_1,\underline{x}_2(s),\bar{\theta}\right)-v_A\left(\underline{x}_1,\underline{x}_2(s),\underline{\theta}\right)\geq 0$$

is the efficiency gain of a high type with respect to a low type.

When P_1 discloses information to P_2, an efficient type who misreports at $t=1$, not only gets $\underline{x}_1$ instead of $\bar{x}_1$, but also affects the beliefs of the follower and hence of the expected rent $E_{\underline{\delta}(s)}R(s)$.

On the other hand, when contracting with a high type, P_1 realizes that the payoff the agent obtains with P_2 depends on the terms of the contract P_1 and A sign in the first stage. Clearly, for the low type the continuation game is irrelevant since any additional surplus from the interaction with P_2 will be extracted by the latter with $\underline{t}_2$.

Conversely, for the high type, the informational rent vis à vis P_2, conditional on signal s is

$$R_2(s)=\widehat{v}^A\left(\bar{x}_1,\underline{x}_2(s),\bar{\theta}\right)-\widehat{v}^A\left(\underline{x}_1,\underline{x}_2(s),\underline{\theta}\right)$$

where, as in Section 3,

$$\widehat{v}^A\left(\bar{x}_1,\underline{x}_2(s),\bar{\theta}\right)=v_A\left(\bar{x}_1,\underline{x}_2(s),\bar{\theta}\right)-v_A\left(\bar{x}_1,0,\bar{\theta}\right)$$

$$\widehat{v}^A\left(\underline{x}_1,\underline{x}_2(s),\underline{\theta}\right)=v_A\left(\underline{x}_1,\underline{x}_2(s),\underline{\theta}\right)-v_A\left(\underline{x}_1,0,\underline{\theta}\right)$$

represent the additional value of $\underline{x}_2(s)$ respectively for the high and the low type.

The expected rent for the high type vis à vis P_2, conditional on reporting truthfully to P_1, is thus $E_{\overline{\delta}(s)}R_2(s)$.

It follows that the leader must leave to the high type an expected rent equal to $ER_1 = E_{\underline{\delta}(s)}R(s) - E_{\overline{\delta}(s)}R_2(s)$.

A last comment on the constraint in P_1's reduced program. This represents the incentive constraint for the low type; as in dynamic single-principal frameworks, P_1 cannot reduce $\overline{t}_1$ too much if she does not want to induce the low type to mimic the high type. This constraint is never binding when P_1 does not signal to P_2; nevertheless, it might bind when some information is released and $E_{\underline{\delta}(s)}\underline{x}_2(s) \gg E_{\overline{\delta}(s)}\underline{x}_2(s)$.

The following definition identifies an important benchmark.

Definition 2 *The two decisions, x_1 and x_2, are unrelated if for any $x \in \Re^2_+$ and for any $\theta \in \Theta$:*

1) A's marginal value for x_2 does not depend on x_1, $\frac{\partial^2 v_A(x_1,x_2,\theta)}{\partial x_1 \partial x_2} = 0$ (contract externalities);

2) P_1 is not directly interested in x_2, $\frac{\partial v_1(x_1,x_2,\theta)}{\partial x_2} = 0$ (direct externalities).

Proposition 7 relates the role of externalities to information transmission between principals.

Proposition 7 *If the two decisions are unrelated, then P_1 never finds profitable to disclose information to P_2.*

Proof. See Appendix.

This result is very general and deserves a few comments. Even if P_1 can increase the agent's rent vis à vis P_2 when she discloses information, she can never appropriate this rent when $v_A(x_1, x_2, \theta)$ is separable in the two decisions. This is a consequence of P_1's own incentive problem.

To make the point clear, consider a situation in which prior beliefs are such that P_2, should she receive no information from P_1, she would set $\underline{x}_2 = 0$[11]. In this case, A receives no rent from his interaction with P_2. But then P_1 could design a signalling mechanism such that she sends a signal s_1 with probability $\overline{\delta}$ in case she observes a high type and with probability $\underline{\delta} = 1$ in case she observe a low type. Clearly, for any $s \neq s_1$, $\underline{x}_2(s) = 0$ since P_2 understands that she is facing a high type with certainty. However, for $\overline{\delta}$ low enough, signal s_1 becomes sufficiently informative of a low type and induces P_2 to leave some positive rent to a high type: the posterior probability of dealing with a high type after observing s_1 is sufficiently low and it becomes profitable to set $\underline{x}_2(s_1) > 0$. In this case, P_1 can induce P_2 to leave to A a strictly positive rent. Information transmission is profitable for P_1 since she can make the agent pay for the extra surplus with P_2 by increasing $\overline{t}_1$. Indeed, this would be the end of the story if there were no asymmetry of information between A and P_1.

[11] In the contract theory literature this situation is often referred to as the "shut-down case" and arises when $p > \widehat{p}$ where $\widehat{p}$ is a critical value that depends on all parameters of the model.

When also P_1 cannot tell the agent's type, she is obliged to provide A with incentives for truthtelling. The efficient type can in fact always pretend he is inefficient. Since in this case signal s_1 is sent with probability one, by mimicking the inefficient type, the efficient agent can obtain the same rent $R_2(s_1)$ with certainty rather that with probability $\overline{\delta}$. To avoid false reports, P_1 is therefore obliged to decrease $\overline{t}_1$ by $R_2(\underline{s_1})$. It follows that the net effect of information disclosure on $\overline{t}_1$ is thus $-(1-\overline{\delta(s_1)})R_2(s_1) < 0$ so that P_1 is strictly better off by keeping secret all the information she learns from A.

Proposition 7 generalizes this reasoning by showing that for any pair of probability distributions $\overline{\delta}(s), \underline{\delta}(s)$ such that P_1 (partially) informs P_2, the most favorable signals for A (signals such that $R_2(s)$ is high) are always more likely under $\underline{\delta}(s)$ than $\overline{\delta}(s)$ so that information disclosure exasperates P_1's incentives problem.

Proposition 7 is reminiscent of one in Baron and Besanko (1984) for a dynamic principal-agent framework. They show that under full commitment the optimal dynamic contract is the twofold repetition of the static contract. In other words, principal's *self one* should not inform *self two* about what she learns in the first stage. If a principal lacks of the commitment not to use such information to turn the contractual relationship to her own advantage, or to renegotiate the contract in the second period if the two parties both want to alter the initial agreement, then the optimal contract must be modified in order not to fully inform the principal about the agent's type (see, for example, Laffont and Tirole (1988, 1990)).

In our framework there are two different principals acting at $t = 1$ and $t = 2$. The agent's payoff vis à vis P_1 can differ from that vis à vis P_2 and the two principals do not need to have the same preferences. Nevertheless, when the agent's payoff is separable in x_1 and x_2 we obtain the same result.

Proposition 8 holds even if we allow P_1 to sell the information she obtains from A to P_2. Suppose, for example, that P_1 has the possibility to make P_2 pay for a better information structure. When the agent's payoff is separable in x_1 and x_2, one can show that P_1 is still worse off if she releases any information. The idea is always the same. When preferences are separable, P_1 can never appropriate the surplus generated by the finer information structure she gives to P_2 because of her incentives problem.

We can conclude that there must be externalities between the two contractual relationships to induce P_1 to disclose information to P_2. In the rest of this section we concentrate on a special class, *contract externalities,* that are typical of contracting with common agency. We acknowledge that also direct payoff-relevant externalities may be relevant in explaining why a principal may be interested in disclosing the information she learns from her agent. Nevertheless, by leaving these effects out of the analysis we succeed in isolating the rent seeking effect that characterizes the design of sequential contracts under asymmetric information.

Assume that $v_A(x_1, x_2, \theta)$ is such that the two decisions x_1 and x_2 are complements so that the marginal value of x_2 increases with the decision x_1. If P_1's mechanism is such that $\overline{x}_1 > \underline{x}_1$, then although it remains true that the efficient type can induce a higher expected decision with P_2 when he announces he is inefficient [as suggested in the discussion of Proposition 7, $E_{\underline{\delta}(s)}(\underline{x}_2(s)) \geq E_{\overline{\delta}(s)}(\underline{x}_2(s))$], this does not imply that he guarantees himself a higher expected rent since by mimicking the low type the efficient agent reduces x_1 and therefore the value of x_2. It follows that under strict complementarity the leader can indeed benefit from disclosing some information to P_2. In order to clarify this we reintroduce the simple linear-quadratic model used in Section 3.

Example: Information Disclosure on Consumer's Shopping Activity.

Let the common buyer have preferences $v_A(x_1, x_2, \theta) = \theta(x_1 + x_2) + x_1 x_2$. Think of the two principals as e-sellers who post websites in which they specify the price-schedule for their product or service, $t_i(x_i)$, and a disclosure policy $\delta(.)$. Both sellers have the same cost function $v_i(x_i) = -\frac{1}{2}x_i^2$.

In this setting the signal the second seller receives from the first seller can represent, for example, the agent's choice of the quality of P_1's product, the number of units that he bought, the price paid, or any other information that is correlated with the agent's shopping activity, like for example the path followed in the website.

P_1 must leave to the high-valuation buyer an informational rent equal to

$$\begin{aligned} ER_1(\underline{x}_1, \underline{\delta}(s), \bar{x}_1, \overline{\delta}(s)) &= E_{\underline{\delta}(s)}R(s) - E_{\overline{\delta}(s)}R_2(s) \\ &= \Delta\theta\underline{x}_1 + \Delta\theta E_{\underline{\delta}(s)}\underline{x}_2(s) - (\Delta\theta + \overline{x}_1 - \underline{x}_1)E_{\overline{\delta}(s)}\underline{x}_2(s). \end{aligned}$$

As shown in Proposition 6, the leader's optimal contract, π_1^* must solve[12]

$$\begin{cases} \underset{\pi_1}{Max}\ p\left(\bar{\theta}\bar{x}_1 - \frac{1}{2}\bar{x}_1^2\right) + (1-p)\left(\underline{\theta}\underline{x}_1 - \frac{1}{2}\underline{x}_1^2\right) - pER_1(\underline{x}_1, \underline{\delta}(s), \bar{x}_1, \overline{\delta}(s)) \\ \text{s.t.} \\ \bar{x}_1 - \underline{x}_1 \geq E_{\underline{\delta}(s)}\underline{x}_2(s) - E_{\overline{\delta}(s)}\underline{x}_2(s). \qquad (\underline{IC}_1) \end{cases}$$

This program would still be difficult to solve for any possible disclosure policy δ that P_1 may use to signal information to P_2. The following lemma proves that P_1 does not need to use more than two signals, say s_1 and s_2.

Lemma 3 *The optimal disclosure policy, δ^*, is such that $\overline{\delta}(s)^*$ and $\underline{\delta}(s)^*$ are two Bermoulli distributions. Furthermore, for one of the two signals P_2 "shuts down" the low type, i.e. $\underline{x}_2(s) = 0$.*

Proof. See Appendix.

The idea behind this result is simple. The leader wants to manipulate the follower's beliefs in order to increase the agent's rent with P_2 and in so doing

[12] The monotonicity constraint for x_1 is automatically implied by $(\underline{IC}_1)$ since $E_{\underline{\delta}(s)}\underline{x}_2(s) - E_{\overline{\delta}(s)}\underline{x}_2(s) \geq 0$, as suggested in the proof of Proposition 7.

reduce the net informational rent that she has to pay to the agent. For this aim, the leader decides to let P_2 have access to information that is correlated with the agent's preferences. Lemma 3 proves that the best way to increase the expected rent of the agent vis à vis the second principal is to give to P_2 an informational structure such that one of the two signals (say s_1) is really informative of the low type and induces P_2 to give away a high informational rent, i.e. a significant price discount. In this case, the other signal (say s_2) is necessarily informative of a high type and hence induces P_2 to set $\underline{x}_2(s_2) = 0$.

Let $\bar{\delta} = \Pr(s_1/\bar{\theta}) = 1 - \Pr(s_2/\bar{\theta})$ and $\underline{\delta} = \Pr(s_1/\underline{\theta}) = 1 - \Pr(s_2/\underline{\theta})$ be the probability of signal s_1 respectively for the high and the low-type contracts. The optimal disclosure policy is depicted in Figure 1.

$$\begin{array}{lcl} \underline{\theta}: & \xrightarrow{\underline{\delta}} & s_1 \\ & \searrow \nearrow & \\ \bar{\theta}: & \xrightarrow[1-\bar{\delta}]{} & s_2 \end{array} \qquad \text{(Figure 2)}$$

Furthermore, let (Π^C, δ) be the set of all mechanisms that satisfy the following constraints:

(a) $\underline{\theta} + \underline{x}_1 - \frac{p}{1-p}\frac{\bar{\delta}}{\underline{\delta}}(\Delta\theta + \bar{x}_1 - \underline{x}_1) \geq 0,$

(b) $\underline{\theta} + \underline{x}_1 - \frac{p(1-\bar{\delta})}{(1-p)\underline{\delta}}(\Delta\theta + \bar{x}_1 - \underline{x}_1) \leq 0,$

(c) $(\bar{\delta}, \underline{\delta}) \in [0,1]^2,$

(d) $\underline{x}_1 \geq 0,$

(e) $\bar{x}_1 - \underline{x}_1 \geq (\underline{\delta} - \bar{\delta})\, \underline{x}_2(s_1) \qquad (\underline{IC}_1)$

Using Lemma 3, the leader's optimal mechanism $\pi_1^*, \delta^* = \{\underline{x}_1^*, \bar{x}_1^*, \bar{\delta}^*, \underline{\delta}^*\}$ solves

$$\begin{aligned} \underset{(\pi_1,\delta)\in(\Pi^C,\delta)}{Max} \quad & p\left(\bar{\theta}\bar{x}_1 - \tfrac{1}{2}\bar{x}_1^2\right) + (1-p)\left(\underline{\theta}\underline{x}_1 - \tfrac{1}{2}\underline{x}_1^2\right) - p\Delta\theta\underline{x}_1 + \\ & -p\left[\Delta\theta\underline{\delta} - (\Delta\theta + \bar{x}_1 - \underline{x}_1)\bar{\delta}\right]\left[\underline{\theta} + \underline{x}_1 - \tfrac{p}{1-p}\tfrac{\bar{\delta}}{\underline{\delta}}(\Delta\theta + \bar{x}_1 - \underline{x}_1)\right]. \end{aligned}$$

Note that in general any pair $\bar{\delta}, \underline{\delta}$ with $\bar{\delta} = \underline{\delta}$ represents a contract in which P_1 does not signal any information to P_2. However, since we write the program such that for s_2 $\underline{x}_2(s_2) = 0$, then the no-signalling case is uniquely defined by probabilities $\bar{\delta}^* = \underline{\delta}^* = 1$.

Under full information the last two terms in P_1's objective do not exist. In the standard principal-agent screening model the rent P_1 leaves to the efficient agent is simply $\Delta\theta\underline{x}_1$; hence, the last term, $\left[\Delta\theta\underline{\delta} - (\Delta\theta + \bar{x}_1 - \underline{x}_1)\bar{\delta}\right]\underline{x}_2(s_1)$, represents the additional rent introduced by common agency.

The expected rent the leader leaves to the agent, ER_1, can be rewritten as

$$ER_1 = \Delta\theta\underline{x}_1 - \bar{\delta}(\Delta_\theta x_1)\underline{x}_2(s_1) - \Delta\theta\underline{x}_2(s_1)\Delta_\theta\delta,$$

where $\Delta_\theta x_1 = \overline{x}_1 - \underline{x}_1 > 0$ and $\Delta_\theta \delta = \overline{\delta} - \underline{\delta} < 0$[13].

The second term in ER_1 represents the rent P_1 is able to extract from P_2 via the agent. Given that $\underline{x}_2(s_2) = 0$, this rent exists only with signal s_1 and it is then weighted with probability $\overline{\delta}$.

The third term is the negative incentives-effect generated by information transmission; by choosing the contract for the low type, the high type can obtain a rent with P_2, $\Delta\theta \underline{x}_2(s_1)$, with probability $\underline{\delta}$ instead of $\overline{\delta}$. In order to prevent mimicking the high type must therefore increase ER_1 by $\Delta\theta \underline{x}_2(s_1)\Delta_\theta \delta$. It is interesting to note that this effect is larger the more the information transmitted to P_2 (i.e. the larger is $\Delta_\theta \delta$).

When P_1 transmits information to P_2, $\underline{x}_2(s_1) > \underline{x}_2(p)$, where $\underline{x}_2(p)$ stands for the equilibrium choice of x_2 in case P_2 does not receive any information from P_1. This policy has a trade-off. On one hand, disclosing information results in a reduction of ER_1 due to the increase of $\underline{x}_2(s_1)$; on the other hand, disclosing information increases ER_1 because of the extra informational rent P_1 must leave to A for incentives-reasons. If the first effect dominates the second, then releasing some information becomes profitable.

With the following proposition we show that this may well be the case and we characterize a parameters' region such that the solutions of the previous program involves *information transmission* between the two principals.

Proposition 8 *In the parameter region C, the leader offers to the agent a mechanism, π_1^*, δ^*, that involves information transmission to the follower.*

π_1^*, is such that

$$\bar{x}_1^* = \bar{x}_1^{FI}, \\ \underline{x}_1^* = \underline{x}_1^{FI} + \frac{2(1-p) - \Delta\theta(1+2p)}{1-p} \leq \underline{x}_1^{FI}.$$

The optimal disclosure policy is

$$\bar{\delta}^* = \Pr(s_1/\overline{\theta}) = \frac{(1-p)[\Delta\theta - 2(1-p)]}{p[2(1-p) - 3\Delta\theta]}, \\ \underline{\delta}^* = \Pr(s_1/\underline{\theta}) = 1.$$

Conditional on receiving signal s, the follower offers to the agent a contract $\pi_2^*(s)$ such that:

$$\overline{x}_2^*(s) = \overline{x}_2^{FI} \text{ for any } s, \\ \underline{x}_2^*(s_1) = \underline{x}_2^{FI} - \frac{2p\Delta\theta}{1-p} < \underline{x}_2^{FI}, \\ \underline{x}_2^*(s_2) = 0.$$

Region C is defined by:

(i) $\frac{2(1-p)}{1+2p} \leq \Delta\theta \leq \frac{2(1-p)\left[\sqrt{1+p} - p\right]}{1+p(1-p)}$,

(ii) $\frac{-\left(9p+5p^2+1\right)\Delta\theta^2 + 2\left(6p - 7p^2 - p^3 + 2\right)\Delta\theta - 4 + 4p + 4p^2 - 4p^3}{p(1-p)[2(1-p) - 3\Delta\theta]} \geq 0$

[13] This derives from the fact that $\underline{x}_2(s_2) = 0$, as suggested in Lemma 3.

Proof. See Appendix.

The parameter region C simply ensures the existence of an interior solution in which all constraints (a) to (e) are satisfied[14].

Information transmission enables the leader to extract more rent from A by increasing the informational rent the follower leaves to the agent.

In this trade example, the first seller provides another seller with information that is directly correlated with the consumer's preferences to extract more surplus from the agent; indeed the final price the customer pays for P_1's product incorporates not only the value of the product per se, but also the gains that the agent obtains in future contractual relationships with other sellers.

Note that P_1 is strictly better off by disclosing information on the agent's type even if she does not make P_2 pay for the signal the latter receives. In case P_1 has the possibility to "sell" information on consumer's preferences to other sellers, we expect to observe in equilibrium a higher information flow (i.e. a lower $\bar{\delta}^*$).

Implementation.

The optimal disclosure policy δ^* has a simple implementation. Let s_1 correspond to a "secret" contract and s_2 to a "public" contract. When the agent picks up option $\underline{x}_1^*$, P_1 does not disclose the agent's choice. Conversely, when A selects option $\overline{x}_1^*$, then the principal discloses the terms of the contract with probability $1 - \bar{\delta}^*$. Alternatively, P_1 could offer to the agent the choice between three options. The first one, $(\underline{x}_1^*, \underline{t}_1^*)$ is never disclosed and it is chosen by unsophisticated consumers (low type). The second, $(\overline{x}_1^*, \overline{t}_1^*(s_1))$ is selected by sophisticated customers and is also secret. The third, $(\overline{x}_1^*, \overline{t}_1^*(s_2))$ is made public, in the sense that P_1 notifies to P_2 that A selected this option. By playing with the transfers $\overline{t}_1^*(s_1)$ and $\overline{t}_1^*(s_2)$ the seller can make the agent indifferent between $(\overline{x}_1^*, \overline{t}_1^*(s_1))$ and $(\overline{x}_1^*, \overline{t}_1^*(s_2))$, so that the latter randomizes with probability $\bar{\delta}^*$. As suggested in Proposition 8, letting P_2 observe the realizations of π_1 with a positive probability reduces, on average, the distortions that are due to the asymmetry of information and induces P_2 to increase the price discount that she offers to her customers. Clearly, those consumers who accept to disclose the terms of their contracts do not receive future rents; it follows that P_1 must compensate them through a discount on t_1.

It is important to point out that in our model it is the leader who offers to disclose the information on the agent's type. This decision, although motivated by profit-maximization, has a direct impact on welfare and consumer's surplus, as suggested in the following proposition.

Proposition 9 *Strategic information transmission between two sellers of complementary products or services increases welfare and consumer surplus.*

[14] It is possible to verify that region C is not empty (one can take for example $p = .5$ and explicitly solve the double inequality).

Proof. See Appendix.

Clearly, if P_2 receives a finer information structure, then she reduces the distortions of her contract. Since information transmission enables P_1 to reduce the net informational rent she leaves to the agent, the trade-off between efficiency and rent-extraction in her optimal mechanism also moves in favor of efficiency. This results in an increase of welfare. If information disclosure occurs with the customer's approval, it also increases consumer surplus, as suggested in many privacy policy webpages. Not surprising, as we argue here, it also favors those vendors who organize it.

Commitment.

The analysis of Sections 3 and 4 has been performed assuming full commitment. This means that we ruled out the possibility that P_1 discloses more information than actually allowed by the contract π_1. When P_1 is not directly interested in x_2, nor can she sell the information to P_2, as we assumed at the end of Section 4, then she has no reason to deviate from the disclosure policy $\delta(.)$. Conversely, if P_1 can collude with P_2 and share the extra surplus that the latter derives from a better information structure, then cheating on $\delta(.)$ becomes attractive. In this case, the optimal collusion-proof contract must be such that P_1 herself does not fully learn from the agent's choice. This requires the agent to randomize over (x_1, t_1), as suggested in the literature on the ratchet effect (see, for example, Freixas, Guesnerie and Tirole (1985), Hart and Tirole (1988), Laffont and Tirole (1988), Malcomson and Spynnewyn (1988), among others).

We also implicitly assumed that P_1 does not collude with the agent. As suggested in Caillaud, Jullien and Picard (1995), P_1 could publicly announce π_1 and δ and then sign a secret side contract with A. For example, once she has manipulated P_2's beliefs, P_1 could offer to A an agreement on the basis of which she sends only the most favorable signal, s_1. Equivalently, she could modify x_1 to extract further surplus from the agent. If the commitment assumption is removed, then we are back to a contracting game in which π_1, δ must be a best response to π_2. In such a case the optimal contracts would be the ones derived in Propositions 3 and 4 (with x_2 rather than $x_2(\pi_1)$). When P_1 lacks of any commitment not to privately renegotiate π_1, δ with A, information transmission to P_2 does not occur in equilibrium, since it cannot be credible (the coalition between A and P_1 can always improve upon δ by selecting only payoff-maximizing signals). The commitment towards π_1, δ seems reasonable when the interaction between P_1 and A is not too personal, as in case of a seller that screens many consumers with a nonlinear price schedule. It is much harder to defend in case of a personalized one-shot relationship. Nevertheless, if P_1 is involved in a long-term interaction with P_2, like in case of two differentiated sellers, then private renegotiation with the agent is also associated to a reputational cost.

5 Conclusions

This paper has examined the dynamic interaction between two principals that sequentially design their contractual relationships with a common agent.

We have shown that the optimal contracts can be characterized by an endogenous information flow between the two principals.

When receiving information that is correlated with the agent's type and his past contractual activity, a downstream principal may be induced to leave out a higher (expected) informational rent to the agent. For example, in case of price-discrimination, providing a downstream seller with information that is correlated with consumers's preferences may well result in a lower distortion on the price-schedule which eventually favors consumer's surplus.

In our model the disclosure of information is organized by an upstream principal, for example a seller that previously traded with the same agent. This principal has her own interest in favoring the agent with the second principal since she can appropriate the extra surplus that is generated by the disclosure of information. Although the leader is concerned only about her own payoff, in releasing information she also increases welfare by reducing the overall distortions that are due to the asymmetry of information.

The dynamic common agency model can accommodate direct externalities between the two mechanism designers. If the upstream principal is also directly affected by the decision of the downstream principal, there might be other reasons for information transmission that add to the rent-extraction one we proposed in this paper. For example, consider the case of a jointly-financed project. Two large investors that provide funding to a common borrower usually design contracts that (at least partially) release to each other the information that is obtained from the common borrower. The latter usually has better information than external investors on the characteristics of the project like, for example, the probability of success or its final cost. Under adverse selection, information sharing between investors is usually motivated either by a rent-extraction behavior, or by the need of coordinating the two investments. In this paper we provide a rationale based on the first effect. The possibility that principals communicate the agent's private information in order to better coordinate their policies also represents an interesting area for future research.

References

[1] Baron, D., 1985, "Non-Cooperative Regulation of a Non-Localized Externality", *Rand Journal of Economics*, 16, 553-568.

[2] Bagwell, K., 1995, "Commitment and Observability in Games", *Games and Economic Behavior*, 8, 271-280.

[3] Baron, D. and D. Besanko, 1984, "Commitment and Fairness in a Continuing Relationship", *Review of Economic Studies*, 413-436.

[4] Bernheim, B.D. and M.D. Whinston, 1986, "Common Agency", *Econometrica*, 54, 923-942.

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

[6] Caillaud, B., B. Jullien and P. Picard, 1995, "Competing Vertical Structures: Precommitment and Renegotiation", *Econometrica*, 63, 3, 621-646.

[7] Calzolari, G. and Pavan, A., 2000a, "Monopoly with Resale", mimeo, University of Toulouse.

[8] Calzolari, G. and Pavan, A., 2000b, "A Markovian Revelation Principle for Sequential Common Agency", mimeo, University of Toulouse.

[9] Dixit, A., G. Grossman and E. Helpman, 1997, "Common Agency and Coordination: General Theory and Application to Government Policy Making", *Journal of Political Economy*, 105, 752-769.

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

[11] Freixas, X., R. Guesnerie and J. Tirole, 1985, "Planning under Incomplete Information and the Ratchet Effect", *Review of Economic Studies*, 52, 173–192.

[12] Gal-Or, E., 1985, "First and Second Mover Advantages", *International Economic Review*, 26, 649-653.

[13] Grossman, S. and O. Hart, 1983, "An Analysis of the Principal-Agent Problem", *Econometrica*, 51, 7-45.

[14] Hart, O. and J. Tirole, 1988, "Contract renegotiation and Coasian Dynamics", *Review of Economic Studies*, 55, 509-540.

[15] Laffont, J.J. and J. Tirole, 1987, "The Dynamics of Incentives Contracts", *Econometrica*, 56, 1153–1175.

[16] Laffont, J.J. and J. Tirole, 1988, "Comparative Statics of the Optimal Dynamic Incentives Contract", *European Economic Review*, 31, 901-926.

[17] Laffont, J.J. and J. Tirole, 1990, "Adverse Selection and Renegotiation in Procurement", *Review of Economic Studies*, 57, 597-625.

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

[19] Martimort, D., 1992, "Multi-Principaux avec Anti Selection", *Annales d'Economie et Statistique*, 28, 1-38;

[20] Martimort, D. 1996, "Exclusive Dealing, Common Agency and Multiprincipals Incentive Theory", *Rand Journal of Economics*, 27, 1-31.

[21] Martimort, D., 1999, "Renegotiation Design with Multiple Regulators", *Journal of Economic Theory*, 88, 261-293.

[22] Martimort, D. and L. Stole, 1999a, "The Revelation and Taxation Principles in Common Agency Games", mimeo University of Chicago and IDEI.

[23] Martimort, D. and L. Stole, 1999b, "Contractual Externalities and Common Agency Equilibria", mimeo University of Chicago and IDEI.

[24] Malcomson, J. and F. Spynnewyn, 1988, "The Multiperiod Principal-Agent Problem", *Review of Economic Studies*, 55, 391-407.

[25] Mezzetti, C. ,1997, "Common Agency with Horizontally Differentiated Principals", *Rand Journal of Economics*, 254-275.

[26] Padilla, A. and M. Pagano, 1997, "Endogenous Communication Among Lenders and Entrepreneurial Incentives", *Review of Financial Studies*, 10, 205-236.

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

[28] Prat. A. and A. Rustichini, 1998, "Sequential Common Agency", Tilburg CentER for Economic Research Discussion Paper 9895.

[29] Raith, M., 1996, "A General Model of Information Sharing in Oligopoly", *Journal of Economic Theory*, 71, 269-288.

[30] Stole, L. A. 1991, "Mechanism Design Under Common Agency", mimeo, University of Chicago.

Appendix: Proofs. Proof of Lemma 1.

- First, assume that a (pure-strategy) equilibrium exists in which $\bar{x}_1 < \underline{x}_1$ and $\bar{x}_2 \geq \underline{x}_2$. We show that $\mathcal{P}_1$ cannot be incentive compatible since the agent benefits from lying with P_1 and then truthfully reporting to P_2. Adding constraints ($\overline{IC}_1$) and ($\underline{IC}_1$) one gets

$$v_A(\bar{x}_1, \bar{x}_2, \bar{\theta}) - v_A(\underline{x}_1, \bar{x}_2, \bar{\theta}) \geq v_A(\bar{x}_1, \underline{x}_2, \underline{\theta}) - v_A(\underline{x}_1, \underline{x}_2, \underline{\theta}).$$

  This condition is never verified for $\bar{x}_1 < \underline{x}_1$ and $\bar{x}_2 \geq \underline{x}_2$ because of Assumptions A3 and A5.

- Second, assume that $\bar{x}_i < \underline{x}_i$ for both $i = 1, 2$.

  Considers the case in which A lies to both principals. Adding constraints ($\overline{IC}_1$) and ($\underline{IC}_1$) one gets

$$v_A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2, \bar{\theta}\right) \geq v_A\left(\bar{x}_1, \bar{x}_2, \underline{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right),$$

  which is never compatible with $\bar{x}_i < \underline{x}_i$ for both $i = 1, 2$ because of Assumption A3.

- Finally, consider the case in which $\bar{x}_2 < \underline{x}_2$ and $\bar{x}_1 \geq \underline{x}_1$.

  The sum of ($\overline{IC}_2$) and ($\underline{IC}_2$) requires that

$$v_A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) + v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) \geq v_A\left(\bar{x}_1, \underline{x}_2, \bar{\theta}\right) + v_A\left(\underline{x}_1, \bar{x}_2, \underline{\theta}\right).$$

  This condition is never satisfied for $\bar{x}_2 < \underline{x}_2$ and $\bar{x}_1 \geq \underline{x}_1$ because of Assumptions A3 and A5.

It follows that any pure-strategy equilibrium, $(\underline{x}_1, \overline{x}_1, \underline{x}_2, \overline{x}_2)$ of the game must have $\bar{x}_1 \geq \underline{x}_1$ and $\bar{x}_2 \geq \underline{x}_2$. ■

Proof of Proposition 3.

This is the standard principal-agent screening problem applied to a common-agency framework. The solution is in three steps.

- First, we show that if constraints $(\overline{IC}_2)$ and $(\underline{IR}_2)$ are satisfied so is $(\overline{IR}_2)$.

  Using

  $$\begin{aligned}\overline{U}_A &\equiv v_A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) - \bar{t}_1 - \bar{t}_2 \\ \underline{U}_A &\equiv v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) - \underline{t}_1 - \underline{t}_2,\end{aligned}$$

  constraint $(\overline{IC}_2)$ can be rewritten as

  $$\overline{U}_A \geq \underline{U}_A - v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) + \underline{t}_1 + v_A\left(\bar{x}_1, \underline{x}_2, \bar{\theta}\right) - \bar{t}_1.$$

  From $(\underline{IR}_2)$ we have that $\underline{U}_A \geq v_A\left(\underline{x}_1, 0, \underline{\theta}\right) - \underline{t}_1$. If $(\overline{IC}_2)$ and $(\underline{IR}_2)$ are satisfied, then

  $$\overline{U}_A \geq v_A\left(\underline{x}_1, 0, \underline{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) + v_A\left(\bar{x}_1, \underline{x}_2, \bar{\theta}\right) - \bar{t}_1 \geq v_A\left(\bar{x}_1, 0, \bar{\theta}\right) - \bar{t}_1$$

  for Assumptions A3 and A5.

- At the solution $(\underline{IR}_2)$ and $(\overline{IC}_2)$ bind and $(\underline{IC}_2)$ is slack.

  Using the expressions for $\overline{U}_A$ and $\underline{U}_A$, P_2's objective function can be rewritten as

  $$\begin{aligned}U_2 = \; & p\left[v_2(\bar{x}_1, \bar{x}_2, \bar{\theta}) - \overline{U}_A + v_A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) - \bar{t}_1\right] + \\ & +(1-p)\left[v_2(\underline{x}_1, \underline{x}_2, \underline{\theta}) - \underline{U}_A + v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) - \underline{t}_1\right].\end{aligned}$$

  In a similar way, constraint $(\underline{IC}_2)$ can be reduced to

  $$\overline{U}_A \leq \underline{U}_A + u^A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) - u^A\left(\underline{x}_1, \bar{x}_2, \underline{\theta}\right) + \underline{t}_1 - \bar{t}_1,$$

  which together with $(\overline{IC}_2)$ gives

  $$\begin{aligned}&\underline{U}_A + v_A\left(\bar{x}_1, \underline{x}_2, \bar{\theta}\right) - v_A\left(\underline{x}_1, \underline{x}_2, \underline{\theta}\right) + \underline{t}_1 - \bar{t}_1 \leq \overline{U}_A \\ &\overline{U}_A \leq \underline{U}_A + v_A\left(\bar{x}_1, \bar{x}_2, \bar{\theta}\right) - v_A\left(\underline{x}_1, \bar{x}_2, \underline{\theta}\right) + \underline{t}_1 - \bar{t}_1.\end{aligned}$$

  The r.h.s. term of this double inequality is (weakly) larger than the l.h.s. by monotonicity in decisions (Lemma 1). Since U_2 is decreasing in both rents $\overline{U}_A$ and $\underline{U}_A$, then for any $\underline{x}_2$ and $\overline{x}_2$ it is always optimal to make $(\underline{IR}_2)$ bind and to reduce $\overline{U}_A$ at its lower bound, which requires setting $(\overline{IC}_2)$ binding.

- P_2's program reduces to

$$\begin{cases} \underset{\{\underline{x}_2, \bar{x}_2\}}{Max} & p[v_2(\bar{x}_1, \bar{x}_2, \bar{\theta}) + v_A(\bar{x}_1, \bar{x}_2, \bar{\theta}) - v_A\left(\overline{x}_1, 0, \overline{\theta}\right)] - pR_2(\underline{x}_2) + \\ & +(1-p)\left[v_2(\underline{x}_1, \underline{x}_2, \underline{\theta}) + v_A(\underline{x}_1, \underline{x}_2, \underline{\theta}) - v_A(\underline{x}_1, 0, \underline{\theta})\right] \\ \text{s.t.} & R_2(\underline{x}_2) = v_A(\bar{x}_1, \underline{x}_2, \bar{\theta}) - v_A\left(\overline{x}_1, 0, \overline{\theta}\right) - \left[v_A(\underline{x}_1, \underline{x}_2, \underline{\theta}) - v_A(\underline{x}_1, 0, \underline{\theta})\right], \\ & \bar{x}_2 \geq \underline{x}_2 \geq 0. \end{cases}$$

The solution to this reduced program gives Proposition 3. ■

Proof of Lemma 2.
First, take a high type that misreports to P_1. We have to prove that

$$v_A(\underline{x}_1, \underline{x}_2, \bar{\theta}) - \underline{t}_2 \geq v_A(\underline{x}_1, \bar{x}_2, \bar{\theta}) - \bar{t}_2$$

Furthermore, since participation with P_2 is voluntary it must be that

$$v_A(\underline{x}_1, \underline{x}_2, \bar{\theta}) - \underline{t}_2 \geq v_A(\underline{x}_1, 0, \bar{\theta}).$$

The transfers $\underline{t}_2$ and $\bar{t}_2$ can be recovered from the proof of Proposition 3. Substituting $\underline{t}_2$ and $\bar{t}_2$ and using the monotonicity of x_2 one can show that the two inequalities are always satisfied under A3 and A5.

Similarly, take a low type who lies to P_1. We have to show that

$$v_A(\bar{x}_1, \underline{x}_2, \underline{\theta}) - \underline{t}_2 \geq v_A(\bar{x}_1, \bar{x}_2, \underline{\theta}) - \bar{t}_2$$

and

$$v_A(\bar{x}_1, \underline{x}_2, \underline{\theta}) - \underline{t}_2 \geq v_A(\bar{x}_1, 0, \underline{\theta}).$$

Using $\underline{t}_2$, $\bar{t}_2$ the first inequality reduces to

$$v_A(\bar{x}_1, \bar{x}_2, \bar{\theta}) - v_A(\bar{x}_1, \underline{x}_2, \bar{\theta}) \geq v_A(\bar{x}_1, \bar{x}_2, \underline{\theta}) - v_A(\bar{x}_1, \underline{x}_2, \underline{\theta})$$

which is always true for $\bar{x}_2 \geq \underline{x}_2$ and A3. The second inequality is always true under A5. We can conclude that an agent who lies to P_1 always receives allocation $(\underline{x}_2, \underline{t}_2)$ in the out-of-equilibrium continuation game with P_2. ■

Proof of Proposition 4.
To simplify notation we drop the dependence of P_2's optimal mechanism π_2^* on π_1 when not explicitly needed.

- If constraints $(\overline{IC}_1)$ and $(\underline{IR}_1)$ are satisfied, so is $(\overline{IR}_1)$. Rewrite constraint $(\overline{IC}_1)$ to get

$$\overline{U}_A \geq \underline{U}_A + v_A(\underline{x}_1, \underline{x}_2, \bar{\theta}) - v_A(\underline{x}_1, \underline{x}_2, \underline{\theta}).$$

Clearly, if $(\underline{IR}_1)$ is verified so is $(\overline{IR}_1)$ for Assumption A4.

- We are left with $(\overline{IC}_1)$, $(\underline{IR}_1)$ and $(\underline{IC}_1)$. We show that only the first two bind. With standard substitutions for $\underline{t}_2$ and $\overline{t}_2$, $(\overline{IC}_1)$ and $(\underline{IC}_1)$can be reduced to the following double inequality

  $$\underline{U}_A+v_A(\underline{x}_1,\underline{x}_2,\bar{\theta})-v_A(\underline{x}_1,\underline{x}_2,\underline{\theta}) \leq \overline{U}_A \leq \underline{U}_A+v_A(\bar{x}_1,\underline{x}_2,\bar{\theta})-v_A(\bar{x}_1,\underline{x}_2,\underline{\theta}),$$

  where the l.h.s. term is smaller than the r.h.s. term for Assumption A3. Since P_1's objective function is decreasing in both $\overline{U}_A$ and $\underline{U}_A$, it is optimal to make both $(\overline{IC}_1)$ and $(\underline{IR}_1)$ bind. In this case, $(\underline{IC}_1)$ is slack and can be neglected.

- Substituting $(\overline{IC}_1)$ and $(\underline{IR}_1)$ into P_1's objective we obtain that π_1^* is the solution of the reduced program

  $$\begin{cases} \underset{\{\underline{x}_1,\bar{x}_1\}}{Max}\ p\Big[\ v_1(\bar{x}_1,\overline{x}_2(\pi_1),\bar{\theta})+v_A(\bar{x}_1,0,\bar{\theta})\ \Big]-pR_1(\bar{x}_1,\underline{x}_1) \\ \qquad +(1-p)\left[v_1(\underline{x}_1,\underline{x}_2(\pi_1),\underline{\theta})+v_A(\underline{x}_1,0,\underline{\theta})\right] \\ \text{s.t.} \\ \bar{x}_1 \geq \underline{x}_1 \geq 0. \end{cases}$$

  ■

Proof of Proposition 5.

As suggested in Proposition 4, the leader's optimal contract is the solution of

$$\begin{cases} \underset{\{\underline{x}_1,\bar{x}_1\}}{Max}\ p\left(\bar{\theta}\bar{x}_1-\frac{1}{2}\bar{x}_1^2\right)+(1-p)\left(\underline{\theta}\underline{x}_1-\frac{1}{2}\underline{x}_1^2\right)-p[\Delta\theta\underline{x}_1-(\bar{x}_1-\underline{x}_1)\underline{x}_2(\bar{x}_1,\underline{x}_1)] \\ \text{s.t.} \\ \text{(a)}:\ \underline{x}_2(\bar{x}_1,\underline{x}_1)=\max\left\{\underline{\theta}+\underline{x}_1-\frac{p}{1-p}\left(\Delta\theta+\bar{x}_1-\underline{x}_1\right),0\right\}, \\ \text{(b)}:\ \bar{x}_1 \geq \underline{x}_1 \geq 0. \end{cases}$$

- We solve this program by comparing the maximal payoff P_1 can achieve by designing contracts $\{\underline{x}_1,\bar{x}_1\}$ such that $\underline{x}_2(\bar{x}_1,\underline{x}_1)=0$ with the maximal payoff P_1 can obtain with mechanisms that induce $\underline{x}_2(\bar{x}_1,\underline{x}_1)>0$. To this aim, we deliberately neglect all constraints and verify them ex-post.

- For any contract, $\{\underline{x}_1,\bar{x}_1\}$, such that $\underline{x}_2(\bar{x}_1,\underline{x}_1)=0$, P_1's program reduces to the standard single-principal screening problem whose solution is

  $$\begin{aligned} \bar{x}_1 &= \bar{\theta}, \\ \underline{x}_1 &= \max\left\{\underline{\theta}-\tfrac{p}{1-p}\Delta\theta,0\right\}. \end{aligned}$$

  Replacing $\bar{x}_1$ and $\underline{x}$, into the objective function and using $\bar{\theta}=1+\frac{\Delta\theta}{2}$ and $\underline{\theta}=1-\frac{\Delta\theta}{2}$, we obtain that the maximal payoff for any contract that induces $\underline{x}_2(\bar{x}_1,\underline{x}_1)=0$ is

  $$U_1=\begin{cases} \frac{(3p+1)\Delta\theta^2-4(1-p)\Delta\theta+4(1-p)}{8(1-p)} & \text{if}\quad 0\leq\Delta\theta\leq\frac{2(1-p)}{1+p}, \\ \frac{p}{2}(1+\frac{\Delta\theta}{2})^2 & \text{if}\quad \frac{2(1-p)}{1+p}\leq\Delta\theta\leq 2. \end{cases}$$

- If, instead, P_1 selects a mechanism such that $\underline{x}_2(\bar{x}_1, \underline{x}_1) > 0$, then $\underline{x}_2(\bar{x}_1, \underline{x}_1) = \underline{\theta} + \underline{x}_1 - \frac{p}{1-p}(\Delta\theta + \bar{x}_1 - \underline{x}_1)$.
  In this case, the contract that maximizes P_1's objective is recovered from the system of FOCS[15]

$$\begin{cases} p(\bar{\theta} - \bar{x}_1) - p\left[-\underline{x}_2(\bar{x}_1, \underline{x}_1) - (\bar{x}_1 - \underline{x}_1)\frac{\partial \underline{x}_2(\bar{x}_1, \underline{x}_1)}{\partial \bar{x}_1}\right] = 0 \\ (1-p)(\underline{\theta} - \underline{x}_1) - p\left[\Delta\theta + \underline{x}_2(\bar{x}_1, \underline{x}_1) - (\bar{x}_1 - \underline{x}_1)\frac{\partial \underline{x}_2(\bar{x}_1, \underline{x}_1)}{\partial \underline{x}_1}\right] = 0 \end{cases}$$

  Substituting the reaction function $\underline{x}_2(\bar{x}_1, \underline{x}_1)$ and its derivatives into the system we get:

$$\bar{x}_1 = \bar{x}_1^{FI} + \frac{2(1-p) - \Delta\theta(1+2p)}{1+p},$$
$$\underline{x}_1 = \underline{x}_1^{FI}.$$

  The payoff associated to this contract is

$$U_1 = \frac{(4p^3 - p^2 + 1)\Delta\theta^2 + 4(4p^3 - 3p^2 - 1)\Delta\theta + 4(4p^3 - 9p^2 + 4p + 1)}{8(1-p^2)}.$$

- The optimal mechanism is obtained by comparing the value functions associated to these two focal contracts.

- i) For $0 \leq \Delta\theta \leq \frac{2(1-p)}{1+p}$, then

$$\frac{(4p^3-p^2+1)\Delta\theta^2+4(4p^3-3p^2-1)\Delta\theta+4(4p^3-9p^2+4p+1)}{8(1-p^2)} \geq \frac{(3p+1)\Delta\theta^2-4(1-p)\Delta\theta+4(1-p)}{8(1-p)}$$

  if and only if

$$\Delta\theta^2\left[1 + p(1-p)\right] + 4p(1-p)\Delta\theta - 4(1-p)^2 \leq 0.$$

  The left hand side of this condition identifies a second order polynomial with roots $\Delta\theta^+ = \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)} \geq 0$ and $\Delta\theta^- = \frac{-2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)} < 0$.
  It follows that the optimal contract is such that $\underline{x}_2(\bar{x}_1, \underline{x}_1) > 0$ if

$$0 \leq \Delta\theta \leq \frac{2(1-p)\left[\sqrt{1+p} - p\right]}{1 + p(1-p)} \leq \frac{2(1-p)}{1+p},$$

  which corresponds to region A.
  On the other hand, the optimal contract induces $\underline{x}_2(\bar{x}_1, \underline{x}_1) = 0$ if

$$\frac{2(1-p)\left[\sqrt{1+p} - p\right]}{1 + p(1-p)} \leq \Delta\theta \leq \frac{2(1-p)}{1+p}.$$

  It remains to verify that all constraints are satisfied in these two regions. In region A, $\bar{x}_1 = \bar{x}_1^{FI} + \frac{2(1-p)-\Delta\theta(1+2p)}{1+p}$ and $\underline{x}_1 = \underline{x}_1^{FI}$. Constraint (b) requires that

$$\Delta\theta \leq \frac{2(1-p)}{p}.$$

[15] One can verify that the program is concave so that FOCS are also sufficient.

Since $\frac{2(1-p)}{p} > \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)}$, this constraint is verified. Constraint (a) requires that $\underline{x}_2(\bar{x}_1, \underline{x}_1) = \underline{x}_2^{FI} - \frac{p[2(1-p)+\Delta\theta]}{1-p^2} > 0$. This is the case for

$$\Delta\theta \leq \frac{2(1-p)}{1+p(1-p)}.$$

Since $\frac{2(1-p)}{1+p(1-p)} > \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)}$, we can conclude that all constraints are verified in region A.

For $\frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)} \leq \Delta\theta \leq \frac{2(1-p)}{1+p}$, then $\bar{x}_1 = \bar{\theta}$, and $\underline{x}_1 = \underline{\theta} - \frac{p}{1-p}\Delta\theta > 0$. Clearly constraint (b) is verified. It remains to check that indeed $\underline{x}_2(\bar{x}_1, \underline{x}_1) = 0$. This is the case for $\Delta\theta \geq \frac{2(1-p)^2}{1+p(1-p)}$. Since $\frac{2(1-p)^2}{1+p(1-p)} < \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)}$, then all constraints are satisfied also in this region.

- ii) For $\frac{2(1-p)}{1+p} \leq \Delta\theta \leq 2$, then

$$\frac{(4p^3 - p^2 + 1)\Delta\theta^2 + 4(4p^3 - 3p^2 - 1)\Delta\theta + 4(4p^3 - 9p^2 + 4p + 1)}{8(1-p^2)} \geq \frac{p}{2}(1+\frac{\Delta\theta}{2})^2,$$

  if $\Delta\theta \leq \Delta\theta_1$ and $\Delta\theta \geq \Delta\theta_2$, where $\Delta\theta_1 = \frac{2(1-p)\left(5p^2+2p+1-2\sqrt{5}p\sqrt{(p+1)}\right)}{5p^3-p^2+1-p}$ and $\Delta\theta_2 = \frac{2(1-p)\left(5p^2+2p+1+2\sqrt{5}p\sqrt{(p+1)}\right)}{5p^3-p^2+1-p}$.

  Since $\Delta\theta_1 < \frac{2(1-p)}{1+p}$, then the optimal contract is such that $\underline{x}_2(\bar{x}_1, \underline{x}_1) > 0$ only if $\Delta\theta \geq \Delta\theta_2$. However, $\Delta\theta_2 > \frac{2(1-p)}{1+p(1-p)}$, and therefore $\underline{x}_2(\bar{x}_1, \underline{x}_1)$ is never strictly positive for $\Delta\theta \geq \Delta\theta_2$. It follows that the unique candidate is $\bar{x}_1 = \bar{\theta}$, and $\underline{x}_1 = 0$. This contract satisfies both constraints (a) and (b) and it is thus the optimal contract for this region.

- Combining i) with ii) we can conclude that in region B the optimal contract is characterized by $\bar{x}_1 = \bar{\theta}$, $\underline{x}_1 = \max\left\{\underline{\theta} - \frac{p}{1-p}\Delta\theta, 0\right\}$.

- The follower's decisions are obtained by replacing the leader's decisions into the reaction functions derived in Proposition 3. ■

Proof of Proposition 6.

First we prove that if $(\overline{IC}_1)$ and $(\underline{IR}_1)$ are satisfied so is $(\overline{IR}_1)$. Using Lemma 2 we obtain that $(\overline{IC}_1)$ is equivalent to

$$\overline{U}_A \geq \underline{U}_A + E_{\underline{\delta}(s)}\left[\Delta_\theta v_A(\underline{x}_1, \underline{x}_2(s))\right] > 0,$$

because of Assumption A4 and $(\underline{IR}_1)$.

Second, we show that $(\overline{IC}_1)$ and $(\underline{IR}_1)$ bind. Using Lemma 2 and $\underline{t}_2(s)$, $\bar{t}_2(s)$, $(\underline{IC}_1)$ reduces to

$$\overline{U}_A \leq \underline{U}_A + E_{\bar{\delta}(s)}\left[\Delta_\theta v_A(\bar{x}_1, \underline{x}_2(s))\right].$$

Replacing $\underline{t}_1(s)$, $\bar{t}_1(s)$ with $\overline{U}_A$ and $\underline{U}_A$ into P_1's objective function it is immediate to see that it is optimal to make both $(\overline{IC}_1)$ and $(\underline{IR}_1)$ bind. ■

Proof of Proposition 7.

Take the reduced program as in Proposition 6. When the two decisions are independent, the payoff of the agent is separable in the two decisions, i.e. $v_A(x_1, x_2, \theta) = g_A^1(x_1, \theta) + g_A^2(x_2, \theta)$. In this case the agent's rent conditional on signal s is

$$R(s) = \Delta_\theta v_A(\underline{x}_1, \underline{x}_2(s)) = \Delta_\theta g_A^1(\underline{x}_1) + \Delta_\theta g_A^2(\underline{x}_2(s))$$

and

$$R_2(s) = \Delta_\theta g_A^2(\underline{x}_2(s)).$$

It follows that P_1 must give to the efficient type an expected rent

$$\begin{aligned} ER_1(\underline{x}_1, \underline{\delta}(s), \bar{x}_1, \overline{\delta}(s)) &= E_{\underline{\delta}(s)} R(s) - E_{\overline{\delta}(s)} R_2(s) \\ &= \Delta_\theta g_A^1(\underline{x}_1) + \left[E_{\underline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s)) - E_{\overline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s))\right], \end{aligned}$$

where for any $s \in S$, $\underline{x}_2(s) = \max\{\underline{x}_2', 0\}$ with $\underline{x}_2'$ implicitly defined by

$$\frac{\partial v_2(\underline{x}_2', \underline{\theta})}{\partial x_2} + \frac{\partial g_A^2(\underline{x}_2', \underline{\theta})}{\partial x_2} - \left(\frac{\mu(\overline{\theta}/s)}{\mu(\underline{\theta}/s)}\right)\left[\frac{\partial g_A^2\left(\underline{x}_2', \overline{\theta}\right)}{\partial x_2} - \frac{\partial g_A^2(\underline{x}_2', \underline{\theta})}{\partial x_2}\right] = 0.$$

The optimal contract π_1^* must be characterized by two probability distributions that minimize $\left[E_{\underline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s)) - E_{\overline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s))\right]$.

Since $v_2(\underline{x}_2', \underline{\theta})$ and $g_A^2(\underline{x}_2', \underline{\theta})$ are concave and $\frac{\partial g_A^2(\underline{x}_2', \overline{\theta})}{\partial x_2} - \frac{\partial g_A^2(\underline{x}_2', \underline{\theta})}{\partial x_2} > 0$ under A3, $\underline{x}_2(s)$ is decreasing in the hazard rate $\frac{\mu(\overline{\theta}/s)}{\mu(\underline{\theta}/s)} = \frac{p\overline{\delta}(s)}{(1-p)\underline{\delta}(s)}$.

The hazard rate is increasing in $\overline{\delta}(s)$ and decreasing in $\underline{\delta}(s)$. But then $\Delta_\theta g_A^2(\underline{x}_2(s))$ is decreasing in $\overline{\delta}(s)$ and increasing in $\underline{\delta}(s)$.

Suppose to start with $\overline{\delta}(s) = \underline{\delta}(s)$ for all $s \in S$ and then increase by ε the probability of signal s_j conditional on $\overline{\theta}$ and decrease by ε the probability of signal s_k, again conditional on $\overline{\theta}$. In other words, let $\overline{\delta}'(s_j) = \overline{\delta}(s_j) + \varepsilon$ and $\overline{\delta}'(s_k) = \overline{\delta}(s_k) - \varepsilon$. In this case

$$\left[E_{\underline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s)) - E_{\overline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s))\right] = \varepsilon\left[\Delta_\theta g_A^2(\underline{x}_2(s_k)) - \Delta_\theta g_A^2(\underline{x}_2(s_j))\right].$$

Since the hazard rate in s_k is smaller than in s_j, i.e. $\frac{\mu(\overline{\theta}/s_k)}{\mu(\underline{\theta}/s_k)} < \frac{\mu(\overline{\theta}/s_j)}{\mu(\underline{\theta}/s_j)}$, then $[\Delta_\theta g_A^2(\underline{x}_2(s_k)) - \Delta_\theta g_A^2(\underline{x}_2(s_j))] > 0$. Repeating the argument for any pair of signals, we have that

$$\left[E_{\underline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s)) - E_{\overline{\delta}(s)} \Delta_\theta g_A^2(\underline{x}_2(s))\right] > 0$$

as long as $\overline{\delta}(s) \neq \underline{\delta}(s)$ and it is minimized for $\underline{\delta}(s) = \overline{\delta}(s)$ $\forall s \in S$ which implies no information transmission between the two principals. Finally, since $(\underline{IC}_1)$ is always satisfied when x_1 is monotonic and $\underline{\delta}(s) = \overline{\delta}(s)$ $\forall s \in S$, this is indeed a maximal point with respect to $\underline{\delta}(s)$ and $\overline{\delta}(s)$. ■

Proof of Lemma 3.

We prove this lemma by showing that for any pair of probability measures $\delta^*(./\bar{\theta}), \delta^*(./\underline{\theta})$ that give strictly positive probabilities to more than two signals, there always exists another pair of probability measures with $\bar{\delta} = \Pr(s_1/\bar{\theta}) = 1 - \Pr(s_2/\bar{\theta})$ and $\underline{\delta} = \Pr(s_1/\underline{\theta}) = 1 - \Pr(s_2/\underline{\theta})$ such that P_1 is (weakly) better off.

Clearly, this is true if it is optimal for P_1 not to signal any information to P_2.

Let us then concentrate on optimal contracts that involve information transmission between the two principals.

For any $s \in S$, $\underline{x}_2(s) = \max\left\{\underline{\theta} + \underline{x}_1 - \frac{p}{1-p}\frac{\delta^*(s/\bar{\theta})}{\delta^*(s/\underline{\theta})}\left[\Delta\theta + \bar{x}_1 - \underline{x}_1\right], 0\right\}$, or equivalently, $\underline{x}_2(s) = \max\left\{a - b\frac{\delta^*(s/\bar{\theta})}{\delta^*(s/\underline{\theta})}, 0\right\}$ where $a = \underline{\theta} + \underline{x}_1$ and $b = \frac{p}{1-p}\left[\Delta\theta + \bar{x}_1 - \underline{x}_1\right]$.

Let $S^{'} = \{s \in S \,/\underline{x}_2(s) > 0\}$ and $\overline{S}^{'} = S \backslash S^{'}$.

Assume $\delta^*(./\bar{\theta}), \delta^*(./\underline{\theta})$ are not Bernoulli so that they assign positive probability to more than two signals, and let $\#(S^{'}) = m > 2$. Let $\bar{\delta} = \sum_{s\in S^{'}} \delta^*(s/\bar{\theta})$ and $\underline{\delta} = \sum_{s\in S^{'}} \delta^*(s/\underline{\theta})$ be respectively the probability that $\underline{x}_2(s) > 0$ under $\delta^*(./\bar{\theta})$ and $\delta^*(./\underline{\theta})$.

P_1's reduced program shows that P_1's objective function is strictly increasing in $E_{\delta^*(./\bar{\theta})}\underline{x}_2(s)$ and decreasing in $E_{\delta^*(./\underline{\theta})}\underline{x}_2(s)$. Thus, for $\delta^*(./\bar{\theta}), \delta^*(./\underline{\theta})$ to be part of P_1's optimal mechanism, it is necessary that there does not exist another pair of probability measures $\delta^{'}(./\bar{\theta}), \delta^{'}(./\underline{\theta})$ that increase P_1's objective function while relaxing the constraint $(\underline{IC}_1)$. We show that this is indeed the case only if $\delta^*(./\bar{\theta}), \delta^*(./\underline{\theta})$ are two Bernoulli distributions.

$$E_{\delta^*(./\underline{\theta})}\underline{x}_2(s) = \sum_{s\in S^{'}}\left[a - b\frac{\delta^*(s/\bar{\theta})}{\delta^*(s/\underline{\theta})}\right]\delta^*(s/\underline{\theta}) = a\underline{\delta} - b\bar{\delta}.$$

Similarly,

$$E_{\delta^*(./\bar{\theta})}\underline{x}_2(s) = \sum_{s\in S^{'}}\left[a - b\frac{\delta^*(s/\bar{\theta})}{\delta^*(s/\underline{\theta})}\right]\delta^*(s/\bar{\theta}) = a\bar{\delta} - b\sum_{s\in S^{'}}\frac{\delta^{*2}(s/\bar{\theta})}{\delta^*(s/\underline{\theta})}.$$

Suppose P_1 replaces $\delta^*(./\bar{\theta}), \delta^*(./\underline{\theta})$ with a pair of Bernoulli of parameters $\bar{\delta}$ and $\underline{\delta}$ so that $\bar{\delta} = \Pr(s_1/\bar{\theta}) = 1 - \Pr(s_2/\bar{\theta})$, and $\underline{\delta} = \Pr(s_1/\underline{\theta}) = 1 - \Pr(s_2/\underline{\theta})$. Let $E_{\bar{\delta}}\underline{x}_2(s)$, $E_{\underline{\delta}}\underline{x}_2(s)$ be respectively the expectation of $\underline{x}_2(s)$ under $\bar{\delta}$ and $\underline{\delta}$.

We claim that $E_{\bar{\delta}}\underline{x}_2(s) \geq E_{\delta^*(./\bar{\theta})}\underline{x}_2(s)$ and $E_{\underline{\delta}}\underline{x}_2(s) = E_{\delta^*(./\underline{\theta})}\underline{x}_2(s)$. To prove this claim we first show that $\underline{x}_2^*(s_2) = 0$ when $\bar{\delta}$ and $\underline{\delta}$ are the two Bernoulli distributions defined above. Let $\hat{\delta}$ be the critical ratio such that $a - b\hat{\delta} = 0$. By construction, for all $s \in \overline{S}^{'}$ we have that $\frac{\delta^*(s/\bar{\theta})}{\delta^*(s/\underline{\theta})} \geq \hat{\delta}$ and therefore $\underline{x}_2(s) = 0$. It follows that $\sum_{s\in\overline{S}^{'}} \delta^*(s/\bar{\theta}) \geq \hat{\delta}\sum_{s\in\overline{S}^{'}} \delta^*(s/\underline{\theta})$, or equivalently $1 - \bar{\delta} \geq \hat{\delta}(1 - \underline{\delta})$. Hence under the probability measures $\bar{\delta}$, $\underline{\delta}$, for $s = s_2$ we have $\frac{1-\bar{\delta}}{1-\underline{\delta}} \geq \hat{\delta}$ so that $\underline{x}_2^*(s_2) = 0$.

It follows that $E_{\underline{\delta}}\underline{x}_2(s) = \underline{\delta}\left[a - b\frac{\bar{\delta}}{\underline{\delta}}\right] = a\underline{\delta} - b\bar{\delta} = E_{\delta^*(./\underline{\theta})}\underline{x}_2(s)$.

Similarly, $E_{\bar{\delta}}\underline{x}_2(s) = a\bar{\delta} - b\frac{\bar{\delta}^2}{\underline{\delta}}$.

To prove the lemma it suffices to show that $E_{\bar{\delta}}\underline{x}_2(s) \geq E_{\delta^*(./\bar{\theta})}\underline{x}_2(s)$ or, equivalently,

$$\sum_{s\in S'} \frac{\delta^{*2}(s/\bar{\theta})}{\delta^*(s/\underline{\theta})} \geq \frac{\left[\sum_{s\in S'} \delta^*(s/\bar{\theta})\right]^2}{\sum_{s\in S'} \delta^*(s/\underline{\theta})} = \frac{\bar{\delta}^2}{\underline{\delta}}$$

We make use of the following mathematical property. For any strictly positive scalars a, b, c, d, $\frac{a^2}{b} + \frac{c^2}{d} \geq \frac{(a+c)^2}{b+d}$. In fact, rearranging we have $a^2 d(b+d) + c^2 d(b+d) \geq bd(a+c)^2$ which is equivalent to $(ad - cb)^2 \geq 0$.

This simple property shows that P_1 can increase the expected value of $\underline{x}_2(s)$ replacing the initial pair of probability measures $\delta^*(./\bar{\theta})$, $\delta^*(./\underline{\theta})$ with a pair $\delta'(./\bar{\theta})$, $\delta'(./\underline{\theta})$ such that for all the first $m-2$ signals $s \in S'$ $\delta^*(s/\bar{\theta}) = \delta'(s/\bar{\theta})$ and $\delta^*(s/\underline{\theta}) = \delta'(s/\underline{\theta})$ and for the last two signals $\delta'(s_{m-1}/\bar{\theta}) = \delta^*(s_{m-1}/\underline{\theta}) + \delta^*(s_m/\underline{\theta})$ and $\delta'(s/\underline{\theta}) = \delta^*(s_{m-1}/\underline{\theta}) + \delta^*(s_m/\underline{\theta})$.

Repeating this argument recursively we can conclude that P_1 must use a pair of Bernoulli probability distributions. ■

Proof of Proposition 8.

For any contract such that $\underline{x}_2(s_1) > 0$, then $\underline{x}_2(s_1) = \underline{\theta} + \underline{x}_1 - \frac{p}{1-p}\frac{\bar{\delta}}{\underline{\delta}}(\Delta\theta + \bar{x}_1 - \underline{x}_1)$.

Taking the derivatives for $\bar{x}_1$ and $\bar{\delta}$ and equating them to zero, with simple algebraic manipulations, we obtain

$$\underline{x}_1 = \frac{\underline{\delta}[2(1-p)+\Delta\theta(1-3p)]+\bar{\delta}p(4+6\Delta\theta)}{(1-p)\underline{\delta}+2p\bar{\delta}}$$
$$\bar{x}_1 = \bar{x}_1^{FI} = 1 + \frac{\Delta\theta}{2}$$

Replacing $\underline{x}_1$ and $\bar{x}_1$ into the derivative of the objective function w.r.t. $\underline{\delta}$ we obtain

$$\frac{p^2\left[\Delta\theta\underline{\delta}(1-p) + \bar{\delta}2p(1+\Delta\theta) - \bar{\delta}(2+\Delta\theta)\right]^2}{(1-p)\left[\underline{\delta}(1-p) + 2p\bar{\delta}\right]^2}$$

which is always positive. It follows that $\underline{\delta}^* = 1$.

Taking the first order condition for $\bar{\delta}$ and substituting $\underline{\delta}^* = 1$, $\bar{x}_1^* = 1 + \frac{\Delta\theta}{2}$, $\underline{x}_1 = \frac{2(1-p)+\Delta\theta(1-3p)+\bar{\delta}p(4+6\Delta\theta)}{(1-p)+2p\bar{\delta}}$ we obtain $\bar{\delta}^*$. Substituting $\bar{\delta}^*$ we finally get $\underline{x}_1^*$ and $\underline{x}_2^*(s_1)$.

Now we prove that there exist a parameters' region in which this interior solution satisfies all the constraints in Π^C.

First note that when $\underline{\delta} = 1$, constraint (b) is immediately satisfied.

Constrained (c) requires that $\bar{\delta}^* = \frac{(1-p)[\Delta\theta - 2(1-p)]}{p[2(1-p)-3\Delta\theta]} \in [0, 1]$.

$\bar{\delta}^* \geq 0$ for

$$\frac{2}{3}(1-p) \leq \Delta\theta \leq 2(1-p),$$

$\bar{\delta}^* \leq 1$ for

$$\frac{2(1-p)}{1+2p} \leq \Delta\theta$$

It follows that $\bar{\delta}^* \in [0,1]$ for

$$\frac{2(1-p)}{1+2p} \leq \Delta\theta \leq 2(1-p).$$

Moreover, $\underline{x}_2(s_1) > 0$ and $\underline{x}_1 > 0$ for $\Delta\theta \leq \frac{2(1-p)}{1+p}$.
Hence all constraints (a) to (d) are satisfied in the region defined by:

$$\frac{2(1-p)}{1+2p} \leq \Delta\theta \leq \frac{2(1-p)}{1+p}.$$

It remains to verify that P_1 is not better off by designing a mechanism such that $\underline{x}_2^*(s_1) = \underline{x}_2^*(s_2) = 0$. Following the proof of Proposition 5, this is never the case for $\Delta\theta \leq \frac{2(1-p)\left[\sqrt{1+p}-p\right]}{1+p(1-p)} \leq \frac{2(1-p)}{1+p}$.

Under condition (i) the optimal contract is indeed the one of Proposition 8 and it satisfies all constraints (a) to (d).

Finally, condition (ii) assures that also constraint (e) is verified. ■

Proof of Proposition 8.

Let us assume that the parameters of the model are as in region $C \subset A$. If P_1 does not disclose information to P_2, the equilibrium allocations are given by Proposition 5 (region A). Under *no information transmission* (NIT) the three players obtain an expected payoff equal to

$$\begin{aligned} U_1^{*NIT} &= \frac{4(1-\Delta\theta)+\Delta\theta^2+16p-36p^2-\Delta\theta^2p^2+16\Delta\theta p^3+4\Delta\theta^2p^3-12\Delta\theta p^2+16p^3}{8(1-p^2)}, \\ U_2^{*NIT} &= \frac{4(1-\Delta\theta)+\Delta\theta^2+8p-12p^2-\Delta\theta^2p^2+4\Delta\theta p^3-\Delta\theta^2p^3+2\Delta\theta^2p}{2(1+p)(1-p^2)}, \\ U_A^{*NIT} &= \frac{\Delta\theta p[6-3\Delta\theta-2p(2+\Delta\theta)-p^2(2-3\Delta\theta)]}{2(1-p^2)}. \end{aligned}$$

Conversely, when π_1 involves information transmission (IT) we have

$$\begin{aligned} U_1^{*IT} &= \frac{20(1-\Delta\theta)-4p(9-\Delta\theta)+\Delta\theta^2(5+11p+4p^2)+16p^2(1+\Delta\theta)}{8(1-p)} \\ U_2^{*IT} &= \frac{4(1-\Delta\theta)(1-p)+(1+3p)\Delta\theta^2}{2(1-p)} \\ U_A^{*IT} &= \frac{5\Delta\theta p[2(1-p)-\Delta\theta(1+p)]}{2(1-p)} \end{aligned}$$

Comparing the expected payoffs we obtain that

$$\begin{aligned} U_1^{*IT}-U_1^{*NIT} &= \frac{[2(1-p)-\Delta\theta(1+p)]}{2(1-p^2)} > 0, \\ U_2^{*IT}-U_2^{*NIT} &= \frac{p(-4(1+\Delta\theta)+3\Delta\theta^2+8p+4\Delta\theta p+8\Delta\theta^2p-4p^2+4\Delta\theta^2p^2)}{2(1+p)(1-p^2)} > 0^{16}, \\ U_A^{*IT}-U_A^{*NIT} &= \frac{\Delta\theta p(1+2p)(2(1-p)-\Delta\theta(1+p))}{1-p^2} > 0. \end{aligned}$$

It follows that all players gain with information transmission. ■