SIRE DISCUSSION PAPER
SIRE-DP-2010-59

Optimal contracting with private information on cost expectation and variability

Daniel Danau
University of St Andrews

Annalisa Vinellay
University of Bari
Optimal contracting with private information on cost expectation and variability

Daniel Danau*  Annalisa Vinella†

Abstract

We study the screening problem that arises in a framework where, initially, the agent is privately informed about both the expected production cost and the cost variability and, at a later stage, he learns privately the cost realization. The specific set of relevant incentive constraints, and so the characteristics of the optimal mechanism, depend finely upon the curvature of the principal's marginal surplus function as well as the relative importance of the two initial information problems. Pooling of production levels is optimally induced with respect to the cost variability when the principal's knowledge imperfection about the latter is sufficiently less important than that about the expected cost.

Keywords: Cost uncertainty; Multidimensional screening; Sequential screening

J.E.L. Classification Numbers: D81, D82, D86

*University of St Andrews (School of Economics & Finance and CRIEFF), Castlecliffe, 1 the Scores, St Andrews, Fife, KY16 9AR (UK). Corresponding author. E-mail: Daniel.Danau@st-andrews.ac.uk
†University of Bari (Department of Economics and Mathematics), Via C. Rosalba 53, 70124 Bari (Italy). E-mail: a.vinella@dse.uniba.it
1 Introduction

Before undertaking new activities in uncertain environments, whether to be run for public or private purposes, firms typically develop a feasibility analysis that is meant to assess prospective costs. This analysis consists in determining not only the expected cost, but also the variability of the cost around the expected value, which provides a measure of the uncertainty associated with the activity.\(^1\) The outcome that is obtained is not necessarily publicly observable. Therefore, when firms perform the analysis for a delegated activity, they are likely to have an information advantage about the two assessed cost components \textit{vis-à-vis} the delegating party. Information problems of this kind may plague any contractual relationship in which some good or service is procured from an outside supplier under uncertainty about production costs. In particular, they are quite relevant for government agencies (regulators, public authorities and other institutions) dealing with firms (regulated monopolies, franchisees) for the execution of activities of general interest. As an illustration, in transportation projects, entrusted firms present forecasts that are generally overoptimistic. Not only this may reflect the presence of technical errors. It may also follow from the strategic manipulation of the truly estimated expected cost and degree of uncertainty, unless incentives to behave opportunistically are contractually removed.\(^2\)

Although it is natural that, in uncertain environments, agents hold private information about both the expected value of some parameter that matters in the relationship with the principal and its variability, the literature has not yet studied how principals should design screening mechanisms to properly address information issues of this kind. Riordan and Sappington [15] and Spulber [17] focus on situations in which, at the outset of the relationship with the principal, the agent is privately informed about the sole expected cost of production. Miravete [14] - [13] and Courty and Li [8] take a similar approach, though in a different context. They tackle the issue of private knowledge about the expected consumption benefit, looking at situations in which the agent is a consumer who purchases a product from a monopolist. Courty and Li [8] also analyze the case in which the consumer observes the variability of his perspective consumption benefit but not its expected value.

In this article, we characterize the optimal incentive mechanism for a production activity that a principal requires from an agent who is privately informed about both the expected cost and the variability of the cost at the contracting stage. In line with previous works, we make the study truly positive by focusing on situations in which, at a

\(^1\) A feasibility analysis can of course be required also to assess prospective benefits. However, in this study, we solely focus on costs.

\(^2\) Several studies about transportation projects provide evidence that, in the latter, costs and errors about expected costs turn out to be systematically bigger than originally estimated. It is argued that wrong predictions are largely due to firms’ strategic behaviour (see, for instance, Flyvbjerg [10], Flyvbjerg, Bruzelius and Rothengatter [11] and Flyvbjerg, Holm and Buhl [12]).
later stage, the agent observes privately also the realized production cost, which can be either low (the good state) or high (the bad state), i.e. we allow for sequential learning on the agent’s side. Hence, overall, the agent holds two pieces of private information (jointly representing the first-stage two-dimensional type) at the outset of the relationship with the principal and learns an additional piece of information once uncertainty vanishes.

By representing this information structure, we bridge the strand of literature on sequential learning problems (that we mentioned above) with the studies on simultaneous multidimensional information issues. The latter explore situations in which, at the contracting stage, the agent holds more than one piece of information related to either one activity (Armstrong [1], Asker and Cantillon [4]) or two activities (Dana [9], Armstrong and Rochet [2]). Nevertheless, the agent faces no uncertainty so that sequential learning is ruled out in those models.

To pick the most suitable kind of mechanisms, among those that could be adopted under the information structure of our interest, we take into account that the principal prefers to collect a report every time the agent acquires some private information prior to the policy choice, rather than to ignore it (for this argument, see Baron and Besanko [5], for instance). In our framework, this means that, on the one hand, the initial report should be two-dimensional, as in simultaneous multidimensional screening problems. On the other hand, the agent should be required to disclose information twice: when the relationship begins and after uncertainty vanishes. That is, screening should occur sequentially, as it is the case in Riordan and Sappington [15] and Courty and Li [8]. In line with these requirements, we focus on a mechanism whereby a prior menu of optional schedules, among which the agent chooses by reporting the first-stage two-dimensional type, is followed by a menu of specific contractual options, from which the agent draws the final policy by reporting the second-stage realization of the cost. In principle, while the sequential nature of information release should inhibit the agent’s ability to behave opportunistically, the two-dimensional nature of the agent’s initial information, together with the fact that it concerns a single activity, should restrict the principal’s ability to screen.\(^3\) How the optimal screening mechanism looks like in the framework of our interest is thus far from obvious.

To understand the main features of the optimal scheme, it is important to identify the variables that drive the agent’s incentives to (mis)represent private information and that, at the same time, constitute screening devices for the principal. With private information concerning the cost, the production levels to be assigned in the good and in the bad state are naturally used as second-stage screening devices. To induce truth-telling about the state realization, the former production level must be at least as large as the latter for each first-stage type. Once this requirement is met, provided that the agent

\(^3\)On this point, see also the discussion in Rochet and Stole [16], who review some of the papers recalled in the text.
is unlimitedly liable, the principal can easily find a payment scheme that addresses the second-stage information issue whatever the first-stage decisions. Yet it remains to ensure that the production levels chosen in compliance with the second-stage requirement, are first-stage incentive-compatible as well. This is more easily done by referring, at the first stage, to two quantity-related variables/devices, namely the expected production and the expected difference between the good- and the bad-state production levels (which is non-negative, given the second-stage requirement aforementioned). While the former would be standard in any sequential screening problem with privately known expected costs, the latter becomes relevant because of the information problem on the cost variability and is thus specific to our framework. Under these circumstances, the usual trade-off between allocation efficiency and extraction of information rents is expressed in expected terms and is thus linked to the quantity-related variables affecting the agent’s initial incentives.

Our results reveal that the solution to this trade-off and, hence, the features of the optimal mechanism depend finely on the joint work of two elements. One element resides in the characteristics of the principal’s preferences for the good produced by the agent. It is specifically represented by the curvature of the marginal surplus function. The other element is given by the relative importance of the principal’s knowledge imperfection about the two cost components. It is represented by the so-called "spread index", namely a ratio having, at numerator, the spread between possible cost means and, at denominator, the spread between possible cost variabilities.

Let us illustrate why these two elements matter jointly. Each rent unit that the principal can save at the first stage by distorting quantities at the second stage is associated with a loss in the expected marginal surplus and in the expected difference between the marginal surpluses of those quantities in the two possible states. Therefore, the efficiency losses that the principal incurs by inducing distortions depend upon the curvature of the marginal surplus function. On the other hand, the rent-extraction benefits that the principal obtains through the allocation distortions are related to which dimension of private information is more costly to her. This occurs because, while both expected production and expected difference increase with the good-state production level, they move in opposite directions as the bad-state production level is varied. A raise in the latter induces an increase in the rent that is given up to prevent cheating on cost expectation through the increase in expected production; it triggers a decrease in the rent that is given up to prevent cheating on cost variability through the decrease in the expected difference. The way the bad-state production is fixed in the optimal mechanism, in addition to the good-state production, captures the need to compromise these two contrasting effects. The rent-extraction benefits are related to the magnitude of the spread index through this channel.

Once the joint relevance of the two core elements is understood, one has the key to interpret the general findings of our analysis. They are described as follows. The agency
cost that the principal bears is highest when her marginal surplus function is very convex and when the spread index is very low. Intuitively, as the marginal surplus function proceeds from a very concave to a very convex shape, the efficiency losses that ensue from distorting the good-state production levels becomes increasingly more important, relative to those ensuing from distorting bad-state production levels, as the latter are smaller. Moreover, having the spread index low means that the two first-stage information problems are both concerning for the principal. Under this circumstance, because of the contrasting effects triggered by variations in the bad-state production levels, it is convenient to point to adjustments mainly in the good-state productions to contain the overall rent. However, this strategy is costly to the principal on efficiency ground, especially with a very convex marginal surplus function, as we said. This difficulty is lessened when the spread index is large i.e., when the information problem about the cost variability at the first stage is relatively weak. Then, the principal can rely more deeply on distortions in the bad-state production levels, thereby retaining more surplus from the agent. Furthermore, private knowledge about cost variability being of little concern, (distorted) productions are optimally differentiated only with respect to the other information dimension. That is, pooling is induced in the production levels that are designed for different realizations of cost variability. Under this result, the optimal mechanism displays similarities with those characterized in contexts with publicly observable cost variability (Riordan and Sappington [15], Courty and Li [8]). The mechanism that we pin down naturally collapses onto the latter in the limit case in which the information problem about the cost variability disappears.

The reminder of the article is organized as follows. After reviewing the mainly related literature hereafter, we present the basic analytical setup in Section 2. In Section 3, we develop a few preliminary steps of analysis and describe the first-best benchmark. We characterize the optimal mechanism in Section 4. We conclude in Section 5. Mathematical details are mainly presented in an Appendix.

1.1 The relationship with the literature

In a early work on agency relationships with information evolution over time, Riordan and Sappington [15] explore the problem of a regulator who auctions out a franchise public-service contract to a firm that holds private information about the expected cost at the tendering stage and privately observes the realized cost at a later stage. Still in an

\[4\]\n
A few more recent studies extend the analysis of Riordan and Sappington [15] to environments in which either the contract includes such additional factors as product quality (Che [7]; see also Che [6] for an overview) or it accounts for various sources of information asymmetries (Armstrong and Sappington [3]) or both (Asker and Cantillon [4]). However, these papers do not tackle problems related to information learning on the agent’s side. The information structure they consider is tantamount to having perfectly correlated first- and second-stage information in the model of Riordan and Sappington [15] (on this point compare footnote 5 in Che [7]).
auction model, Spulber [17] assumes that, at the tendering stage, each participant knows
privately the possible levels of a cost overrun that he may incur at a later stage, whereas
the basic cost is commonly observed. This is tantamount to having the agent privately
informed about either the cost expectation or its variability, with perfect correlation
between the two. Hence, private information at the tendering stage is one-dimensional,
as in the model of Riordan and Sappington [15]. Miravete [14] - [13] and Courty and
Li [8] represent an information problem analogous to that of Riordan and Sappington
[15], in a different setting. Specifically, they consider a monopolist selling a product to a
consumer who knows privately his expected benefit from consumption at the initial stage
and observes privately his actual preference for the product at a later stage. Additionally,
Courty and Li [8] investigate the alternative situation in which the customer is initially
informed about his taste variability, rather than its expected value.\(^5\) None of these
authors considers the possibility that the agent hold two-dimensional private information
(including both the expectation and the variability of the relevant cost/benefit) at the
time he signs the contract with the principal, whereas we do so with regards to the cost of
the agent’s activity. The focus on a sequential mechanism, which we share with Riordan
and Sappington [15] and Courty and Li [8], allows us to account for the principal’s wish
to receive a report every time the agent learns something privately, as we mentioned.
From this standpoint, we rather diverge from Spulber [17] and Miravete [14] - [13]. In
the former model, second-stage information disclosure is unfeasible because contracts are
not enforceable; in the latter, sequential screening is neglected as the goal is to compare
two mechanisms that are based on either first- or second-stage information release only.

Among the studies on simultaneous multidimensional screening, Dana [9] and Arm-
strong and Rochet [2] assume that the agent executes two activities for the principal,
holding private information about the cost of either activity. In this setting, the produc-
tion level of each activity is used as a screening instrument for the corresponding piece
of private information. In Asker and Cantillon [4], the agent runs a single business and
knows privately both the operating and the fixed cost. As the latter is not related to
the produced quantity, the principal has one sole screening device related to the produc-
tion level. Similarly, in Armstrong [1] the agent executes a single activity but because
his two pieces of information (namely, production cost and product demand) are both
related to the production level, the principal uses the (sole) product price to screen both.
In our model, as in Asker and Cantillon [4] and in Armstrong [1], the agent is devoted
to a single activity. Nonetheless, unlike in those papers, two quantity-related screening
devices, namely the expected production and the expected difference between produc-

\(^5\)In details, Courty and Li [8] consider two types of customers having a continuum of possible valuations
of the product and study the following two situations: \((i)\) one type first-order stochastically dominates
the other; \((ii)\) one type faces greater valuation uncertainty than the other in the sense of mean-preserving
spread. The first case corresponds to having private information about the sole cost expectation, the
second case to having private information about the sole cost variability.
tion levels, come up to be available for the principal at the stage in which she faces the two-dimensional information problem.

2 The basic setup

We consider the relationship between a principal (P) and an agent, both risk-neutral, for the production of $q$ units of some good at a payment $t$.\(^6\) The expected unit cost of production $\theta$ is drawn from the set $\{\theta_L, \theta_H\}$ with commonly known probabilities $\nu$ and $1 - \nu$, respectively. We denote $\Delta \theta = \theta_H - \theta_L > 0$. The true unit cost is realized after the contract is signed and before production takes place. It can be either $\theta - \sigma$ or $\theta + \sigma$ with equal probabilities. The former (low cost) represents the "good" state resulting from a positive shock; the latter (high cost) represents the "bad" state resulting from a negative shock.\(^7\) By attaching equal probabilities to these two events, we prevent that the otherwise asymmetric distribution of high and low unit costs impose structure on the optimal mechanism. The parameter $\sigma$ expressing the uncertainty about the cost realization is drawn from the set $\{\sigma_L, \sigma_H\}$ with commonly known probabilities $\mu$ and $1 - \mu$, respectively. Throughout the article, we refer to $\sigma$ as to the variability of the unit cost. We also denote $\Delta \sigma = \sigma_H - \sigma_L > 0$. We hereafter refer to the generic realization of the two cost parameters as to $\theta_i$ and $\sigma_j$, with $i, j \in \{L, H\}$. To simplify the analysis, we take $\theta_L, \theta_H, \sigma_L$ and $\sigma_H$ to be such that $\Delta \theta > \Delta \sigma$. In words, we suppose that the knowledge imperfection about the expected unit cost is more important than that about the cost variability. This means that higher expected cost corresponds to higher true cost, even when it is associated with higher variability and the good state is realized (i.e., $\theta_H - \sigma_H > \theta_L - \sigma_L$).

**Information structure** Before sitting at the contracting table, the agent observes privately the expected unit cost $\theta_i$, $i \in \{L, H\}$, as well as the cost variability $\sigma_j$, $j \in \{L, H\}$. Hence, when the contract is drawn up, he enjoys a double information advantage. We denote $ij$ the agent’s type for any realized pair $(\theta_i, \sigma_j)$ and $\Upsilon \equiv \{LL, LH, HL, HH\}$ the set of feasible types. Moreover, the agent acquires a new information advantage when the state of nature is determined. Indeed, he learns privately whether the latter is $\theta_i - \sigma_j$ or $\theta_i + \sigma_j$.

---

\(^6\)As far as a public project or an activity of general interest is concerned, the agent can be viewed as a contractor, a supplier, a regulated (possibly local) monopoly; the principal as a government agency, a regulator etc.

\(^7\)For instance, these shocks could reflect variations in the price of some input that is used to produce the good. Principals may be (and are typically) unable to observe the prices at which agents purchase production inputs from their suppliers.
The relationship between $P$ and the agent unfolds over two stages but there is no discounting. For each $ij \in \Upsilon$, we let $(\bar{q}_{ij}, \bar{t}_{ij})$ and $(\bar{q}_{ij}, \bar{\tau}_{ij})$ denote the allocations to be implemented at the second stage, respectively in states $\theta_i - \sigma_j$ and $\theta_i + \sigma_j$. Accordingly, the agent’s second-stage profits are given by

$$\pi_{ij} = \bar{t}_{ij} - (\theta_i - \sigma_j) \bar{q}_{ij} \quad (1)$$
$$\pi_{ij} = \bar{\tau}_{ij} - (\theta_i + \sigma_j) \bar{q}_{ij} \quad (2)$$

His first-stage payoff is written

$$\Pi_{ij} = \frac{1}{2}(\pi_{ij} + \bar{\pi}_{ij}). \quad (3)$$

Production of $q$ units of the good by the agent yields gross surplus $S(q)$ to $P$. We take the function $S(\cdot)$ to be three-time differentiable (for technical reasons), with $S' > 0$ and $S'' < 0$, $\forall q \in \mathbb{R}_+$ (as usual). We further assume that $\lim_{q \to 0} S'(q)$ is finite but sufficiently large and $\lim_{q \to +\infty} S'(q)$ sufficiently small to ensure that an interior solution exists.\footnote{We assume finite limits, rather than the more standard Inada conditions, in order not to impose any restrictions on the shape of the function $S'(\cdot)$, which is going to be a crucial aspect in our analysis.} Lastly, to rule out mixed cases, we suppose that $S'''$ has constant sign $\forall q \in \mathbb{R}_+$. The first-stage payoff of $P$ is given by

$$V_{ij} = \frac{1}{2} \left\{ S(q_{ij}) - \bar{t}_{ij} + S(\bar{q}_{ij}) - \bar{\tau}_{ij} \right\}. \quad (4)$$

2.1 The programme of the principal

The Revelation Principle applies. This justifies restricting attention to mechanisms that induce truthful reporting. The optimal mechanism is the solution to the following programme, to be denoted $\Gamma$:

$$\begin{align*}
\max_{(q_{ij}, \bar{q}_{ij}, \bar{\pi}_{ij}, \bar{\tau}_{ij})} & \quad \frac{1}{2} \sum_{ij \in \Upsilon} E_{ij} \left[ (S(q_{ij}) - \bar{t}_{ij}) + (S(\bar{q}_{ij}) - \bar{\tau}_{ij}) \right] \\
\text{subject to} & \quad \Pi_{ij} \geq \frac{1}{2} \left\{ [\bar{q}_{i'j'} - (\theta_i + \sigma_j) \bar{q}_{i'j'}] + [\bar{t}_{i'j'} - (\theta_i - \sigma_j) \bar{q}_{i'j'}] \right\}, \forall i'j' \in \Upsilon \quad (5) \\
\bar{\pi}_{ij} & \geq \bar{\pi}_{ij} + 2\sigma_j \bar{q}_{ij}, \forall ij \in \Upsilon \quad (6) \\
\pi_{ij} & \geq \pi_{ij} - 2\sigma_j \bar{q}_{ij}, \forall ij \in \Upsilon \quad (7) \\
\Pi_{ij} & \geq 0, \forall ij \in \Upsilon. \quad (8)
\end{align*}$$

The objective function in $\Gamma$ mirrors $P$’s purpose to maximize the expected level of surplus net of the payment to be made to the agent for the production of the good. The inequality in (5) represents the set of first-stage incentive constraints, which ensure that each type
The inequalities in (6) and (7) represent the sets of second-stage incentive constraints in the good and in the bad state, respectively, which warrant that each type correctly announces whether a positive or a negative shock has occurred. Finally, (8) is the set of participation constraints under which each type is guaranteed non-negative expected payoff.

The problem compares with the programmes presented in Riordan and Sappington [15] and in Courty and Li [8]. Unlike in the former model, which stylizes an auction setting, here the principal deals with only one agent since the very beginning of the relationship. More importantly, unlike in either model, at the contracting stage the agent holds two pieces of information. Hence, he is to be induced to announce them both correctly.

As usual, for the agent to release information at the second stage (which occurs when both (6) and (7) are satisfied), it is necessary that \( P \) sets \( q_{ij} \geq \bar{q}_{ij} \). That is, the quantity produced in the good state \((\theta_i - \sigma_j)\) is to be at least as large as the quantity produced in the bad state \((\theta_i + \sigma_j)\). Throughout the analysis, we shall sometimes refer to \( q_{ij} \) as the "high" production level and to \( \bar{q}_{ij} \) as the "low" production level.

In addition to being second-stage incentive-compatible, the pair of quantities \((q_{ij}, \bar{q}_{ij})\) must be such that the \( ij \)-agent truth-tells at the first stage. To see under which conditions this occurs, first observe that the expected total cost of the \( ij \)-agent reporting \( i'j' \) is written

\[
EC_{ij}(q_{i'j'}, r_{i'j'}) = \theta_i q_{i'j'} - \sigma_j r_{i'j'},
\]

where

\[
q_{i'j'} \equiv \frac{1}{2}(q_{i'j'} + \bar{q}_{i'j'}) \quad \text{and} \quad r_{i'j'} \equiv \frac{1}{2}(q_{i'j'} - \bar{q}_{i'j'})
\]

respectively denote the expected production and the expected difference between high and low production levels that \( P \) commends when she receives the report \( i'j' \). Using (9), the set of first-stage incentive constraints in (5) is re-expressed as

\[
\Pi_{ij} \geq \Pi_{i'j'} + EC_{ij}(q_{i'j'}, r_{i'j'}) - EC_{i'j'}(q_{i'j'}, r_{i'j'}), \quad \forall ij, i'j' \in \Upsilon.
\]

This formulation highlights that the agent’s incentives to (mis)represent information at the first stage can be related directly to the expected productions and to the expected production differences of the types he might want to mimic \((q_{i'j'} \text{ and } r_{i'j'})\), suggesting that \( P \) will be based on this link to contrast opportunistic behaviour. From this perspective,
provided the second-stage condition $q_{ij} \geq \bar{q}_{ij}$ is met for all $ij \in \mathcal{Y}$, the programme $\Pi$ displays a similarity with the static two-dimensional information problems in which two quantity-related screening devices are at hand because the agent executes two distinct activities for the principal (Dana [9] and Armstrong and Rochet [2]). Here, however, a complication follows from the circumstance that $q$ and $r$ pertain to the same activity and are thus interlinked, involving that the solution to $\Pi$ does not simply replicate the solution that is found for those problems.

**Timing** To sum up, the game between $P$ and the agent unfolds as follows. Prior to sitting at the contracting table, the agent observes privately $\theta_i$ and $\sigma_j$. At the first stage, $P$ offers to the agent the truthful menu of optional contracts $\{(q_{ij}, l_{ij}); (\bar{q}_{ij}, \bar{l}_{ij})\}, \forall ij \in \mathcal{Y}$. The agent reports $ij$ to $P$ and the contract targeted to type $ij$ is signed. Both parties fully commit to this contract. At the second stage, the agent observes privately whether a positive or a negative shock has affected the cost (and so whether the realized state is good $(\theta_i - \sigma_j)$ or bad $(\theta_i + \sigma_j)$) and reports it to $P$. Accordingly, out of the stipulated contract, either the allocation $(q_{ij}, l_{ij})$ or the allocation $(\bar{q}_{ij}, \bar{l}_{ij})$ is effected.

### 3 A few preliminary steps of analysis

As usual in full-commitment frameworks without limited liability on the agent’s side, $P$ has no difficulty at designing transfers $l_{ij}$ and $\bar{l}_{ij}$ that are second-stage incentive-compatible, for any given pair of production levels satisfying the required monotonicity condition $q_{ij} \geq \bar{q}_{ij}$ and for any given rent $\Pi_{ij}$ to be given up at the first stage. Thus, once it is ensured that $q_{ij} \geq \bar{q}_{ij}$, the analysis can be developed focusing on the first-stage information problem.

The four possible first-stage types are "ordered" as follows:

$$LH > LL > HH > HL.$$  \hspace{1cm} (11)

This ranking reflects increasing values of the expected total cost of production, which we interpret as decreasing degrees of efficiency in executing the activity. More precisely, $LH$ is the most efficient type and $HL$ the least efficient type because, for any given pair $(q_{ij}, r_{ij})$, the expected cost is lower the lower $\theta_i$ and the higher $\sigma_j$. Types $LL$ and $HH$ both display an intermediate degree of efficiency but type $LL$ is more efficient than type $HH$ because $q_{ij} > r_{ij}$ (by definition) and $\Delta \theta > \Delta \sigma$ (by assumption).

To characterize the solution to $\Pi$, it is necessary to identify the constraints that are potentially binding. Once the second-stage monotonicity condition is satisfied, the sole relevant participation constraint is that of the least efficient type (here type $HL$). This is usually the case in sequential screening problems (compare Courty and Li [8]; see also
It is also the case in simultaneous multidimensional screening problems in contexts where the agent executes two activities (see Armstrong and Rochet [2], for instance). Moreover, as it is typical of the latter category of problems, relevant incentive constraints are some downward constraints whereby more efficient types not be willing to mimic some less efficient type. In our setting, these are the incentive constraints whereby type $LH$ not be tempted to choose any other type’s contract, type $LL$ not be attracted by either the type–$HH$ or the type–$HL$ contract and type $HH$ not be interested in the type–$HL$ contract. Accordingly, the rents of the four possible types are written as follows (see Appendix for details):

$$
\begin{align*}
\Pi_{HL} & = 0 \\
\Pi_{HH} & = \Delta \sigma_{HL} \\
\Pi_{LL} & = \max \{ \Pi_{LL,1}; \Pi_{LL,2} \} \\
\Pi_{LH} & = \max \{ \Pi_{LH,1}; \Pi_{LH,2}; \Pi_{LH,3} \},
\end{align*}
$$

where

$$
\Pi_{LL,1} = \Delta \theta q_{HL} \quad \text{and} \quad \Pi_{LL,2} = \Delta \sigma_{HL} + \Delta \theta q_{HH} - \Delta \sigma_{HH}
$$

as well as

$$
\Pi_{LH,1} = \Delta \theta q_{HH} + \Delta \sigma_{HL}, \quad \Pi_{LH,2} = \Pi_{LL} + \Delta \sigma_{LL} \quad \text{and} \quad \Pi_{LH,3} = \Delta \theta q_{HL} + \Delta \sigma_{HL}.
$$

The solution to $\Gamma$ is differently characterized, depending upon the specific values that the payoffs $\Pi_{LL}$ and $\Pi_{LH}$ take in (12a) to (12d). We shall enter into these details in a moment, after describing the first-best allocation, which $P$ would effect in a context of complete information.

### 3.1 The first-best allocation

At first best (FB hereafter), the $ij$–agent produces the quantities $q_{ij}^*$ and $\overline{q}_{ij}$ such that $S'(q_{ij}^*) = \theta_i - \sigma_j$ and $S'(\overline{q}_{ij}) = \theta_i + \sigma_j$, respectively, where the star is appended to indicate FB values. Moreover, he receives no rent i.e., $\Pi_{ij}^* = 0, \forall ij \in \Upsilon$.

For future reference, it is useful to notice that, given the symmetry of the cost distribution, at FB, the expected marginal benefit from the type–$ij$ production equals the expected unit cost i.e., $\frac{1}{2}[S'(q_{ij}^*) + S'(\overline{q}_{ij})] = \theta_i, \forall ij \in \Upsilon$, and is thus unrelated to the second cost dimension. On the other hand, the expected difference between marginal benefits from high and low productions equals the variability of the unit cost i.e., $\frac{1}{2}[S'(\overline{q}_{ij}) - S'(q_{ij}^*)] = \sigma_j, \forall ij \in \Upsilon$, and is thus unrelated to the first cost dimension.

Recall that, under asymmetric information, the agent’s incentives to (mis)represent the first-stage type are directly linked to the expected production $q$ and to the expected
production difference $r$. With regards to these expected quantities, it is further noteworthy that FB levels are not necessarily ordered according to the types’ efficiency ranking. To be more precise, FB expected productions ($q_{ij}^*$) are not necessarily ordered according to the efficiency ranking in (11), despite that the expected cost increases with $q$ (compare (9)). In turn, FB expected differences ($r_{ij}^*$) are not necessarily ranked in inverse order with respect to (11), despite that the expected cost decreases with $r$ (compare (9) again).

Actually, the way the FB values of $q$’s and $r$’s are ordered depends upon the shape of the principal’s marginal surplus function. For instance, $q_{LH}^* \geq q_{HL}^*$ if and only if $S_L(\cdot)$ is convex, whereas $r_{HH}^* \geq r_{LH}^*$ if and only if $S_L(\cdot)$ is concave. Remarkably, this point delivers a preliminary hint about the optimal (second-best) mechanism. Under asymmetric information, the distortions that the principal induces in the values of $q_{ij}$ and $r_{ij}$; through those in the values of $g_{ij}$ and $\tilde{q}_{ij}$, depend finely on the characteristics of the principal’s preferences, provided the associated efficiency losses are strictly related to the shape of the marginal surplus function. This aspect will become fully apparent throughout the analysis below.

4 The second-best mechanism

We have previously mentioned that the second-best (SB hereafter) optimal mechanism specifies differently, according to the particular values that $\Pi_{LL}$ and $\Pi_{LH}$ take. In the latter, two rent components are combined, namely $\Delta \theta q$, which is given up to prevent that the first dimension of private information be overstated, and $\Delta \sigma r$, which is given up to prevent that the second dimension be understated (see (13) and (14)).

Based on the FB analysis we deduced that the SB values of $q_{ij}$ and $r_{ij}$, $\forall ij \in \Upsilon$, are set according to the characteristics of the principal’s surplus function. Thus, rents clearly depend on those characteristics as well. This is not the sole determining element though. In addition, it matters how important the knowledge imperfection about the expected unit cost ($\Delta \theta$) is, relative to that about the cost variability ($\Delta \sigma$). This rests on the circumstance that any variation in the high-cost production level that the generic first-stage type $ij$ is assigned at the second stage, triggers opposite effects on the levels of $q_{ij}$ and $r_{ij}$ and, hence, on the magnitude of the rent components $\Delta \theta q_{ij}$ and $\Delta \sigma r_{ij}$. As a raise in the former comes along with a reduction in the latter (and vice versa), the optimal allocation choice cannot abstract from considerations about how $\Delta \theta$ compares with $\Delta \sigma$.

From now on, we shall conveniently use the ratio $\frac{\Delta \theta}{\Delta \sigma}$ as a "spread index" capturing the size of $\Delta \theta$ relative to that of $\Delta \sigma$.

11The $q$’s and the $r$’s that we refer to are specifically those included in the rents $\Pi_{LL}$ and $\Pi_{LH}$. Type subscripts are omitted for the sake of shortness.
4.1 Characterization of the mechanism

In what follows, we present the SB mechanism moving from the case in which the function $S' (\cdot)$ is very concave to that in which it is very convex, for diverse ranges of values of the spread index. The superscript $sb$ is used throughout to denote SB values, when different from FB values.

**Proposition 1** Assume $\frac{1-\mu}{\mu+\nu(1-\mu)} < \frac{\Delta \theta}{\Delta \sigma} < \frac{1-\nu \mu}{\nu}$. Further assume that $S' (\cdot)$ is "sufficiently" concave, such that $q_{HL}^{sb} > q_{HH}^{sb}$ and $r_{HL}^{sb} > r_{LL}^{sb}$. The optimal mechanism is characterized as follows:

- Rents are given by $\Pi_{HL}^s$, $\Pi_{HH}^s$, $\Pi_{LL,1}^{sb}$ and $\Pi_{LH,3}^{sb}$.
- Quantities are such that $q_{HL}^{sb} \geq q_{LL}^{sb}$ and $r_{HL}^{sb} \geq r_{LL}^{sb}$. i) If these two conditions are strictly satisfied, then only the type–HL productions are distorted away from the FB levels, such that $q_{HL}^{sb} < q_{HL}^s$ and $r_{HL}^{sb} < r_{HL}^s$. Otherwise, the type–LH production levels are also distorted, such that: ii) if $q_{HL}^{sb} = q_{HL}^s$, then $q_{HL}^{sb} > q_{LL}^s$; iii) if $r_{HL}^{sb} = r_{HL}^s$, then $r_{HL}^{sb} > r_{HL}^s$. Case (i) arises when $S' (\cdot)$ is not too concave as well as when $\frac{\Delta \theta}{\Delta \sigma} = \delta$, for some given $\delta \in \left(\frac{1-\mu}{\mu+\nu(1-\mu)}, \frac{1-\nu \mu}{\nu}\right)$. Case (ii) arises when $S' (\cdot)$ is very concave and $\frac{\Delta \theta}{\Delta \sigma} < \delta$, case (iii) when $S' (\cdot)$ is very concave and $\frac{\Delta \theta}{\Delta \sigma} > \delta$.

The proposition describes the characteristics of the optimal mechanism when type $LL$ and type $LH$ have both incentives to mimic type $HL$ (so that the associated incentive constraints are binding). Let us first focus on type $LL$. Recall that this type is tempted to report either $HL$ or $HH$. The former case arises when, at SB, $\Pi_{LL,1} > \Pi_{LL,2}$ i.e., when

$$\Delta \theta (q_{HL}^{sb} - q_{HH}^{sb}) \geq \Delta \sigma (r_{HL}^{sb} - r_{HH}^{sb}).$$

This condition is interpreted as follows. By over-reporting $\theta$, while truthtelling on $\sigma$, type $LL$ would obtain a marginal gain equal to $\Delta \theta (q_{HL} - q_{HH})$. At the same time, he would renounce to the benefit $\Delta \sigma (r_{HL} - r_{HH})$ i.e., to the profit designed for type $HH$ not being tempted to mimic type $HL$, net of the loss that would be associated with over-reporting $\sigma$. Overall, the $LL$–agent would prefer to overstate only the first information dimension (i.e., to report $HL$) as long as the net benefit associated with this lie is larger than that he would obtain from overstating also the second dimension (i.e., from reporting $HH$)

Before moving to consider type $LH$, we find it interesting to discuss the circumstances under which it is optimal to concede the rent $\Pi_{LL,1}$ to the $LL$–agent (and thus (15) holds). We previously mentioned that one relevant circumstance is the magnitude of the spread index. To clarify this aspect with specific regards to the situations considered in Proposition 1, it is useful to rewrite (15) as

$$(\Delta \theta + \Delta \sigma) (q_{HL}^{sb} - q_{HH}^{sb}) \geq (\Delta \theta - \Delta \sigma) (q_{HH}^{sb} - q_{HL}^{sb}).$$

13
For this inequality to hold, the difference \( q_{sb}^{HH} - q_{HL}^{sb} \) which is positive, should be low enough as compared to the difference \( q_{HL}^{sb} - q_{HH}^{sb} \). The latter should then be positive as well. However, this is not necessarily the case. The reason is that, as long as type \( LL \) has an interest in mimicking type \( HL \), raising the quantity that the latter is required to produce if the bad state is realized, triggers two opposite effects on the incentives to declare \( HL \) at the first stage. While the temptation to cheat on \( \theta \) is exacerbated (\( \Delta \theta q_{HL} \) is increased), it becomes less attractive to cheat also on \( \sigma \) (\( \Delta \sigma r_{HL} \) is decreased). On the other hand, requiring type \( HH \) to produce more in the bad state only tightens the incentives to cheat on \( \theta \). However, this is not necessarily the case. The reason is that, as long as type \( HL \) has an interest in mimicking type \( HH \), raising the quantity that the latter is required to produce if the bad state is realized, triggers two opposite effects on the incentives to declare \( HL \) at the first stage. While the temptation to cheat on \( \theta \) is exacerbated (\( \Delta \theta q_{HL} \) is increased), it becomes less attractive to cheat also on \( \sigma \) (\( \Delta \sigma r_{HL} \) is decreased). On the other hand, requiring type \( HH \) to produce more in the bad state only tightens the incentives to cheat on \( \theta \). The second effect of a raise in \( q_{HL} \) is more important the lower \( r_{HL} \). This explains why \( P \) sets \( q_{HL}^{sb} \) larger than \( q_{HH}^{sb} \) (so that (15) can be satisfied) if and only if the spread index is sufficiently small (precisely, if and only if it falls below the threshold \( \frac{1}{\mu + \nu(1-\mu)} \)).

As we deduced from the FB analysis, another relevant circumstance resides in the characteristics of \( P \)'s preferences for the good, which affect the solution to \( \Gamma \) on efficiency ground. We shall now look more closely at this aspect with specific regards to the context of Proposition 1. Recall that, as usual, the marginal surplus function is decreasing in quantity. When it is concave, the rate of decrease of the marginal surplus is larger the bigger the initial quantity. In the framework here considered, it means that, being quantities such that at optimum \( q_{HL} > q_{HH} \), the benefit that is additionally generated reduces more by moving from \( q_{HL} \) to some higher quantity than it does by moving from \( q_{HH} \). This further involves that, as \( S'(\cdot) \) becomes more concave, the wedge between the production levels of these first-stage types optimally decreases in either state, the rate of decrease being higher in the good state. In other words, keeping the difference \( q_{HL} - q_{HH} \) large, relative to the difference \( q_{HH} - q_{HL} \) (hence, increasing the difference \( q_{HL} - q_{HH} \)), enables \( P \) to reach a higher surplus. This illustrates why, when \( S'(\cdot) \) is sufficiently concave (the situation considered in Proposition 1), \( P \) optimally designs a mechanism such that (16) (or, equivalently, (15)) is satisfied.

We next consider type \( LH \). Being most efficient, this type might be willing to report any of the other types. Looking at the rents listed in (12a) to (12d), one deduces that type \( LH \) has incentives to mimic type \( HL \), rather than either \( LL \) or \( HH \), if and only if \( q_{HL} > q_{HH} \) together with \( r_{HL} > r_{LL} \). As the former inequality was discussed above, to avoid redundancy, we now focus on the latter. At optimum, type \( LL \) produces more than type \( HL \) in the good state (\( q_{LL} > q_{HL} \)) and, provided that \( \frac{\Delta \theta}{\Delta \sigma} > \frac{1}{\mu + \nu(1-\mu)} \), also in the bad state (\( q_{LL} > q_{HH} \)). As before, for more concave \( S'(\cdot) \), the difference between low production levels is kept large, at optimum, as compared to that between high production levels. Thus, if \( S'(\cdot) \) is sufficiently concave, then not only \( q_{HL} > q_{HH} \) but also \( r_{HL} > r_{LL} \) and type \( LH \) is actually tempted to mimic the least efficient type.

In definitive, Proposition 1 describes how the optimal mechanism is characterized in situations in which all types are willing to declare \( HL \). Accordingly, \( P \) distorts \( q_{HL} \) and
\(r_{HL}\) downward in order to contain rents.

A particular case arises when the marginal surplus function is very concave. Recall that, as \(S'(\cdot)\) becomes more concave, for the concerned (first-stage) types, \(P\) reduces the wedge between production levels more in the good than in the bad state. If \(P\) were to insist on the contract described so far when \(S'(\cdot)\) is very concave, then the wedge between \(\bar{q}_{LH}\) and \(\bar{q}_{HL}\) (which is positive for \(\frac{\Delta_0}{\Delta_{\sigma}} < \frac{1-\mu}{\mu+\nu(1-\mu)}\) and negative in the converse case) would become so important (in absolute value), relative to the wedge between \(q_{LH}\) and \(q_{HL}\) (which is positive), that either type \(LL\) or type \(HH\) would display unusual incentives to pretend to be more efficient. Specifically, with \(\bar{q}_{LH} > q_{LH}\), type \(LL\) would obtain a net benefit equal to \(\Delta \sigma (r_{HL}-r_{LH})\) at the first stage by mimicking the most efficient type rather than declaring \(HL\). On the other hand, with \(\bar{q}_{HL} > q_{HL}\), type \(HH\) would be tempted to mimic type \(LH\) to obtain the net benefit \(\Delta \sigma (q_{HL}-q_{LH})\) at the first stage. To remove the temptation to cheat, \(P\) optimally adjusts production levels at the second stage so that \(r_{HL}^b=r_{LH}^b\), with \(r_{LH}\) upward distorted, in the former case, and \(q_{HL}^b=q_{LH}^b\), with \(q_{LH}\) upward distorted, in the latter.\(^{12}\) This result is intriguing in that it describes a situation in which \(P\) distorts quantities (also) for the most efficient type to provide desirable incentives to types that exhibit an intermediate degree of efficiency.

**Corollary 1** Assume \(\frac{\Delta_0}{\Delta_{\sigma}} \leq \frac{1-\mu}{\mu+\nu(1-\mu)}\). Further assume that \(S'(\cdot)\) is sufficiently concave, such that \(q_{HL}^b > q_{HH}^b\). Then, \(r_{HL}^b = r_{LL}^b\). The optimal mechanism has the same characteristics as described in Proposition 1 for the case of \(\frac{\Delta_0}{\Delta_{\sigma}} < \delta\), except that \(\Pi_{LH,2}^b = \Pi_{LH,3}^b\) and that the type–\(LL\) production levels are such that \(r_{LL}^b < r_{LH}^b\).

Recall that the magnitude of the spread index reflects the importance of the information problem about the expected unit cost relative to that about the cost variability. The lower the index, the more costly it is to remove incentives to understate \(\sigma\), relative to removing incentives to overstate \(\theta\). In particular, when \(\frac{\Delta_0}{\Delta_{\sigma}} \leq \frac{1-\mu}{\mu+\nu(1-\mu)}\), under the mechanism described above, type \(LH\) would rather be attracted by the type–\(LL\) contract. To prevent cheating at the lowest cost, \(P\) now needs to concede the same rent for the \(LH\)–agent not reporting \(LL\) as she does for the \(LH\)–agent not reporting \(HL\) (i.e., \(\Pi_{LH,2}^b = \Pi_{LH,3}^b\)) and, unlike in the previous situation, to induce distortions in the type–\(LL\) production levels.

The next step of analysis is to consider situations in which \(S'(\cdot)\) is less concave than required in Proposition 1. To begin with, two particular cases can be identified, namely that in which \(q_{HL}^b = q_{HH}^b\) together with \(r_{HL}^b > r_{LL}^b\) and that in which \(q_{HL}^b > q_{HH}^b\) together with \(r_{HL}^b = r_{LL}^b\). In the former case, \(\Pi_{LH,1}^b = \Pi_{LH,3}^b\); in the latter, as in Corollary 1, \(\Pi_{LH,2}^b = \Pi_{LH,3}^b\). Because these cases are explained similarly to Corollary 1, we do not insist on them. We rather move to explore the more general situation in which the

\(^{12}\)For instance, as compared to the situation previously described, \(P\) reduces the gap between the production levels of type \(HL\) and raises that between those of type \(LH\) to warrant that \(r_{HL} = r_{LH}\).
concavity of the marginal surplus function is sufficiently weak that \( q^b_{HH} > q^b_{HL} \) jointly with \( r^b_{LL} > r^b_{HL} \). At this aim, it is useful to introduce the following definitions:

\[
\gamma_1 = \frac{[S'(q^b_{HL}) - S'(q^b_{LL})] - [S'(q^b_{HH}) - S'(q^b_{HL})]}{2(1-\nu)\mu \Delta \sigma}
\]

\[
\gamma_2 = \frac{[S'(q^b_{HH}) - S'(q^b_{HL})] - [S'(q^b_{HH}) - S'(q^b_{HL})]}{2(1-\nu)\mu \Delta \theta}
\]

**Proposition 2** Assume \( \frac{\Delta \theta}{\Delta \sigma} < \frac{1}{\sqrt{2}} \left[ \frac{1}{\nu} + (1 - \mu) \gamma_1 \right] \). Further assume that \( S'(\cdot) \) is neither very concave, such that both \( q^b_{HH} > q^b_{HL} \) and \( r^b_{LL} > r^b_{HL} \), nor very convex, such that the condition (15) holds together with

\[
\Delta \theta(q^b_{HH} - q^b_{HL}) = \Delta \sigma(r^b_{LL} - r^b_{HL}).
\]

The optimal mechanism is characterized as follows:

- rents are given by \( \Pi^b_{HL}, \Pi^b_{HH}, \Pi^b_{LL,1} \) and \( \Pi^b_{LL,2} = \Pi^b_{LH,2} \);
- production levels are distorted for types \( HL, HH \) and \( LL \) such that \( q^b_{HL} < q^b_{HH}, r^b_{LL} < r^b_{HH} \),

\[
r^b_{HL} < r^b_{LL}, q^b_{HH} < q^b_{HL}, r^b_{LH} < r^b_{HH}, r^b_{LL} < r^b_{HH},
\]

respectively.

The proposition states that, as \( S'(\cdot) \) becomes almost linear, on top of the inequalities \( q^b_{HH} > q^b_{HL} \) and \( r^b_{LL} > r^b_{HL} \) that we have already discussed, the additional condition (19) holds at SB, meaning that type \( LH \) is assigned the same rent for not mimicking either type \( HH \) or type \( LL \) (i.e., \( \Pi^b_{LH,1} = \Pi^b_{LH,2} \). To contain the rent accruing to type \( LH \), which is lowest when (19) is satisfied, \( P \) distorts output away from the FB levels for types \( HL, HH \) and \( LL \). The rent would not be minimized if production levels were distorted either for the sole types that could be announced by overstating (namely, \( HL \) and \( HH \)) or for the sole types that could be announced by understating \( \sigma \) (namely, \( HL \) and \( LL \).

To clarify this point, it is first helpful to insist on the peculiarity of the trade-off between efficiency losses and rent-extraction benefits in our framework. Here the basic trade-off involves second-stage allocations and first-stage incentives. Specifically, for each feasible first-stage type \( ij \), at the second stage quantities are distorted away from FB levels trading off, on one side, the ensuing loss in expected marginal surplus

\[
\frac{1}{2} [S'(q_{ij}) + S'(\bar{q}_{ij})]
\]

against the first-stage rents that depend upon the expected quantity \( q_{ij} \) and, on the other, the ensuing loss in expected difference between marginal surpluses

\[
\frac{1}{2} [S'(q_{ij}) - S'(\bar{q}_{ij})]
\]

against the first-stage rents that depend upon \( r_{ij} \).

We can now be based on this consideration to illustrate how the trade-off is solved in the situations considered in Proposition 2, taking into account the incentives to (mis)represent information that an agent of type \( LH \) displays. Suppose first that he has no interest in declaring \( LL \). Then, only the type—\( HL \) and the type—\( HH \) production
levels are conveniently distorted. As from the previous explanation, \( q_{HL} \) and \( \overline{q}_{HL} \), and thus \( q_{HL} \), are adjusted so that the resulting loss in expected surplus is traded off against the reduction in the rent \( \Delta \theta q_{HL} \), which \( P \) concedes to prevent type \( LL \) from declaring \( HL \). Similarly, \( q_{HH} \) and \( \overline{q}_{HH} \), and thus \( q_{HH} \), are adjusted so that the resulting loss in expected surplus is traded off against the reduction in the rent \( \Delta \theta q_{HH} \), which type \( LH \) receives for not announcing \( HH \). This shows that, for all \( j \in \{ L, H \} \), the distortions induced in the type\(-Hj\) production levels (and, implicitly, in \( q_{Hj} \)) serve the purpose of removing the type\(-Lj\) incentives to overstate \( \theta \) at lowest cost, which requires that, as at FB, the expected marginal surplus be equal across \( Hj \)-types i.e.,

\[
S'(q_{HL}^{sb}) + S'(\overline{q}_{HL}^{sb}) = S'(q_{HH}^{sb}) + S'(\overline{q}_{HH}^{sb}).
\]  (20)

Next suppose that type \( LH \) has no incentive to report \( HH \) so that only the type\(-HL \) and the type\(-LL \) production levels are distorted. Then, with analogous reasoning, for all \( i \in \{ L, H \} \), the distortions induced in the type\(-iL \) quantities (and, implicitly, in \( r_{iL} \)) serve the purpose of removing the type\(-iH \) incentives to understate \( \sigma \) at lowest cost, which requires that, again as at FB, the expected difference in marginal surplus be equal across \( iL \)-types i.e.,

\[
S'(q_{HL}^{sb}) - S'(q_{HH}^{sb}) = S'(q_{LL}^{sb}) - S'(q_{LL}^{sb}).
\]  (21)

Using (20) and (21), one can clarify why \( P \) induces distortions for types \( HL, HH \) and \( LL \) when the marginal surplus function is (close to) linear. If distortions were to concern types \( HL \) and \( HH \) only, then it would be optimal to set production levels so that both types provide the same output in expectation \( (q_{HH} = q_{HL}) \). However, in that case, the rent designed for type \( LH \) would not remove his incentives to report \( LL \), unless \( q_{HH} \) were raised above \( q_{HL} \). This would require that more important distortions be induced for type \( HL \) than for type \( HH \) so that the efficiency\/rent-extraction trade-off described above would no longer be optimally solved. On the other hand, if distortions were to concern types \( HL \) and \( LL \) only, then with \( S'(\cdot) \) linear it would be optimal to set output levels so that the difference between high and low productions is the same for both types in expectation \( (r_{LL} = r_{HL}) \). Yet, in that case, the rent designed for type \( LH \) would not discourage the \( LH \)-agent from mimicking type \( HH \), unless \( r_{LL} \) were increased above \( r_{HL} \) (and thus distorted less than) \( r_{HL} \). Therefore, as long as inefficiencies are induced for only two types when \( S'(\cdot) \) is linear, \( P \) is turned between two costly options, namely distorting \( q_{HH} \) less, relative to \( q_{HL} \), and distorting \( r_{LL} \) less, relative to \( r_{HL} \). To contain the cost, \( P \) seeks to decrease the type\(-LH \) rent by keeping \( q_{HH} \) and \( r_{LL} \) as close as possible to \( q_{HL} \) and \( r_{HL} \), respectively. The least costly strategy is to design an equal rent that prevents type \( LH \) from mimicking both type \( HH \) and type \( LL \).
As $S'(\cdot)$ becomes sufficiently convex, this strategy might no longer be appropriate. Type $LH$ may be tempted to report either $HH$ or $LL$, rather than being indifferent between the two lies, which removes the trade-off between the candidate type—$LH$ rents. However, it is not always clear which situation will actually arise, as the following corollary states.

**Corollary 2** Condition (19) is not necessarily satisfied for intermediary degrees of convexity of $S'(\cdot)$ for which (15) holds. Suppose it is not, indeed. Then, $\Pi_{LH,1} > \Pi_{LH,2}$ if and only if $\frac{\partial \phi}{\partial \sigma} > \delta'$, for some given $\delta' > 1$. Moreover, $\Pi_{LH,1} < \Pi_{LH,2}$ if and only if $S'(\cdot)$ is sufficiently convex and $\frac{1}{\mu} < \frac{\partial \phi}{\partial \sigma} < \frac{1}{\nu}$. If there exists some degree of convexity of $S'(\cdot)$ for which $\Pi_{LH,1} > \Pi_{LH,2}$ is feasible, then $\Pi_{LH,1} < \Pi_{LH,2}$ is feasible only for higher degrees of convexity of $S'(\cdot)$. The optimal mechanism has the same characteristics as in Proposition 2 except that, whenever $\Pi_{LH,1} > \Pi_{LH,2}$ ($\Pi_{LH,1} < \Pi_{LH,2}$), type $LH$ is assigned the rent $\Pi_{LH,1}^b$ ($\Pi_{LH,2}^b$) and type $LL$ (type $HH$) the FB production levels.

Keeping $q_{HH}$ and $r_{LL}$ close to $q_{HL}$ and $r_{HL}$, respectively, might no longer be optimal as $S'(\cdot)$ becomes more convex. To see why this is the case, notice that, in the good state, type $HH$ and type $LL$ are both more efficient than is type $HL$.\(^{13}\) Hence, production levels are fixed such that $\frac{q_{HH}^b}{q_{HL}^b}$ and $\frac{q_{LL}^b}{q_{HL}^b}$, which shows that the optimal values of the differences $q_{HH} - q_{HL}$ and $r_{LL} - r_{HL}$ raise with the convexity of $S'(\cdot)$. Remarkably, if (19) is to be maintained, then the difference $r_{LL} - r_{HL}$ is to be enlarged more than is to be the difference $q_{HH} - q_{HL}$, for increasing degrees of convexity of $S'(\cdot)$. However, as $S'(\cdot)$ gets sufficiently convex, (19) no longer needs to hold at SB. For instance, we might rather have $\Delta \theta(q_{HH}^b - q_{HL}^b) > \Delta \sigma(r_{LL}^b - r_{HL}^b)$. When this is the case, for some convex $S'(\cdot)$, decreasing the wedge between the type—$HH$ and the type—$HL$ expected productions is especially costly on efficiency ground. Then, $P$ prefers to leave a higher rent to type $LH$ for not mimicking type $HH$ at the first stage, rather than inducing more important distortions in the type—$HH$ quantities at the second stage. For even more convex $S'(\cdot)$, contrariwise, $\Delta \theta(q_{HH}^b - q_{HL}^b) < \Delta \sigma(r_{LL}^b - r_{HL}^b)$. This means that $P$ is now better off conceding a larger rent to type $LH$ for not announcing $LL$, rather than inducing big distortions in the type—$LL$ production levels as to lower the wedge between expected differences.

Lastly observe that neither of the two possible situations in which one rent dominates the other arises, unless type $LL$ has incentives to mimic type $HH$, as it is the case when (15) holds.

**Proposition 3** Assume that $S'(\cdot)$ is sufficiently convex, in which case condition (15) is not satisfied. The following two situations can arise:

\(^{13}\)Indeed, $\theta_H - \sigma_H < \theta_H - \sigma_L$ together with $\theta_L - \sigma_L < \theta_H - \sigma_L$. 

18
If $S^0(\cdot)$ is not very convex, then $r_{HH}^b > r_{LL}^b$.
Assume $\frac{\Delta \theta}{\Delta \sigma} < \mu + \frac{1-\mu}{\nu \mu}$. The optimal mechanism is characterized as follows:

- rents are given by $\Pi_{HL}^*, \Pi_{HH}^b, \Pi_{LL,2}^b$ and $\Pi_{LH,1}^b$;
- production levels are distorted for types HL and HH such that $r_{HL}^b < r_{HL}^*$ and $q_{HH}^b < q_{HH}^*$.

If $S^0(\cdot)$ is very convex, then $r_{HH}^b \leq r_{LL}^b$.
Assume $\frac{\Delta \theta}{\Delta \sigma} < 1 + \frac{1-\mu}{\nu \mu}$. The optimal mechanism has the same characteristics as above, except that (i) the type-LH rent is $\Pi_{LH,2}^b$ and (ii) production levels are distorted also for type LL such that $r_{LL}^b < r_{LL}^*$.

With $S^0(\cdot)$ sufficiently convex, type LL could gain from cheating on both dimensions of private information. To induce truthtelling, P is forced to concede a higher rent (now $\Pi_{LL,2}^b$) than in previous situations. With (15) not holding, it is straightforward to see that the difference between the type–HH and the type–HL expected quantities is now to be set very large at optimum. As distorting quantities away from the FB levels at the second stage is increasingly more costly, P abandons even more surplus to prevent the type–LL agent from reporting HH. Moreover, for $S^0(\cdot)$ very convex, it is harder to discourage type LH from mimicking type LL than type HH. This requires that a higher rent be given up for not reporting LL, which calls for fixing $r_{LL}^b \geq r_{HH}^b$.

In line with the content of Corollary 2, again it appears that, as $S^0(\cdot)$ becomes more convex, on the one hand, types LL and LH display increasingly stronger incentives to pretend to be less efficient and, on the other, it becomes more costly to distort quantities for decreasing the rents that must be conceded to remove such incentives at the first stage.

In presenting our results so far, we have classified cases according to the curvature of the principal’s marginal surplus function. This has helped us emphasize how the design of the optimal mechanism varies with the principal’s preferences. Besides, we have suggested that the structure of the incentive scheme also depends upon the magnitude of the spread index, which reflects the relative importance of the two first-stage information problems the principal tackles. The following corollary tells more on this latter aspect.

**Corollary 3** Assume that $\frac{\Delta \theta}{\Delta \sigma}$ is sufficiently large to violate any of the conditions stated in Proposition 1 to 3 and in Corollary 2. The optimal mechanism is analogous to that presented in Proposition 3, except that: (i) $q_{HH}^b = q_{HL}^*$ and $r_{HH}^b = r_{HL}^b$ (so that also $\Pi_{LL,1}^b = \Pi_{LL,2}^b$); (ii) $r_{HH}^b > r_{LL}^b$ if and only if $S^0(\cdot)$ is sufficiently concave. Under (ii), if $S^0(\cdot)$ is extremely concave, then $r_{LL}^b = r_{HH}^b$ and the type–LH production levels are distorted such that $r_{LL}^b < r_{LL}^*$. This corollary states that, when the information problem about the expected unit cost is relatively important (i.e., $\Delta \theta$ is sufficiently large with respect to $\Delta \sigma$), the least
costly way to remove the incentives of type $LL$ to report either $HL$ or $HH$ is to pool the production level designed for type $HH$ with that designed for type $HL$ in each possible state.

To illustrate the content of the corollary, first suppose that type $LL$ is solely attracted by the type—$HL$ contract. Then, in preventing that $HH$ be announced untruthfully, $P$ is to be concerned with the first dimension of private information only and so with the magnitude of $\Delta \theta q_{HH}$. On the other hand, because not only type $LL$ but also type $HH$ may be attracted by the type—$HL$ contract, when $HL$ is reported $P$ is to worry about possible cheating on both information dimensions and thus about the magnitude of both $\Delta \theta q_{HL}$ and $\Delta \sigma r_{HL}$. Taking this into account, $P$ fixes $q_{sb}^{HL} > q_{sb}^{HH}$, i.e., in the good state, she preserves the FB ranking between the type—$HH$ and the type—$HL$ output levels. By contrast, in the bad state, the FB ranking persists (i.e., $q_{sb}^{HH} > q_{sb}^{HL}$) if and only if the spread index is sufficiently small because raising the type—$HL$ production has opposite effects on $\Delta \theta q_{HL}$ and $\Delta \sigma r_{HL}$, as already seen. This is the situation described in Proposition 1 and 2 as well as in Corollary 1 and 2. Contrariwise, with $\frac{\Delta \theta}{\Delta \sigma}$ sufficiently large, it would be beneficial to decrease $q_{HL}$ below $q_{HH}$ (i.e., to set $q_{HH} > q_{HL}$) to contain the rent. However, by doing so, $P$ would be unable to solicit information release from type $LL$, independently of the characteristics of the marginal surplus function ((15) would be violated). The best strategy for $P$ is to make the type—$HH$ contract as attractive as the type—$HL$ contract and set the type—$LL$ rent accordingly. This requires that for types $HL$ and $HH$ production levels be pooled in either state (i.e., $q_{sb}^{HH} = q_{sb}^{HL}$ and $q_{sb}^{HH} = q_{sb}^{HL}$ so that $q_{GH}^{HH} = q_{GH}^{HL}$ and $r_{GH}^{HH} = r_{GH}^{HL}$ as well).

Next suppose that type $LL$ is solely tempted to report $HH$, which is the case described in Proposition 3. Then, when faced with the report $HH$, $P$ is to worry about possible lying on either dimension of private information. When faced with the report $HL$, instead, she is to care about possible lying on $\sigma$ only. While the expected rent to be given up is larger the bigger $\Delta \theta q_{HH}$, $\Delta \sigma r_{HH}$, and $\Delta \sigma r_{HL}$, here it does not depend upon $\Delta \theta q_{HL}$. This enables $P$ to raise the type—$HL$ production level in the bad state so that $q_{HL}^{HH} > q_{HL}^{HH}$. Although both such quantities are lowered below the FB levels, the FB ranking is here preserved. Indeed, a more important distortion is induced for type $HH$, provided the type—$HH$ contract is appealing for both types $LL$ and $LH$, whereas the type—$HL$ contract attracts type $HH$ only.

In definitive, the optimal balance between distortions depends upon how costly it is for $P$ to remove the incentives to lie on both dimensions of first-stage information, relative to removing the incentives to lie on one sole such dimension. From Proposition 3 we learn that, as long as the spread index is low, the temptation to cheat on $\sigma$ is relatively more concerning for $P$ and she sets $q_{sb}^{HH} > q_{sb}^{HL}$. Corollary 3 highlights that the converse occurs as the spread index becomes sufficiently large and type $LL$ is attracted by the contract designed for the least efficient type. Then, as above, pooling the production levels of types
and \( HH \) is the least costly strategy that \( P \) can adopt to discourage the \( LL \)--agent from reporting either such type.

Noticeably, the solution in Corollary 3 displays similarities with the optimal mechanisms characterized in Riordan and Sappington [15] and in Courty and Li [8]. We found that, in both possible states, production is fixed at the FB level for a low-expected-cost agent and distorted for a high-expected-cost agent, whatever the cost variability (the \( \sigma \)--pooling result). In the models aforementioned, a similar no-distortion-at-the-top result emerges. For instance, in Riordan and Sappington [15], no distortion is associated with the highest franchise valuation (\( i.e. \), the lowest expected production cost), whatever the second-stage cost realization.\(^{14} \) The outcome in Corollary 3 exactly replicates those of the authors aforementioned in the limit case in which \( \Delta \sigma \) approaches zero. Actually, in that case, our whole model degenerates into one in which the expected cost is privately known to the agent, whereas the cost variability is publicly observable, just as in Riordan and Sappington [15] and in Courty and Li [8].

5 Concluding remarks

We considered a principal-agent relationship in which the agent holds two pieces of private information, namely expected value and variability of the unit production cost, at the outset of the relationship with the principal and learns privately an additional piece of information, namely the realized unit cost, at a later stage. This is the case in a variety of real-world situations (frequently, in regulatory and procurements contexts) in which the cost of performing the concerned activity is uncertain when the contractual relationship begins and the activity is not executed until after uncertainty is solved. We emphasized that, in these settings, the principal tackles an information problem that is multidimensional and sequential at once and characterized the screening mechanism that optimally trades off rent-extraction against expected allocation efficiency.

We found that contractual features depend finely on two elements that are different in nature. One is given by the relative extent of the two first-stage information imperfections, as measured by the spread index, and thus pertains to the knowledge imprecision of the principal about the production environment. The other is given by the shape of the principal’s marginal surplus function and thus pertains to an intrinsic attribute of the principal \( i.e. \), her preferences for the good.

The screening problem that the principal tackles is especially complex when the spread index is low, in which case first-stage cheating is potentially attractive for the agent on either dimension of private information. Because of this, the problem displays similarities with such simultaneous two-dimensional screening problems as that analyzed by Arm-

\(^{14} \)Compare Corollary 3 in Riordan and Sappington [15].
strong and Rochet [2]. Specifically, at the first stage, the principal faces an equally rich set of incentive constraints (provided information is indeed disclosed at the second stage) and has two quantity-related screening devices at hand. However, availability of two such devices does not follow from having the agent execute two distinct activities, as it is the case in Armstrong and Rochet [2]. Rather, it rests on the circumstance that those devices depend strictly upon the two final production levels, which must comply with a second-stage incentive-compatibility requirement. This casts more restrictions on the principal’s strategies so that more combinations of incentive constraints can be binding in our framework. Actually, the first-stage constraints that are binding in our model can be viewed as the counterpart of those in Armstrong and Rochet [2] limitedly to the situations in which the marginal surplus function is not too convex (described in Proposition 1 and 2 and in Corollary 1 and 2). A different subset of constraints comes to matter when the function becomes very convex (Proposition 3), in which case an agent forecasting low mean and low variability of the cost must obtain more surplus for correctly announcing either dimension of private information.

The problem that the principal tackles comes closer to a standard sequential screening problem as the spread index becomes sufficiently large, meaning that the information issue about the cost variability gets sufficiently weak in relative terms. Actually, under these circumstances, the main preoccupation of the principal is to prevent first-stage cheating on the expected cost. It looks thus natural that the optimal mechanism displays similarities with the ones characterized by Riordan and Sappington [15] and Courty and Li [8] with regards to contexts where the agent holds one sole piece of private knowledge at the beginning of the contractual relationship. Nevertheless, the sole case in which the mechanism that we characterized would replicate those in previous sequential screening models, is the limit case in which the principal’s imperfection about the second piece of first-stage information were to vanish. Noticeably, this is also the sole situation in which, like in those models, the curvature of the marginal surplus function would have no impact on the specific combination of binding first-stage incentive constraints. This points to the conclusion that the principal’s preferences come to play a role in shaping the optimal mechanism precisely due to the multidimensional nature of the information problem that is nested at the first stage of the sequential screening procedure.

Two last points are worth making. First, in several studies about simultaneous multi-dimensional screening, it emerges that the degree of correlation between different pieces of private information affects the design of the optimal mechanism. In our model, we assumed away any correlation between expected cost and cost variability. This approach helped us disentangle the effects of the multidimensional nature of private information in the sequential setting we considered. Second, and perhaps more interestingly, Courty and Li [8] suggest through an example that, in situations where an agent forecasting low (high) cost faces high (low) cost variability, there could be a gain in switching from a
deterministic to a random mechanism. In our work, we only focused on deterministic mechanisms as a first step in the study of contract design in the novel framework that we considered. Exploring random mechanisms would be a natural continuation of this study.

References


A Incentive constraints

A.1 Incentive constraints omitted in (5)

Assume that, at the first stage, type $ij$ reports $i'j'$. Conditional on this report, his second-stage incentive constraints are written

$$
\tilde{\pi}_{i'j'} \geq \tilde{\pi}_{i'j'} + 2\sigma_jq_{i'j'} \\
\tilde{\pi}_{i'j'} \geq \tilde{\pi}_{i'j'} - 2\sigma_jq_{i'j'},
$$

(22) (23)

$\tilde{\pi}_{i'j'}$ and $\tilde{\pi}_{i'j'}$ being the corresponding profits in the good and bad state, respectively. Further denote the right-hand side of (5) as

$$
X \equiv \frac{1}{2} \left\{ [\ell_{i'j'} - (\theta_i + \sigma_j)q_{i'j'}] + [\ell_{i'j'} - (\theta_i - \sigma_j)q_{i'j'}] \right\}.
$$

Type $ij$ might want to declare $i'j'$ at the first stage, anticipating that, at the second stage, the following three alternatives will be available (apart from truth-telling whatever the realization of the shock): (1) report a positive shock, whatever the realization; (2) report a negative shock, whatever the realization; (3) report a positive (negative) shock when the true shock is negative (positive). Accordingly, the three first-stage incentive constraints are written as follows:

$$
\Pi_{ij} \geq X + \frac{1}{2}(\tilde{\pi}_{i'j'} - \tilde{\pi}_{i'j'} - 2\sigma_jq_{i'j'}) \\
\Pi_{ij} \geq X + \frac{1}{2}(\tilde{\pi}_{i'j'} + 2\sigma_jq_{i'j'} - \tilde{\pi}_{i'j'}) \\
\Pi_{ij} \geq X + \frac{1}{2}(\tilde{\pi}_{i'j'} - \tilde{\pi}_{i'j'} - 2\sigma_jq_{i'j'}) + \frac{1}{2}(\tilde{\pi}_{i'j'} + 2\sigma_jq_{i'j'} - \tilde{\pi}_{i'j'})
$$

These constraints are all implied by (5), (22) and (23).

A.2 Incentive constraints and rents

Using the definitions of $q_{ij}$ and $r_{ij}$ in the text, (3) is rewritten as $\Pi_{ij} = \frac{1}{2}(\ell_{ij} + \ell_{ij}) - (\theta_iq_{ij} - \sigma_jr_{ij})$. Based on this expression, we can reformulate the incentive constraints in
(5) for all possible types as follows:

\[
\begin{align*}
\Pi_{LL} &\geq \Pi_{HL} + \Delta \theta q_{HL} \quad \text{(IC1)} \\
\Pi_{LL} &\geq \Pi_{LH} - \Delta \sigma r_{LH} \quad \text{(IC2)} \\
\Pi_{LL} &\geq \Pi_{HH} + \Delta \theta q_{HH} - \Delta \sigma r_{HH} \quad \text{(IC3)} \\
\Pi_{HL} &\geq \Pi_{LL} - \Delta \theta q_{LL} \quad \text{(IC4)} \\
\Pi_{HL} &\geq \Pi_{HH} - \Delta \sigma r_{HH} \quad \text{(IC5)} \\
\Pi_{HL} &\geq \Pi_{LH} - \Delta \theta q_{LH} - \Delta \sigma r_{LH} \quad \text{(IC6)} \\
\Pi_{LH} &\geq \Pi_{HH} + \Delta \theta q_{HH} \quad \text{(IC7)} \\
\Pi_{LH} &\geq \Pi_{LL} + \Delta \sigma r_{LL} \quad \text{(IC8)} \\
\Pi_{LH} &\geq \Pi_{HL} + \Delta \theta q_{HL} + \Delta \sigma r_{HL} \quad \text{(IC9)} \\
\Pi_{HH} &\geq \Pi_{LH} - \Delta \theta q_{LH} \quad \text{(IC10)} \\
\Pi_{HH} &\geq \Pi_{HL} + \Delta \sigma r_{HL} \quad \text{(IC11)} \\
\Pi_{HH} &\geq \Pi_{LL} - \Delta \theta q_{LL} + \Delta \sigma r_{LL}. \quad \text{(IC12)}
\end{align*}
\]

Downward incentive constraints are \((\text{IC1}), (\text{IC3}), (\text{IC7})\) to \((\text{IC9})\) and \((\text{IC11})\). Assuming that all other incentive constraints are slack and that \(\Pi_{HL} = 0\), we obtain the rents in \((\text{12a})\) to \((\text{12d})\). With the latter, the participation constraints in \((\text{8})\) are trivially satisfied for all types but \(HL\).

**B. Proof of Proposition 1**

We begin by solving the reduced problem \(\Gamma'\), to be presented below. We subsequently find the conditions under which \(q_{HL} > q_{HH}\) and \(r_{HL} > r_{LL}\). We then identify the situations in which, under the assumptions of the proposition, the solution to \(\Gamma'\) satisfies all constraints in the general problem \(\Gamma\) as well as the exceptional case in which the solution to \(\Gamma'\) does not solve \(\Gamma\). We lastly identify the distortions at the solution.

**B.1 The reduced problem \(\Gamma'\)**

Let \(\Gamma'\) denote the reduced problem in which binding first-stage incentive constraints are those leading to the rents \((\text{12a}), (\text{12b}), \Pi_{LL,1}\) in \((\text{12c})\) and \((\text{12d})\). We rewrite \((\text{12d})\) as

\[
\Pi_{LH} = \gamma_1 (\Delta \theta q_{HH} + \Delta \sigma r_{HL}) + \gamma_2 (\Delta \theta q_{HL} + \Delta \sigma r_{LH}) + \gamma_3 (\Delta \theta q_{HL} + \Delta \sigma r_{HL}),
\]

with \(0 \leq \gamma_k \leq 1\), \(k = 1, 2, 3\), and \(\gamma_1 + \gamma_2 + \gamma_3 = 1\). Embodying the rents above, the objective function becomes

\[
\sum_{i,j \in \mathcal{T}} E_{ij}[V_{ij}] - \nu \mu \Pi_{LL,1} - (1 - \nu) (1 - \mu) \Pi_{HH} - \nu (1 - \mu) \Pi_{LH}.
\]

25
Optimizing with respect to quantities yields \( q_{2LH}^* \) and \( q_{1LH}^* \) for type \( LH \), together with

\[
S'(q_{HL}) = \theta_H - \sigma_L + \frac{\nu}{1 - \nu} \mu (1 - \mu) (\gamma_2 + \gamma_3) \Delta \theta + \frac{1 - \mu}{\mu} \frac{1 - \nu + \nu (\gamma_1 + \gamma_3)}{1 - \nu} \Delta \sigma
\]

\[
S'(\overline{q}_{HL}) = \theta_H + \sigma_L + \frac{\nu}{1 - \nu} \mu (1 - \mu) (\gamma_2 + \gamma_3) \Delta \theta - \frac{1 - \mu}{\mu} \frac{1 - \nu + \nu (\gamma_1 + \gamma_3)}{1 - \nu} \Delta \sigma
\]

for type \( HL \)

\[
S'(q_{LL}) = \theta_L - \sigma_L + \gamma_2 \frac{1 - \mu}{\mu} \Delta \sigma
\]

\[
S'(\overline{q}_{LL}) = \theta_L + \sigma_L - \gamma_2 \frac{1 - \mu}{\mu} \Delta \sigma
\]

for type \( LL \) and

\[
S'(q_{HH}) = \theta_H - \sigma_H + \gamma_1 \frac{\nu}{1 - \nu} \Delta \theta
\]

\[
S'(\overline{q}_{HH}) = \theta_H + \sigma_H + \gamma_1 \frac{\nu}{1 - \nu} \Delta \theta
\]

for type \( HH \). Both in the quantity solution here above and throughout the rest of the proof, the superscript \( sb \) is omitted.

**B.2 The conditions under which \( q_{HL} > q_{HH} \) and \( r_{HL} > r_{LL} \)**

Using (24), conditions \( q_{HL} > q_{HH} \) and \( r_{HL} > r_{LL} \) are found to be jointly equivalent to \( \gamma_3 = 1 \). We replace \( \gamma_3 = 1 \) into the solution. We find \( q_{HH} > q_{HL} \) and, because \( \frac{\Delta \sigma}{\Delta \theta} < \frac{1 - \nu}{\nu} \), \( \overline{q}_{HL} > \overline{q}_{HH} \). We calculate

\[
[S'(q_{HL}) - S'(q_{HH})] - [S'(\overline{q}_{HH}) - S'(\overline{q}_{HL})] = \frac{\nu}{(1 - \nu) \mu} \Delta \theta,
\]

from which we deduce that \( \overline{q}_{HL} - \overline{q}_{HH} > q_{HH} - q_{HL} \) (or, equivalently, \( q_{HL} > q_{HH} \)) if and only if \( S'(\cdot) \) is sufficiently concave. Furthermore, \( q_{HL} > q_{HH} \) and, because \( \frac{\Delta \sigma}{\Delta \theta} > \frac{1 - \mu}{\mu + \nu(1 - \mu)} \), \( \overline{q}_{HL} > \overline{q}_{HH} \). We calculate

\[
[S'(q_{HL}) - S'(q_{HH})] - [S'(\overline{q}_{HL}) - S'(\overline{q}_{HH})] = \frac{1 - \mu}{(1 - \nu) \mu} \Delta \sigma,
\]

from which we deduce that \( \overline{q}_{HL} - \overline{q}_{HH} > q_{HH} - q_{HL} \) (or, equivalently, \( r_{HL} > r_{LL} \)) if and only if \( S'(\cdot) \) is sufficiently concave.

**B.3 Check incentive constraints in \( \Gamma \)**

(Ic1) is binding and (Ic3) is slack because (15) is satisfied (as proved in the main text). (Ic2) is equivalent to \( r_{HL} \geq r_{HH} \). (Ic4) is equivalent to \( q_{LL} \geq q_{HL} \). (Ic5) to
\( r_{HH} \geq r_{HL}, \) (IC6) to \( \Delta \theta q_{LH} + \Delta \sigma r_{LH} \geq \Delta \theta q_{HL} + \Delta \sigma r_{HL}, \) (IC7) to \( q_{HL} \geq q_{HH}, \) (IC8) to \( r_{HL} \geq r_{LL}, \) (IC10) to \( q_{LL} \geq q_{HL} \) and (IC12) to \( \Delta \theta q_{LL} + \Delta \sigma r_{LL} \geq \Delta \theta q_{HL} + \Delta \sigma r_{HL}. \) (IC2) and (IC10) imply (IC6). (IC4) and (IC8) imply (IC1). Moreover, (IC7) and (IC8) hold because \( q_{HL} > q_{HH} \) and \( r_{HL} > r_{LL}. \)

Check (IC2). We have \( q_{LH} > q_{HL}. \) Suppose \( \frac{\Delta \theta}{\Delta \sigma} > \frac{1 - \nu}{\nu + \mu (1 - \nu)} \). Equivalently, \( \bar{q}_{LH} > \bar{q}_{HL}. \) We calculate

\[
[S'(q_{LH}) - S'(\bar{q}_{LH})] - [S'(q_{HL}) - S'(\bar{q}_{HL})] = 2[1 + \frac{1 - \mu}{\mu (1 - \nu)}] \Delta \sigma,
\]

from which we deduce that (IC2) holds if and only if \( S'(\cdot) \) is not very concave. Moreover, from the solution to \( \Gamma' \), \( r_{LH} \geq r_{LL}. \) Hence, if \( S'(\cdot) \) is sufficiently concave for the condition \( r_{HL} \geq r_{LL} \) (as stated in the proposition) to hold, then it is not necessarily the case that \( r_{HL} \geq r_{LL}. \) Suppose \( \frac{\Delta \theta}{\Delta \sigma} < \frac{1 - \nu}{\nu + \mu (1 - \nu)} \) and so \( \bar{q}_{LH} < \bar{q}_{HL}. \) Then, \( r_{LH} > r_{HL} \) whenever the shape of \( S'(\cdot). \)

Check (IC4) and (IC5). They are both satisfied, the former because \( \bar{q}_{LL} > q_{HL} \) and \( \bar{q}_{HL} > q_{HL}, \) the latter because both \( q_{LL} > q_{HL} \) and \( \bar{q}_{HL} > \bar{q}_{HL}. \)

Check (IC10). From the above check of (IC2), \( q_{LH} > q_{HL}. \) Hence, (IC10) is satisfied whenever \( \frac{\Delta \theta}{\Delta \sigma} > \frac{1 - \nu}{\nu + \mu (1 - \nu)} \). Assume \( \frac{\Delta \theta}{\Delta \sigma} \leq \frac{1 - \nu}{\nu + \mu - \nu} \) and calculate

\[
[S'(q_{LH}) - S'(\bar{q}_{LH})] - [S'(\bar{q}_{HL}) - S'(\bar{q}_{HL})] = \frac{2}{1 - \nu} \left( 1 + \frac{\nu}{\mu} \right) \Delta \theta.
\]

We deduce that \( q_{LH} > q_{HL} > \bar{q}_{HL} > \bar{q}_{HL} \) so that (IC10) is satisfied if and only if \( S'(\cdot) \) is not too concave. Moreover, \( q_{LH} > q_{HH} \) as we assumed that \( \gamma_3 = 1. \)

Hence, there exist some intermediary degrees of concavity of \( S'(\cdot) \) for which both \( q_{HL} \geq q_{HH} \) and \( q_{HL} \geq q_{HL} \) at the solution to \( \Gamma'. \) When \( S'(\cdot) \) is very concave, only the condition \( q_{HL} \geq q_{HH} \) is satisfied.

\subsection*{B.3.1 When the solution to \( \Gamma' \) violates incentive constraints in \( \Gamma \)}

Take \( \frac{\Delta \theta}{\Delta \sigma} \leq \frac{1 - \nu}{\nu + \mu - \nu} \) and \( S'(\cdot) \) very concave such that, as from the previous proof, the solution to \( \Gamma' \) violates (IC10). Thus, (IC10) is binding at the solution. (IC11) might be slack instead. The best guess is that the binding constraints are the same as in \( \Gamma', \) except for (IC10) and (IC11).

Check (IC11). It is rewritten as \( q_{HL} \geq q_{HL}. \) Suppose it is slack. Then, the agent’s expected rent is given by

\[
[\nu + (1 - \nu) (1 - \mu)] \Delta \theta q_{HL} + (1 - \mu) \Delta \sigma r_{HL} - (1 - \nu) (1 - \mu) \Delta \theta q_{HL}.
\]

Embodying this expression into \( P' \)’s objective function and calculating the first-order condition with respect to the type-\( LH \) and to the type-\( HL \) quantities, we find that \( q_{LH} > q_{HL} \) and \( \bar{q}_{LH} > \bar{q}_{HL} \) so that \( q_{LH} > q_{HL}. \) This contradicts the assumption that
(IC11) is slack. Hence, whenever (IC10) is binding, (IC11) is binding too and \( q_{HL} = q_{LH} \).

Next take \( \frac{\Delta \theta}{\Delta \sigma} > \frac{1-\mu}{\mu+\mu-\nu} \) and \( S'(\cdot) \) very concave, such that the solution to \( \Gamma' \) violates (IC2). Take (IC2) to be binding. The best guess is that the other binding constraints are the same as before with the sole possible exception of (IC1). This is equivalent to \( r_{HL} \geq r_{LH} \) and might be slack. Taking (IC1) to be slack, the quantity solution leads again to a contradiction. Hence, \( r_{HL} = r_{LH} \).

B.4 Distortions

Take \( S'(\cdot) \) to be not very concave, such that both \( r_{LH} > r_{HL} \) and \( q_{LH} > q_{HL} \). At the solution to \( \Gamma' \) with \( \gamma_3 = 1 \), we obtain \( q^*_{LH}, r^*_{LH} \) together with the quantity distortions stated in the proposition. Next take \( S'(\cdot) \) to be very concave and \( q_{HL} = q_{LH} \). Under this equality constraint, we have \( q_{LH} < q^*_{LH} \). Lastly take \( S'(\cdot) \) to be very concave and \( r_{HL} = r_{LH} \). Under this equality constraint, we have \( r_{LH} < r^*_{LH} \).

C Proof of Corollary 1

Take \( \gamma_3 = 1. \) With \( \frac{\Delta \theta}{\Delta \sigma} \leq \frac{1-\mu}{\mu+\nu(1-\mu)} \) and given the proof of Proposition 1, we have \( \bar{q}_{LL} \leq \bar{q}_{HL} \). Because \( q^*_{LL} > q^*_{HL} \), we also have \( r_{LL} > r_{HL} \), which contradicts the hypothesis that \( \gamma_3 = 1 \). Recall the assumption that \( q_{HL} > q_{HH} \), meaning that \( \gamma_1 = 0 \). Hence, if \( \gamma_3 = 0 \) then \( \gamma_2 = 1 \). However, the proof of Proposition 3 below shows that, with \( S'(\cdot) \) concave, \( \gamma_2 \neq 1 \). Hence, \( 0 < \gamma_3 < 1 \) and \( 0 < \gamma_2 < 1 \), in which case \( \Pi_{LH,2}^{sh} = \Pi_{LH,3}^{sh} \). The condition \( r_{LL} < r^*_{LL} \) is found from the quantity solution to \( \Gamma' \) with \( \gamma_1 = 0 \).

Now check when \( q_{HL} > q_{HH} \). Replacing \( \gamma_1 = 0 \) into the quantity solution to \( \Gamma' \), we see that \( q^*_{HH} > q^*_{HL} \) together with \( \bar{q}_{HL} > \bar{q}_{HH} \). We calculate

\[
[S'(q_{HL}) - S'(q^*_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{HL})] = 2 \frac{\nu}{(1-\nu)\mu} (\gamma_2 + \gamma_3) \Delta \theta,
\]

from which we deduce that \( q_{HL} > q_{HH} \) if and only if \( S'(\cdot) \) is sufficiently concave.

D Proof of Proposition 2

We learnt that \( q_{HH} \geq q_{HL} \) together with \( r_{LL} \geq r_{HL} \) as long as \( S'(\cdot) \) is not very concave (recall the proof of Proposition 1). This means that, in that case, \( \gamma_3 < 1 \). As the concavity of \( S'(\cdot) \) becomes less pronounced, the two conditions come to hold as strict inequalities so that \( \gamma_3 = 0 \). Moreover, in the proof of Proposition 3 below we show that neither \( \gamma_1 = 1 \) nor \( \gamma_2 = 1 \) unless \( S'(\cdot) \) is sufficiently convex. Hence, because \( \gamma_1 > 0 \) and \( \gamma_2 > 0 \), (19) holds. In what follows, we first find the values of \( \gamma_1 \) and \( \gamma_2 \) and then check that the remaining incentive constraints in \( \Gamma \) are satisfied.
D.1 The values of $\gamma_1$ and $\gamma_2$

Replacing $\gamma_3 = 0$ into the quantity solution of $\Gamma'$, we obtain

$$\begin{align*}
[S'(q_{HL}) - S'(\bar{q}_{LL})] - [S'(\bar{q}_{HL}) - S'(\bar{q}_{LL})] &= 2\frac{1 - \mu}{\mu}\left(1 - \gamma_2 + \frac{\nu}{1 - \nu}\gamma_1\right)\Delta\sigma,
[S'(q_{HL}) - S'(q_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{HL})] &= 2\frac{\nu}{1 - \nu}\left(1 - \gamma_1 + \frac{1 - \mu}{\mu}\gamma_2\right)\Delta\theta.
\end{align*}$$

Replacing $\gamma_2 = 1 - \gamma_1$ into the first equation and $\gamma_1 = 1 - \gamma_2$ into the second, we find the expressions in (17) and (18).

D.2 Check incentive constraints in $\Gamma$

(IC1) and (IC3) are satisfied (as before). (IC4) is rewritten as $q_{LL} \geq q_{HL}$, (IC5) as $r_{HH} \geq r_{HL}$ and (IC12) as $\Delta q_{LL} + \Delta r_{HL} \geq \Delta q_{HL} + \Delta r_{LL}$. (IC7) and (IC8) are here binding. Using (IC8) in (IC2), the latter becomes $r_{HL} \geq r_{LL}$. Using (IC8) in (IC6), the latter becomes $\Delta q_{LL} + \Delta r_{HL} \geq \Delta q_{HL} + \Delta r_{LL}$. Using (IC7) in (IC9), the latter becomes $q_{HH} \geq q_{HL}$ and thus holds. Using (IC7) in (IC10), the latter is rewritten as $q_{LL} \geq q_{HH}$. Combining (19) with (IC4), we see that (IC12) is satisfied. Using (IC9), (IC10) and (IC2) in (IC6), the latter is found to be satisfied as well.

We have $q_{LL} > q_{HL}$ and $\bar{q}_{LL} > \bar{q}_{HL}$ so that $r_{HL} \geq r_{LL}$. Hence, (IC2) is satisfied. Moreover, $q_{HH} > q_{HL}$ and, because $\Delta q < \frac{1}{\gamma_2}\left[\frac{1 - \nu}{\nu} + (1 - \mu)\gamma_1\right]\Delta\sigma$, $\bar{q}_{HL} > \bar{q}_{HH}$ so that $r_{HH} \geq r_{HL}$. Thus, (IC5) is satisfied. Also, both $q_{LL} > q_{HH}$ and $\bar{q}_{LL} > \bar{q}_{HH}$ so that $q_{HL} \geq q_{HH}$ and (IC10) is satisfied.

Check (IC4). We have $\bar{q}_{LL} > \bar{q}_{HH}$ and, as from the proof hereafter, $q_{LL} > q_{HL}$. It follows that $q_{LL} \geq q_{HH}$. Then, the condition $q_{HH} > q_{HL}$ implies (IC4), which is satisfied.

Proof of $q_{LL} > q_{HL}$

Suppose $q_{HH} > q_{HL}$. Start from a situation in which $0 < \gamma_1 < 1$ and $0 < \gamma_2 < 1$ i.e., (19) is satisfied. Assume that $S'(\cdot)$ becomes more convex. If $S'(q_{ij})$ and $S'(\bar{q}_{ij})$ remain unchanged for types $HL$, $LL$ and $HH$, then the differences $q_{HH} - q_{HL}$ and $r_{LL} - r_{HL}$ increase to $q_{HH} - q_{HL} + \Delta q$ and to $r_{LL} - r_{HL} + \Delta r$, respectively. To have (19) still satisfied, $S'(q_{ij})$ and $S'(\bar{q}_{ij})$ should be changed so that $\Delta q \Delta q = \Delta \sigma \Delta r$. Because $\Delta q > \Delta \sigma$, this requires that $\Delta q$ be set smaller than $\Delta r$, which in turn calls for $q_{LL} > q_{HH}$.

\(^{15}\)For the sake of clarity, we describe how $S'(q_{ij})$ and $S'(\bar{q}_{ij})$ change. First suppose that $\Delta q \Delta q > \Delta \sigma \Delta r$ prior to any change. Then, it is necessary that $\Delta q$ be decreased and/or $\Delta r$ increased for the equality to be established. This is made by decreasing the difference $[S'(q_{HL}) - S'(q_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{HL})]$ and/or increasing the difference $[S'(q_{HL}) - S'(q_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{HL})]$. As from (17) and (18), it means that $\gamma_2$ is decreased and $\gamma_1$ increased. Changes are to be made in the opposite direction if $\Delta q \Delta q < \Delta \sigma \Delta r$, instead.
E Proof of Corollary 2

We first show that, for $S'(\cdot)$ sufficiently convex but (15) satisfied, the solution might be such that either $\gamma_1 = 1$ or $\gamma_2 = 1$. We find the conditions under which $\gamma_1 = 1$ and those under which $\gamma_2 = 1$. In either case, we characterize the solution to $\Gamma'$ and show that it satisfies the constraints of problem $\Gamma$. We lastly compare the conditions under which the two cases are respectively feasible.

E.1 The conditions under which $\gamma_1 = 1$

Take (15) to hold. The necessary conditions under which $\gamma_1 = 1$ are

$$\Delta \theta (q_{HH} - q_{HL}) > \Delta \sigma (r_{LL} - r_{HL}).$$

(25)

Check $q_{HH} > q_{HL}$. From the solution to $\Gamma'$ with $\gamma_1 = 1$, we have

$$[S'(q_{HL}) - S'(q_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{HL})] = 0.$$  

(26)

Thus, $q_{HH} > q_{HL}$ if and only if $S'(\cdot)$ is strictly convex and $q_{HH} = q_{HL}$ with $S'(\cdot)$ linear.

Check $r_{LL} > r_{HL}$. From the solution to $\Gamma'$, we have $\bar{q}_{LL} > q_{HL}$. Take $\theta_{LL} < \bar{q}_{HL}$. Then, $r_{LL} > r_{HL}$ whatever the shape of $S'(\cdot)$. Take now $\bar{q}_{LL} > \bar{q}_{HL}$ and calculate

$$[S'(q_{HL}) - S'(q_{LL})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{LL})] = 2 \frac{1 - \mu}{(1 - \nu)\mu} \Delta \sigma.$$  

(27)

In this case, $r_{LL} > r_{HL}$ as long as $S'(\cdot)$ is not too concave.

Putting together the conditions for $q_{HH} > q_{HL}$ with those for $r_{LL} > r_{HL}$, we deduce that (25) is feasible only if $S'(\cdot)$ is strictly convex. Moreover, from (26) and (27), we deduce that the difference $q_{HH} - q_{HL}$ is independent of both $\Delta \theta$ and $\Delta \sigma$, whereas the difference $r_{LL} - r_{HL}$ is independent of $\Delta \theta$ and positively related to $\Delta \sigma$. Therefore, for some given degree of convexity of $S'(\cdot)$ for which (15) is satisfied, $\gamma_1 = 1$ if and only if $\frac{\Delta \theta}{\Delta \sigma}$ is sufficiently large. Lastly, (25) and (15) hold jointly only if $r_{HH} > r_{LL}$. Using $\bar{q}_{LL} > \bar{q}_{HH}$ and $q_{LL} > q_{HH}$, we calculate

$$[S'(q_{HH}) - S'(q_{LL})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{LL})] = -2 \Delta \sigma,$$

which means that $r_{HH} \geq r_{LL}$ if and only if $S'(\cdot)$ is not very convex. In conclusion, having both $\gamma_1 = 1$ and (15) satisfied is feasible for some $S'(\cdot)$ convex but not too convex and, at the same time, $\frac{\Delta \theta}{\Delta \sigma}$ sufficiently large.

E.1.1 Check incentive constraints in $\Gamma$

(IC1) and (IC3) are satisfied (as before). With $\gamma_1 = 1$ in $\Gamma'$, (IC2) is equivalent to $\Delta \theta q_{HL} + \Delta \sigma r_{LH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HL}$, (IC4) to $q_{LL} \geq q_{HL}$, (IC5) to $r_{HH} \geq r_{HL}$,
(IC6) to \( \Delta \theta q_{LL} + \Delta \sigma r_{LL} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HL} \), (IC10) to \( q_{LL} \geq q_{HH} \) and (IC12) to \( \Delta \theta q_{LL} + \Delta \sigma r_{HL} \geq \Delta \theta q_{HL} + \Delta \sigma r_{LL} \).

We have

\[
S'(q_{HH}) - S'(q_{LL}) = S'(\bar{q}_{HH}) - S'(\bar{q}_{LL})
\]

so that \( r_{LL} > r_{HH} \) because \( S'(\cdot) \) is convex. Together with (IC3), which is equivalent to \( \Delta \theta q_{HL} + \Delta \sigma r_{HH} > \Delta \theta q_{HH} + \Delta \sigma r_{HL} \), this implies that (IC2) is satisfied. Moreover,

\[
[S'(\bar{q}_{HL}) - S'(q_{LL})] - [S'(\bar{q}_{HH}) - S'(q_{HH})] = \frac{2}{1 - \nu} \Delta \theta.
\]

This says that \( q_{LL} \geq q_{HL} \) for \( S'(\cdot) \) convex, in which case (IC4) is satisfied. From (26), \( r_{HH} \geq r_{HL} \) and so (IC5) is satisfied. From (IC2), (IC10) and \( q_{HH} > q_{HL} \), (IC6) is satisfied. Also, \( \bar{q}_{LL} > q_{HH} \), together with \( \bar{q}_{HL} > \bar{q}_{HH} \) so that (IC10) is satisfied. Lastly, both \( \bar{q}_{LL} > \bar{q}_{HH} \) and \( \bar{q}_{HL} > q_{HH} \) so that \( q_{LL} > q_{HH} \). Together with (25), this implies that (IC12) is satisfied.

**E.2 The conditions under which \( \gamma_2 = 1 \)**

Take again (15) to hold. The conditions under which \( \gamma_2 = 1 \) are \( r_{LL} > r_{HL} \) and

\[
\Delta \sigma (r_{LL} - r_{HL}) > \Delta \theta (q_{HH} - q_{HL}). \tag{28}
\]

Check \( r_{LL} > r_{HL} \). We have

\[
S'(q_{HL}) - S'(q_{LL}) = S'(\bar{q}_{HL}) - S'(\bar{q}_{LL})
\]

so that \( r_{LL} > r_{HL} \) if and only if \( S'(\cdot) \) is strictly convex and \( r_{LL} = r_{HL} \), with \( S'(\cdot) \) linear.

Check \( q_{HH} - q_{HL} \). We have \( q_{HH} > q_{HL} \), and with \( \frac{\Delta \theta}{\Delta \sigma} < \frac{1 - \nu}{\nu} \) (as stated in the corollary), \( \bar{q}_{HL} > \bar{q}_{HH} \). We calculate

\[
[S'(q_{HL}) - S'(q_{HH})] - [S'(\bar{q}_{HH}) - S'(q_{HH})] = 2 \frac{\nu}{(1 - \nu) \mu} \Delta \theta,
\]

from which we deduce that \( q_{HH} > q_{HL} \) as long as \( S'(\cdot) \) is not very concave.

Putting together the circumstances under which \( r_{LL} > r_{HL} \) and those under which \( q_{HH} > q_{HL} \), we notice that (28) is satisfied only if \( S'(\cdot) \) is convex. Moreover, as \( S'(\cdot) \) becomes more convex, (28) is satisfied only if \( r_{LL} - r_{HL} \) increases faster, as compared to \( q_{HH} - q_{HL} \). This is the case if and only if \( q_{LL} > q_{HH} \). Because

\[
S'(q_{HH}) - S'(q_{LL}) = \Delta \theta - \frac{\Delta \sigma}{\mu},
\]

we have \( q_{LL} > q_{HH} \) if and only if \( \frac{\Delta \theta}{\Delta \sigma} > \frac{1}{\mu} \). Together with the condition \( \frac{\Delta \theta}{\Delta \sigma} < \frac{1 - \nu}{\nu} \), it is necessary that \( \frac{\Delta \theta}{\Delta \sigma} \in \left( \frac{1}{\mu}, \frac{1 - \nu}{\nu} \right) \). For any value of \( \frac{\Delta \theta}{\Delta \sigma} \) within this interval, there exists a
threshold on the convexity of $S' (·)$ above which (28) is satisfied. Lastly, (28) holds jointly with (15) only if $r_{LL} > r_{HH}$, which is the case if and only if $S' (·)$ is sufficiently convex.

**E.2.1 Check incentive constraints in $\Gamma'$**

(IC1) and (IC3) are satisfied (as before). With $\gamma_2 = 1$ in $\Gamma'$, (IC2) is equivalent to $r_{LH} \geq r_{LL}$, (IC4) to $q_{LL} \geq q_{HL}$, (IC5) to $r_{HH} \geq r_{LL}$, (IC6) to $\Delta \theta q_{HL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HL} + \Delta \sigma r_{LL}$, (IC10) to $\Delta \theta q_{HL} + \Delta \sigma r_{HL} \geq \Delta \theta q_{HL} + \Delta \sigma r_{LL}$ and (IC12) to $\Delta \theta q_{HL} + \Delta \sigma r_{HL} \geq \Delta \theta q_{HL} + \Delta \sigma r_{LL}$. From (IC2), $r_{LL} > r_{HL}$ and (IC10), (IC6) is satisfied. Moreover,

$$S'(q_{LL}) - S'(q_{HL}) = S'(q_{LH}) - S'(q_{LL}),$$

from which $q_{LL} > q_{HL}$. Together with (IC12), it implies that (IC10) is satisfied.

We check now the remaining constraints. We have $q_{LH} > q_{LL}$ and $q_{HH} > q_{HL}$ so that $r_{LH} > r_{LL}$. Hence, (IC2) is satisfied. We further have $q_{LL} > q_{HL}$ and $q_{LL} > q_{HH}$ so that (IC4) is satisfied as well. Moreover, $q_{HH} > q_{HL}$ and, because $\frac{\Delta \theta}{\Delta \sigma} < \frac{1 - \nu}{\nu}$, $q_{HL} > q_{HH}$ so that (IC5) is satisfied as well. (IC12) is equivalent to

$$(\Delta \theta + \Delta \sigma) (q_{LL} - q_{HL}) + (\Delta \theta - \Delta \sigma) (q_{HH} - q_{HL}) \geq 0,$$

which holds true because $q_{LL} > q_{HL}$ together with $q_{LL} > q_{HH}$.

**E.3 Compare the conditions for $\gamma_1 = 1$ with those for $\gamma_2 = 1$**

We found that $\gamma_1 = 1$ (and so $\gamma_2 = 0$) for some (not too high) degree of convexity of $S' (·)$, such that $r_{HH} > r_{LL}$, and for $\frac{\Delta \theta}{\Delta \sigma}$ sufficiently large. Suppose these conditions are met and fix the maximum degree of convexity of $S' (·)$ for which $r_{HH} > r_{LL}$ with $\gamma_1 = 1$. Further suppose that, for some degree of convexity of $S' (·)$, $\gamma_2 = 1$. We found that, for this to be the case, it is necessary that $q_{LL} > q_{HH}$. Recall also that $q_{LL} > q_{HH}$ at the solution to $\Gamma'$. We calculate

$$[S'(q_{HH}) - S'(q_{LL})] - [S'(q_{HH}) - S'(q_{LL})] = -2(1 + \gamma_2 \frac{1 - \mu}{\mu}) \Delta \sigma.$$

This shows that the condition $r_{LL} > r_{HH}$ that we found to be necessary for $\gamma_2 = 1$, is tightened as $\gamma_2$ increases. Therefore, moving from the maximum degree of convexity of $S' (·)$ for which $r_{HH} > r_{LL}$ with $\gamma_1 = 1$ (and $\gamma_2 = 0$) to some degree of convexity of $S' (·)$ for which $\gamma_2 = 1$, it must be the case that the latter degree of convexity is higher.

**F Proof of Proposition 3**

Assume $S' (·)$ is sufficiently convex that (15) is not satisfied at the solution to $\Gamma'$, in which case the previous solution does not satisfy the constraints in $\Gamma$. In this case, (IC3)
is more stringent than (IC1) and is thus binding. We first present the reduced problem, to be denoted $\Gamma''$, in which (IC3) is binding. We then present the solution and show that, under the assumptions of the proposition, the conditions under which (IC3) implies (IC1) is satisfied. We further analyze the solution and distinguish two cases, according to whether $r_{HH} > r_{LL}$ or $r_{HH} \leq r_{LL}$. We lastly show that, in either case, the solution to $\Gamma''$ satisfies the constraints in $\Gamma$.

F.1 The reduced problem $\Gamma''$

Let $\Gamma''$ denote the reduced problem in which (IC3) is binding and the remaining binding constraints are still those leading to the rents (12a), (12b) and (12d). Under these circumstances, $\Pi_{LL} = \Delta \theta q_{HH} + \Delta \sigma r_{HL} - \Delta \sigma r_{HH}$ and

$$\Pi_{ LH} = \alpha_1 (\Delta \theta q_{HH} + \Delta \sigma r_{HL}) + \alpha_2 (\Delta \theta q_{HH} + \Delta \sigma r_{HL} + \Delta \sigma r_{LL} - \Delta \sigma r_{HH})$$

$$+ \alpha_3 (\Delta \theta q_{HL} + \Delta \sigma r_{HL}),$$

with $0 \leq \alpha_k \leq 1$, $k = 1, 2, 3$, and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Reformulating the objective function in $\Gamma$ to embody the above rents, we can write the first-order conditions with respect to quantities. We obtain $q_{LL}^*$ and $q_{ LH}^*$ together with

$$S'(q_{LL}) = \theta_L - \sigma_L + \frac{1 - \mu}{\mu} \alpha_2 \Delta \sigma$$

$$S'(q_{ LH}) = \theta_L + \sigma_L - \frac{1 - \mu}{\mu} \alpha_2 \Delta \sigma,$$

for type $LL$,

$$S'(q_{HL}) = \theta_H - \sigma_L + \frac{\nu (1 - \mu) \alpha_3 \Delta \theta + [1 - (1 - \nu) \mu] \Delta \sigma}{(1 - \nu) \mu}$$

$$S'(q_{ HL}) = \theta_H + \sigma_L + \frac{\nu (1 - \mu) \alpha_3 \Delta \theta - [1 - (1 - \nu) \mu] \Delta \sigma}{(1 - \nu) \mu},$$

for type $HL$, and

$$S'(q_{HH}) = \theta_H - \sigma_H + \frac{\nu \{[\mu + (1 - \mu) (\alpha_1 + \alpha_2)] \Delta \theta - [\mu + (1 - \mu) \alpha_2] \Delta \sigma\}}{(1 - \nu) (1 - \mu)}$$

$$S'(q_{ HH}) = \theta_H + \sigma_H + \frac{\nu \{[\mu + (1 - \mu) (\alpha_1 + \alpha_2)] \Delta \theta + [\mu + (1 - \mu) \alpha_2] \Delta \sigma\}}{(1 - \nu) (1 - \mu)},$$

for type $HH$.

F.1.1 Show that $\alpha_3 = 0$

Using the expressions of the rents in $\Gamma''$, (IC5) is rewritten as $r_{HH} > r_{HL}$. Together with the hypothesis that (IC3) is binding, this involves that it is necessary to have
\[q_{HH} \geq q_{HL}.\]

At the solution to \(\Gamma''\), \(\overline{q}_{HL} > \overline{q}_{HH}\). Therefore, to have \(q_{HH} \geq q_{HL}\) it is necessary that \(\underline{q}_{HH} > \underline{q}_{HL}\). Assuming that this condition is satisfied, it implies that \(r_{HH} > r_{HL}\). Together with \(q_{HH} \geq q_{HL}\) it means that \(\alpha_3 = 0\). Show now that the assumed condition \(\underline{q}_{HH} > \underline{q}_{HL}\) is satisfied. This is equivalent to \(\frac{\Delta \sigma}{\Delta \sigma} < \frac{1 - \mu}{\mu} + 1\) whenever \(\alpha_1 = 1\) i.e., with \(r_{HH} > r_{LL}\), and to \(\frac{\Delta \sigma}{\Delta \sigma} < \frac{1 - \mu}{\mu} + \mu\) whenever \(\alpha_2 = 1\) i.e., with \(r_{LL} > r_{HH}\), both inequalities being satisfied by assumption of the proposition. At a later stage, we show that there does not exist a situation in which both \(\alpha_2 \neq 1\) and \(\alpha_1 \neq 1\) (i.e., \(r_{HH} = r_{LL}\)). Thus, the condition \(\underline{q}_{HH} > \underline{q}_{HL}\) holds true.

**F.2 The conditions under which (IC3) implies (IC1)**

(IC3) implies (IC1) if and only if

\[
(\Delta \theta - \Delta \sigma) (q_{HH} - \underline{q}_{HL}) \geq (\Delta \theta + \Delta \sigma) (\overline{q}_{HL} - \overline{q}_{HH}).
\]

(29)

Recall from the proof above that \(\overline{q}_{HL} > \overline{q}_{HH}\) and \(\underline{q}_{HH} > \underline{q}_{HL}\). Moreover,

\[
[S'(\overline{q}_{HL}) - S'(\overline{q}_{HH})] - [S'(\overline{q}_{HH}) - S'(\overline{q}_{HL})] = -2\nu \frac{\Delta \theta}{(1 - \nu) (1 - \mu)}.
\]

showing that (29) is satisfied if and only if \(S'(\cdot)\) sufficiently convex. As (29) is the converse of (15), it is satisfied by the assumption of the proposition.

**F.3 Either \(\alpha_1 = 1\) or \(\alpha_2 = 1\)**

As from the expressions of the rents, \(\alpha_1 = 1\) when \(r_{HH} > r_{LL}\) and \(\alpha_2 = 1\) when \(r_{LL} > r_{HH}\). We calculate

\[
S'(\overline{q}_{HH}) - S'(\overline{q}_{HL}) = \Delta \theta + \left(1 + \frac{1 - \mu}{\mu} \alpha_2\right) \Delta \sigma + \frac{\nu \{\Delta \theta + [\mu + (1 - \mu) \alpha_2] \Delta \sigma\}}{(1 - \nu) (1 - \mu)}.
\]

\[
S'(\underline{q}_{HH}) - S'(\underline{q}_{HL}) = \Delta \theta - \left(1 + \frac{1 - \mu}{\mu} \alpha_2\right) \Delta \sigma + \frac{\nu \{\Delta \theta - [\mu + (1 - \mu) \alpha_2] \Delta \sigma\}}{(1 - \nu) (1 - \mu)}.
\]

and we see that \(\overline{q}_{LL} > \overline{q}_{HH}\). Moreover, \(\underline{q}_{LL} > \underline{q}_{HH}\) if and only if

\[
\frac{\Delta \theta}{\Delta \sigma} > \frac{\nu [\mu + (1 - \mu) \alpha_2] + (1 - \nu) (1 - \mu) (1 + \frac{1 - \mu}{\mu} \alpha_2)}{1 - \nu (1 - \mu)}.
\]

This inequality holds true as the right-hand side is smaller than 1. Hence, the condition \(r_{HH} > r_{LL}\) (under which \(\alpha_2 = 0\)) does not hold for all degrees of convexities of \(S'(\cdot)\). We further calculate

\[
[S'(\underline{q}_{HH}) - S'(\underline{q}_{LL})] - [S'(\overline{q}_{HH}) - S'(\overline{q}_{LL})] = -2\Delta \sigma \left\{1 + \frac{\nu [\mu + (1 - \mu) \alpha_2] + \frac{1 - \mu}{\mu} \alpha_2}{(1 - \nu) (1 - \mu)}\right\}.
\]

34
The right-hand side being negative, \( r_{HH} > r_{LL} \) if and only if \( S'(\cdot) \) is not very convex. Moreover, for any given degree of convexity of \( S'(\cdot) \), the bigger \( \alpha_2 \) the more relaxed the condition \( r_{HH} > r_{LL} \). This shows that \( r_{HH} \leq r_{LL} \) and \( \alpha_2 > 0 \) hold jointly if and only if \( \alpha_2 = 1 \) (with \( r_{HH} = r_{LL} \) for a single degree of convexity of \( S'(\cdot) \)). Therefore, the only cases that are feasible at the solution to \( \Gamma'' \) are \( \alpha_1 = 1 \) and \( \alpha_2 = 1 \). The former arises when \( S'(\cdot) \) is not very convex and \( r_{HH} > r_{LL} \). The latter arises when \( S'(\cdot) \) is sufficiently convex to have \( r_{LL} \geq r_{HH} \).

### F.4 Check incentive constraints in \( \Gamma \)

#### F.4.1 Case \( \alpha_1 = 1 \)

(\( IC3 \)) is binding and, as previously shown, implies (\( IC1 \)). Hence, they are both satisfied. (\( IC2 \)) is rewritten as \( r_{LL} \geq r_{HH} \), (\( IC4 \)) as \( \Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HL} \), (\( IC5 \)) as \( r_{HH} \geq r_{HL} \), (\( IC6 \)) as \( \Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HL} \), (\( IC10 \)) as \( q_{LL} \geq q_{HH} \) and (\( IC12 \)) as \( \Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{LL} \).

Check (\( IC2 \)). We calculate

\[
[S'(q_{HH}) - S'(\bar{q}_{HH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{LL})] = -2 \frac{\nu \mu}{(1 - \nu)(1 - \mu)} \Delta \sigma
\]

\[
[S'(q_{HL}) - S'(\bar{q}_{HL})] - [S'(\bar{q}_{HL}) - S'(\bar{q}_{HH})] = -2 \frac{\nu}{(1 - \nu)(1 - \mu)} \Delta \theta.
\]

The right-hand side of the former equality is closer to zero than the right-hand side of the latter. Hence, when \( S'(\cdot) \) is linear, \( r_{LL} - r_{HH} \) is less negative than is \( q_{HH} - q_{HL} \). Moreover, because \( q_{LL} > q_{HH} \), \( r_{LL} - r_{HH} \) is increased faster than is \( q_{HH} - q_{HL} \) as \( S'(\cdot) \) becomes more convex. This means that \( r_{LL} \geq r_{HH} \) holds as long as \( q_{HH} \geq q_{HL} \). (\( IC2 \)) is thus satisfied.

We saw that \( \bar{q}_{LL} > q_{HH} \) and \( \bar{q}_{LL} > \bar{q}_{HH} \) so that \( q_{LL} > q_{HH} \). Together with (\( IC5 \)), this involves that (\( IC4 \)) is satisfied. Moreover, \( r_{HH} > r_{HL} \) (as previously proved) so that (\( IC5 \)) is satisfied. (\( IC10 \), (\( IC2 \) and (\( IC5 \)) imply that (\( IC6 \)) is satisfied as well.

We further have

\[
S'(q_{LL}) - S'(q_{HH}) = \frac{[1 - \mu (1 - \nu)] \Delta \theta + \nu \mu \Delta \sigma}{(1 - \nu)(1 - \mu)}
\]

\[
S'(\bar{q}_{LL}) - S'(\bar{q}_{HH}) = \frac{[1 - \mu (1 - \nu)] \Delta \theta + \nu \mu \Delta \sigma}{(1 - \nu)(1 - \mu)}
\]

from which both \( q_{LL} > q_{HH} \) together with \( r_{HH} > r_{LL} \). (\( IC12 \) is satisfied. Lastly, because \( q_{LL} > q_{HH} \) and \( \bar{q}_{LL} > \bar{q}_{HH} \). Hence, (\( IC10 \)) is satisfied.

#### F.4.2 Case \( \alpha_2 = 1 \)

As before, (\( IC1 \) and (\( IC3 \)) are both satisfied. (\( IC2 \)) becomes \( r_{LL} \geq r_{HH} \), (\( IC4 \)) \( \Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HL} \), (\( IC5 \)) \( r_{HH} \geq r_{HL} \), (\( IC6 \)) \( \Delta \theta q_{LL} + \Delta \sigma (r_{HH} - r_{HL}) \geq \Delta \theta q_{HH} + \Delta \sigma (\bar{r}_{HH} - \bar{r}_{HL}) \).
\[ \Delta \sigma (r_{LL} - r_{HH}), \quad (IC9) \quad \Delta \theta q_{HH} + \Delta \sigma r_{LL} \geq \Delta \theta q_{HL} + \Delta \sigma r_{HH}, \quad (IC10) \quad \Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{LL}. \]

At the solution to \( \Gamma'' \), \( q_{LL} > q_{HH} \) together with \( r_{LL} > r_{HH} \). Hence, \( r_{LL} > r_{HH} \) so that (IC2) is satisfied. Moreover, we know that \( q_{LL} > q_{HH} \) and \( r_{HH} > r_{LL} \) so that \( q_{LL} > q_{HH} \). Together with (IC5), this implies that (IC4) is satisfied. Also, because \( q_{HH} > q_{LL} \) and \( r_{LL} > r_{HH} \), we have \( r_{HH} \geq r_{HL} \) and (IC5) is satisfied.

We further have \( q_{LL} > q_{HL} \) and \( r_{HH} > r_{LL} \) so that \( q_{HH} > q_{HL} \). Together with (IC2) and (IC5), this involves that (IC6) is satisfied. Then, because \( q_{HH} > q_{HL} \) and \( r_{HH} > r_{LL} \), (IC9) is satisfied as well. (IC10) is rewritten as

\[ (\Delta \theta - \Delta \sigma) (q_{HH} - q_{LL}) + (\Delta \sigma + \Delta \theta) (q_{HH} - q_{HL}) + \Delta \sigma (q_{HH} - q_{LL}) + \Delta \sigma (q_{HH} - q_{LL}) \geq 0. \]

As all terms in the left-hand side are positive, (IC10) is satisfied. Lastly, we have

\[ S'(q_{LL}) - S'(q_{HL}) = S'(q_{HH}) - S'(q_{LL}) \]

so that \( q_{LL} \geq q_{HL} \) with \( S'(\cdot) \) convex. Jointly with (IC10), this involves that (IC12) holds.

**G  Proof of Corollary 3**

We first show that, under the assumption of the corollary that the condition of the kind \( \Delta \theta \) is violated in any of the previous propositions and corollaries, the solution to \( \Gamma \) must satisfy \( q_{HH} = q_{HL} \) and \( r_{HH} = r_{HL} \). We then state the reduced problem \( \Gamma'' \) that incorporates these conditions and find its solution. We lastly show that this solution satisfies the incentive constraints in \( \Gamma \).

**G.1 Show that \( q_{HH} = q_{HL} \) and \( r_{HH} = r_{HL} \) whenever \( \Delta \theta \Delta \sigma \geq Y \)**

Recall that (IC1) is binding and that it implies (IC3) if and only if \( S'(\cdot) \) is not very convex and \( \frac{\Delta \theta}{\Delta \sigma} < Y \) (Proposition 1 and 2 as well as Corollary 1 and 2). When this is the case, \( q_{HH} > q_{HL} \) and the condition \( \frac{\Delta \theta}{\Delta \sigma} < Y \) means that \( q_{HL} > q_{HH} \); which is necessary for having (IC1) binding and (IC3) implied by (IC1). Recall also that (IC3) is binding and that it implies (IC1) if and only if \( S'(\cdot) \) is sufficiently convex and \( \frac{\Delta \theta}{\Delta \sigma} \) satisfies again some condition of the kind \( \frac{\Delta \theta}{\Delta \sigma} < Y \) (Proposition 3). When this is the case, \( q_{HL} > q_{HH} \) and the condition \( \frac{\Delta \theta}{\Delta \sigma} < Y \) means that \( q_{HH} > q_{HL} \), which is necessary for having (IC3) binding and (IC1) implied by (IC3). Therefore, when the condition of the kind \( \frac{\Delta \theta}{\Delta \sigma} < Y \) is not satisfied in any of the previous propositions and corollaries, it must be the case that both \( q_{HH} = q_{HL} \) and \( q_{HL} = q_{HH} \), with (IC1) and (IC3) both binding. Equivalently, \( q_{HH} = q_{HL} \) together with \( r_{HH} = r_{HL} \).
G.2 The relaxed problem $\Gamma'''$

(IC7) is now identical to (IC9). Consider the relaxed problem $\Gamma'''$ with rents

\[
\begin{align*}
\Pi_{HL} &= 0 \\
\Pi_{LL} &= \Delta \theta q_{HH} \\
\Pi_{HH} &= \Delta \sigma r_{HH} \\
\Pi_{LH} &= \max \{ \Delta \theta q_{HH} + \Delta \sigma r_{HH}; \Delta \theta q_{HH} + \Delta \sigma r_{LL} \}.
\end{align*}
\]

Let us reformulate the latter rent as

\[
\Pi_{LH} = \beta_1 (\Delta \theta q_{HH} + \Delta \sigma r_{HH}) + \beta_2 (\Delta \theta q_{HH} + \Delta \sigma r_{LL}),
\]

with $\beta_1 + \beta_2 = 1$, $\beta_1 = 1$ if and only if $r_{HH} > r_{LL}$, $\beta_2 = 1$ if and only if $r_{LL} > r_{HH}$ and both $\beta_1 > 0$ and $\beta_2 > 0$ if and only if $r_{LL} = r_{HH}$. The agent’s expected rent is written

\[
\nu \Delta \theta q_{HH} + (1 - \nu + \nu \beta_1) (1 - \mu) \Delta \sigma r_{HH} + \nu (1 - \mu) \beta_2 \Delta \sigma r_{LL}.
\]

At the solution to $\Gamma'''$, the type $LH$ quantities are set at the FB levels. Moreover,

\[
\begin{align*}
S'(q_{LL}) &= \theta_L - \sigma_L + \beta_2 \frac{1 - \mu}{\mu} \Delta \sigma \\
S'(q_{LH}) &= \theta_L + \sigma_L - \beta_2 \frac{1 - \mu}{\mu} \Delta \sigma
\end{align*}
\]

for type $LL$,

\[
\begin{align*}
S'(q_{HH}) &= \theta_H - \sigma_H + \frac{\nu}{(1 - \nu) (1 - \mu)} \Delta \theta + \frac{1 - \nu (1 - \beta_1)}{1 - \nu} \Delta \sigma \\
S'(\bar{q}_{HH}) &= \theta_H + \sigma_H + \frac{\nu}{(1 - \nu) (1 - \mu)} \Delta \theta - \frac{1 - \nu (1 - \beta_1)}{1 - \nu} \Delta \sigma
\end{align*}
\]

for type $HH$, and $q_{LH} = q_{HL}$ together with $\bar{q}_{HL} = \bar{q}_{HH}$ for type $HL$.

G.3 The case of $\beta_1 = 1$

Assume that $\beta_1 = 1$ and so $r_{HH} > r_{LL}$. (IC2) is rewritten $r_{LH} \geq r_{HH}$, (IC4) as $q_{LL} > q_{HH}$, (IC5) as $r_{HH} \geq r_{HL}$, (IC6) as $\Delta \theta q_{LL} + \Delta \sigma r_{LL} \geq \Delta \theta q_{HH} + \Delta \sigma r_{HH}$, (IC10) as $q_{LL} \geq q_{HH}$ and (IC12) as $\Delta \theta q_{LL} + \Delta \sigma r_{HH} \geq \Delta \theta q_{HH} + \Delta \sigma r_{LL}$. (IC2) and (IC10) imply (IC6). (IC4) and $r_{HH} > r_{LL}$ imply (IC12). (IC5) is satisfied with equality.

Check (IC2). We have

\[
S'(q_{HH}) - S'(q_{LH}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta + \frac{1}{1 - \nu} \Delta \sigma
\]
so that \( q_{LH} > q_{HH} \). Moreover,

\[
S'(\bar{q}_{HH}) - S'(\bar{q}_{LH}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta - \frac{1}{1 - \nu} \Delta \sigma
\]

so that \( \bar{q}_{LH} > \bar{q}_{HH} \) if and only if

\[
\frac{\Delta \theta}{\Delta \sigma} > \frac{1 - \mu}{1 - \mu + \nu \mu},
\]

which is true. We calculate

\[
[S'(\underline{q}_{HH}) - S'(\underline{q}_{LH})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{LH})] = \frac{2}{1 - \nu} \Delta \sigma,
\]

(30)

which shows that \( r_{LH} \geq r_{HH} \) (i.e., (IC2) holds) at the solution to \( \Gamma'' \) if and only if \( S'(. \cdot) \) is not too concave.

Check (IC4). We have

\[
S'(\underline{q}_{HH}) - S'(\underline{q}_{LL}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta + \frac{\nu}{1 - \nu} \Delta \sigma
\]

so that \( \underline{q}_{LL} > \underline{q}_{HH} \). Moreover,

\[
S'(\bar{q}_{HH}) - S'(\bar{q}_{LL}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta - \frac{\nu}{1 - \nu} \Delta \sigma,
\]

which is positive and so \( \bar{q}_{LL} > \bar{q}_{HH} \). Hence, \( q_{LL} > q_{HH} \) and so (IC4) is satisfied.

Check the condition \( r_{HH} > r_{LL} \) that we assumed satisfied. Using the above results, we calculate

\[
[S'(\underline{q}_{HH}) - S'(\underline{q}_{LL})] - [S'(\bar{q}_{HH}) - S'(\bar{q}_{LL})] = \frac{2}{1 - \nu} \Delta \sigma,
\]

(31)

which shows that \( r_{HH} > r_{LL} \) if and only if \( S'(. \cdot) \) is sufficiently concave.

**G.3.1 When the solution to \( \Gamma'' \) violates incentive constraints in \( \Gamma \)**

The only conditions that might not hold jointly in \( \Gamma'' \) are \( r_{HH} > r_{LL} \) and \( r_{LH} \geq r_{HH} \) (the constraint (IC2)), both depending upon the shape of \( S'(. \cdot) \). As proved above, the condition \( r_{HH} > r_{LL} \) holds but the condition \( r_{LH} \geq r_{HH} \) does not when \( S'(. \cdot) \) is very concave. In this case, the solution to \( \Gamma \) is such that (IC2) is binding, or, equivalently, \( r_{LH} = r_{HH} \). Moreover, by comparing (31) with (30), we notice that there exist some intermediary degrees of concavity of \( S'(. \cdot) \) for which both \( r_{HH} > r_{LL} \) and \( r_{LH} > r_{HH} \).
G.4 The case of $\beta_2 > 0$

Assume that $\beta_2 = 1$ and so $r_{LL} > r_{HH}$, (IC2) is rewritten as $r_{LH} \geq r_{LL}$, (IC4) as $q_{LL} \geq q_{HH}$, (IC5) as $r_{HH} \geq r_{HL}$, (IC6) as $\Delta \theta q_{LH} + \Delta r_{HH} \geq \Delta \theta q_{HH} + \Delta r_{LL}$, (IC10) as $\Delta \theta q_{LH} + \Delta r_{HH} \geq \Delta \theta q_{HH} + \Delta r_{HH}$ and (IC12) as $\Delta \theta q_{HH} + \Delta r_{HH} \geq \Delta \theta q_{HH} + \Delta r_{LL}$. The condition $q_{HH} \geq q_{HH}$ is necessary for (IC10) to hold. Together with (IC2) imply (IC6). Moreover, (IC5) is satisfied with equality by the assumption of $\Gamma^\mu$.

Check (IC2). We calculate

\[
S'(q_{LL}) - S'(q_{LH}) = \frac{\Delta \sigma}{\mu}
\]

Hence, $q_{LL} > q_{LH}, q_{HH} > q_{LH}$ and so (IC2) is satisfied.

Check (IC4). We calculate

\[
S'(q_{HH}) - S'(q_{LH}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta + \frac{1 - \mu}{\mu} \Delta \sigma
\]

Therefore, $q_{LL} > q_{HH}$. Moreover,

\[
S'(q_{HH}) - S'(q_{LL}) = \frac{1 - \mu (1 - \nu)}{(1 - \nu) (1 - \mu)} \Delta \theta - \frac{1 - \mu}{\mu} \Delta \sigma.
\]

so that $q_{HH} > q_{HH}$ if and only if

\[
\frac{\Delta \theta}{\Delta \sigma} > \frac{(1 - \nu) (1 - \mu)^2}{\mu [1 - \mu (1 - \nu)]}.
\]

We show below that this condition is satisfied and so (IC4) holds.

Check (IC10). We have

\[
S'(q_{HH}) - S'(q_{LH}) = 1 + \frac{\nu}{(1 - \nu) (1 - \mu)} \Delta \theta + \Delta \sigma
\]

\[
S'(q_{HH}) - S'(q_{LL}) = 1 + \frac{\nu}{(1 - \nu) (1 - \mu)} \Delta \theta - \Delta \sigma
\]

so that $q_{HH} > q_{HH}$ and $q_{HH} > q_{HH}$. It follows that $q_{LL} > q_{HH}$. Together with the assumption that $r_{LL} > r_{HH}$, this implies that (IC10) is satisfied.

Check (IC12). This is equivalent to

\[
(\Delta \theta + \Delta \sigma) (q_{LL} - q_{HH}) \geq (\Delta \theta - \Delta \sigma) (q_{HH} - q_{LL}).
\]

Because $q_{LL} > q_{HH}$ and $q_{HH} > q_{HH}$, this condition holds true.

Check the condition $r_{LL} > r_{HH}$ that we assumed satisfied. From the previous cal-
culations of $S'(q_{HH}) - S'(q_{LL})$ and $S' (\bar{q}_{HH}) - S' (\bar{q}_{LL})$, we have $\bar{q}_{LL} > \bar{q}_{HH}$, whereas $q_{LL} > q_{HH}$ if and only if (32) is satisfied. We suppose that (32) holds and we calculate

$$[S'(q_{HH}) - S'(q_{LL})] - [S' (\bar{q}_{HH}) - S' (\bar{q}_{LL})] = -2 \frac{1-\mu}{\mu} \Delta \sigma,$$

which shows that $r_{LL} > r_{HH}$ if and only if $S' (\cdot)$ is sufficiently convex. Moreover, (32) does hold because the case in which $S' (\cdot)$ is so convex that $r_{LL} > r_{HH}$, is associated in the corollary with the hypothesis that the condition $\frac{\Delta \theta}{\Delta \sigma} < 1 + \frac{1-\mu}{\nu \mu}$ in Proposition 3 is not satisfied.

Lastly, we found that $r_{HH} > r_{LL}$ if and only if $S' (\cdot)$ is sufficiently concave, whereas $r_{LL} > r_{HH}$ if and only if $S' (\cdot)$ is sufficiently convex. Therefore, when $S' (\cdot)$ is almost linear, it must be the case that $r_{LL} = r_{HH}$ and $0 < \beta_2 < 1$. The incentive constraints in $\Gamma$ are verified as above.