Sequential screening and the relationship between principal's preferences and agent's incentives

Daniel Danau and Annalisa Vinella

Sequential screening and the relationship between principal's preferences and agent's incentives[*]

Daniel Danau[†]   Annalisa Vinella[‡]

Abstract

In sequential screening problems it is found that, under some regularity conditions, local incentive compatibility constraints are sufficient for implementability. However, this follows from the assumption that the possible distributions of the unknown variable satisfy either first-order stochastic dominance or mean-preserving spread. That assumption is matched with private information about either the expected value or the spread of the variable. In this paper we allow for private information about both parameters. In a setting with four possible cost distributions, two with equal expected values and different spreads and two with different expected values and equal spreads, we show that there can be multiple combinations of binding incentive constraints depending on the principal's preferences. The less concave / more convex that the marginal surplus is, the more that the binding incentive constraints are related to private information about one parameter of the distribution relative to the other. Yet, screening is always two-dimensional. Local incentive constraints are sufficient, as in the literature, only when the marginal surplus is sufficiently convex. We further suggest that, in the same vein as in Consumption theory, the contractual choice can be regarded as mirroring the preference of the decision-maker for a lottery that occasions a higher (certain) cost but grants the possibility of facing more efficient (random) outcomes. Resting on this interpretation, we assess that the benefit of screening the agent in two stages, rather than in the contracting stage only, is higher when the marginal surplus is less concave / more convex.

Keywords: Sequential screening; multidimensional screening; expected cost; spread; marginal surplus; cost overruns

J.E.L. Classification Numbers: D82

[†]Université de Caen Basse-Normandie, Centre de Recherche en Economie et Management, 19 rue Claude Bloch, 14000 Caen (France). *Corresponding author*. E-mail: daniel.danau@unicaen.fr

[‡]University of Bari, Department of Economics and Quantitative methods, Largo Abbazia S. Scolastica, 53, 70124 Bari (Italy). E-mail: annalisa.vinella@uniba.it

1 Introduction

Motivation

By now it is well known that there are important gaps between estimates and real values of the costs of public projects. In a number of empirical investigations underling these gaps, such as the studies on transportation of Flyvbjerg [15] and Flyvbjerg *et al.* [16] - [17], it is argued that "wrong" estimates are plausibly related to the incentives of the delegated firms to manipulate the true estimates *vis-à-vis* the delegating authorities. In the words of Flyvbjerg *et al.* [16]: "the data and tests lead us to reject technical explanations of forecasting errors. Such explanations simply do not fit the data" (p. 286). Whereas the empirical research is suggestive of a poor contractual design, it is not clear how a delegation contract should be structured when the agent is privately informed about the distribution of an unknown variable that matters in the relationship with the principal. Despite that an estimate includes both the point estimate (the mean) and the interval estimate (the spread around the mean) of the unknown variable, hitherto there has been no theoretical study about the optimal screening of agents holding private information about the two parameters of the distribution and being potentially motivated to misrepresent both of them.

In incentive theory, optimal delegation has been studied in situations in which the agent knows either the expected value of some variable or its spread, with results that are now fairly known. Private information about the expected value is matched with the assumption that the possible distributions of the unknown variable (the types) are ranked according to first-order stochastic dominance (as in Riordan and Sappington [33], Courty and Li [9], Eso and Szentes [14], Krahmer and Strausz [24]). Private information about the spread is matched with the assumption that the distributions are ranked in the sense of mean-preserving spread (as in Courty and Li [9], Dai *et al.* [10], Hoffmann and Inderst [20]). Screening occurs sequentially because, as is natural in such settings, in addition to observing the distribution of the variable before signing the contract, the agent will also learn the realized value in a later stage. The stochastic dominance assumption is extremely useful in these contexts. In the same vein as in static problems *à la* Baron and Myerson [5], it entails that, under some regularity conditions, relevant incentive constraints are those whereby higher-order types are not tempted to mimic adjacent lower-order types. Accordingly, only local incentive constraints are binding in the screening problem. The principal should focus on the incentives to exaggerate the expected cost in one case, and to understate the spread in the other.

When both the expected value and the spread of the unknown variable are privately known to the agent, the identification of the relevant incentive constraints becomes quickly more complex because one cannot impose any standard stochastic dominance ordering on the distributions. Then, one cannot hinge on the existing studies to draw conclusions about the incentives to be considered in contractual design. There are thus a few open questions. When delegating a project under incomplete information about the distribution of the cost, should a public authority require the firm to report the entire estimate, *i.e.*, both the mean and the

spread, thus tackling a two-dimensional screening problem? In contractual design, should the authority be concerned with the firm over/understating the expected cost, knowing that such a lie can be combined with a lie about the spread? Similarly, should the authority be concerned with the firm over/understating the spread, knowing that such a lie can be combined with a lie about the expected cost? Our goal is to address these research questions.

By offering replies to these questions, we aim at shedding light on a complementary issue as well. Both in empirical studies like those aforementioned and in public debates a major concern related to the misrepresentation of cost estimates is that cost overruns might result. That is, it might be the case that the cost estimates delivered in the first stage of the relationship are lower than the costs actually realized in a later stage because the real estimates were camouflaged purposely. This seems to suggest that an agent holding private information about the expected value and the spread of the cost might be willing to understate at least one of them. However, it is not clear that the agent would actually display such incentives. For instance, one would expect the agent to benefit if he exaggerates the expected cost when it is low (rather than understating it) and correctly reveals the spread. Our investigation is intended to figure out what incentives cost overruns might exactly mirror (whether those to understate the expected cost, or the spread, or a mix of the two) and should be accounted for accordingly in contractual design.[1]

Setting

We consider the delegation of a production activity by a principal to an agent holding private information about both the expected value and the spread of the unit cost of production. Each parameter can take two values and the unit cost can be drawn from four distributions, two having equal mean but different spreads, and two having equal spread but different means. Thus, provided that the agent is required to report both parameters of the distribution when the contract is signed, both unidimensional and two-dimensional lies are possible, and in either direction. As an illustration, an agent with low expected cost and high spread can exaggerate the expected cost, or he can understate the spread, or he can do both thus cheating in two different directions. In line with Courty and Li [9] and with the sequential screening literature in general, the agent also learns the true cost when it is realized and is required to deliver a second report about the newly acquired information. He produces and is compensated thereafter according to the reports made in the two stages.

Main results

Optimality of a two-dimensional report We find that for the principal it is generally optimal to require the agent to report both the mean and the spread when selecting a specific

---

[1]In the delegation of public projects, public authorities are concerned with cost overruns because they are not necessarily borne by the delegated firm. Indeed, the latter might pass them onto the public authority through contractual renegotiation *ex post*. However, before investigating renegotiation issues, it is necessary to understand whether or not one should expect cost overruns to occur in an optimal full-commitment contract, which is the goal of our study.

contractual option among those in the contractual menu. Whereas this finding is intuitive, the identification of the exact incentives to lie that must be considered in contractual design, and hence of the incentive constraints that are binding in the principal's problem, is far from obvious and leads to a bunch of additional results, which we present below.

The relationship between principal's preferences and agent's incentives   Our study delivers a broad lesson about *how* to identify the relevant incentives to misrepresent information in contractual design. It comes out that those incentives are finely related to the principal's preferences for the good delivered by the agent. The link between preferences and binding incentive constraints is novel in the sequential screening literature, and uncommon in incentive problems in general. Given this link, the similarity between sequential problems and static problems does not apply in our setting, in which the distributions are not ordered in terms of stochastic dominance. Indeed, the screening problem admits multiple solutions (rather than one only) and it depends on the third derivative of the surplus function what solution exactly arises among those that are possible. Specifically, the higher that the third derivative is, the more that the contractual solution mirrors a preoccupation for higher-order types being mimicked relative to lower-order types. There are two important implications to this.

Implication 1: The principal is concerned with the possibility of the agent misrepresenting the spread rather than the expected value, or *vice versa*, depending on whether the degree of private information about the expected value is greater or smaller than that about the spread. Therefore, whereas it is optimal for the principal to screen the agent with respect to both information dimensions, it depends on her own preferences whether she should be more concerned with one or the other dimension being misrepresented.

Implication 2: Local incentive constraints imply global incentive constraints if and only if the principal's marginal surplus is sufficiently convex. Therefore, when this is the case, the type ordering plays a similar role to stochastic dominance in the screening problem, and the similarity between sequential and static screening problems is then restored.

Optimal contract and prevention of cost overruns   The literature provides a clue about the occurrence of cost overruns when private information is unidimensional. That is, one should not expect overruns to occur when private information is about the expected cost. Indeed, when the expected cost is high, the agent is of a "bad" type and has no incentives to understate it. By contrast, one can expect overruns to occur when private information is about the spread. Indeed, when the spread is high, the agent is of a "good" type and might want to understate it, as shown by Courty and Li [9] and the subsequent studies in which mean-preserving spread is assumed. The reason why a high-spread type is "good", and hence desirable to the principal, is understood if it is considered that the principal enjoys an option value by conditioning the volume of trade on the real value of the unknown variable, which will be realized after the contract is drawn up. The higher the spread, the greater the option value to the principal. This is consistent with the theory of investment under uncertainty

(Dixit and Pindyck [12]), according to which uncertainty is beneficial to a decision-maker who can condition her decisions on future contingencies, benefits being more pronounced when uncertainty is high. In our setting, without specific investigation, it is not clear that an agent with high spread is of a good type, provided that the desirability of an agent's type to the principal will also depend on the expected cost. Therefore, one cannot infer that cost overruns would occur if the principal faces a high-spread agent. Nor is it clear that the principal dislikes an agent with high expected cost as the spread will matter as well. However, we find that, regardless of the principal's preferences for the good, conceding a rent to eliminate the agent's incentives to understate the spread suffices to avoid understatement of the expected cost as well.

Extension to more than four types In settings with more than four cost distributions, not only can the agent over/understate either parameter of the distribution. He can also choose the extent to which, for each parameter, the reported value will diverge from the true value. By extending the analysis to such settings, we find that when the difference between the expected costs of any two types is either greater or smaller than that between the respective spreads, thus mirroring a monotonicity on the extent to which one parameter can be misrepresented relative to the other across types, the principal can restrict attention to the same kind of incentives that arise in the four-distribution setting. As an illustration, with more than four distributions, an agent with low expected cost and high spread may want to announce a even higher spread, selecting the exact value to be announced among many possible higher values. However, it turns out that this does not need to be an issue because that type of agent will have stronger incentives to exaggerate the spread in a "local" sense. Therefore, when the extent of private information about one parameter of the distribution is monotonic relative to that about the other parameter, our main results naturally extend to settings with more than four distributions. This shows that the complications we detect in the screening problem are rooted in the possibility of the agent lying on either information dimension in either direction, rather than in the extent to which the agent would misrepresent each dimension.

After drawing the results described so far, we suggest two convenient ways of interpreting the relationship between the principal's preferences and the contractual choice. One interpretation rests on the theory of consumption under uncertainty, the results of which depend as well on the shape of the decision-maker's marginal utility (see, for instance, Leland [27] and Menegatti [30]). The other interpretation refers to monopoly regulation and rests on the correspondence that one can identify between the curvature of the marginal surplus function and the price-elasticity of market demand. Accordingly, the link between preferences and contractual choice in our model is transposed into a link between demand elasticity and optimal regulation. Because regulators refer widely to elasticity estimates, this link is indicative of how they can make use of our theoretical predictions along the current practice.

Related literature

Our study is related to sequential screening problems in which private information is about either the expected value or the spread of the unknown variable, such as those analyzed by Riordan and Sappington [33] and Courty and Li [9]. However, as opposed to those models, we take the signals about the distribution to belong to a discrete set to circumvent the difficulties with identifying the solution which have been found to arise in multidimensional screening problems with a continuum of types (Wilson [35], Armstrong [1], Rochet and Choné [8]).

In the same vein as in the multidimensional screening problems studied by Armstrong and Rochet [3] and Armstrong [2], we identify various combinations of binding incentive constraints at optimum, though in different contexts and with results that are not directly comparable. Armstrong and Rochet [3] consider two distinct activities and private information about the unit cost of each activity. When the two pieces of information are uncorrelated, as in our case, they obtain that only incentives to mimic adjacent types are relevant, as in the unidimensional sequential screening models. In our setting, there is a rich number of combinations of binding incentive constraints, which is due to the fact that the two information dimensions pertain to a distribution. The two pieces of information are uncorrelated but the decision variables (expected production and spread between productions) are interdependent, leading to various solutions that are not comparable to Armstrong and Rochet [3]. Armstrong [2] considers a setting with private information on cost and demand functions of the agent. A link between our study and his study is identified in that some types may be bunched at optimum, because private information on both cost and demand may lead to binding monotonicity conditions, in the spirit of the literature on countervailing incentives (Lewis and Sappington [28] and Maggi and Rodriguez-Clare [29]). We identify that bunching of types may appear at optimum as a way to limit distortions when the private information on one dimension is sufficiently large relative to the other. Hence, types may be bunched together as in Armstrong [2], although not due to binding monotonicity conditions. Moreover, unlike in Armstrong [2], bunching in our setting refers to entire distributions of production, rather than to two quantity levels.

More generally, our paper is related to the literature on the delegation of public activities. Following Riordan and Sappington [33], that literature evolves in two main strands. The first strand includes models in which the firm decides whether to take some action before learning the true cost or the decision is delayed until after the cost is realized (Kjerstad and Vagstad [22], Board [6], Mougeot and Naegelen [31]). The second strand includes models on cost overruns in which the firm is privately informed about the expected cost in the delegation stage, whereas the true cost is observed publicly in a subsequent stage (Spulber [34], Chen and Smith [7]). The issue of cost overruns is also argued to be critical in recent studies on public-private partnerships, such as Laffont [26], Guasch *et al.* [18] - [19], Danau and Vinella [11], Iossa and Martimort [21]. In these studies, the expected cost is commonly known, whereas its realization is observed privately by the firm; none of these studies considers the possibility of the agent observing both the distribution and the true cost privately in the delegation of public activities.

6

1.1 Outline

The remainder of the paper is organized as follows. In section 2 we describe the model and formalize the principal's problem. In section 3 we present a few benchmarks associated with "standard" information structures. In section 4, we figure out the two possible orderings of types; we then look at situations in which screening is sequential only for some types. In section 5 we focus on full sequential screening and identify the relationship between principal's preferences and binding incentive constraints. The optimal contract is characterized for one possible type ordering in section 6. In section 7 we consider a richer set of possible cost distributions. Section 8 provides two alternative interpretations for a deeper understanding of the relationship between preferences and contractual choice. Section 9 concludes. Mathematical details are relegated to an appendix.

2 The model

We consider a principal, P, who delegates an activity to an agent for the provision of a good or service. Production of y units of the good occasions a cost of cy to the agent and is compensated with a payment of t made by P. Consumption of y units of the good yields a gross surplus of $S(y)$ to P, where $S(0) = 0$ and $S(\cdot)$ is three-times continuously differentiable and such that $S'(\cdot) > 0$ and $S''(\cdot) < 0 \ \forall y > 0$.

The relationship between P and the agent takes place in two stages under full commitment. The contract is signed in the *first* stage. At that time, the agent has private information about the exogenous parameters of the distribution of the unit cost $c(\varepsilon) = \theta + \varepsilon\sigma$, say, because of the expertise acquired by running similar activities in the past. The parameters of the distribution are the expected value $\theta > 0$ and the spread $\sigma > 0$. It is common knowledge that θ is either θ_L or θ_H with probabilities ν and $1 - \nu$; σ is either σ_L or σ_H with probabilities μ and $1 - \mu$. We let $\Delta\theta = \theta_H - \theta_L > 0$ and $\Delta\sigma = \sigma_H - \sigma_L > 0$. Furthermore, ε is a shock that will affect the cost realization in the *second* stage. It is commonly known that the shock will be either $\bar{\varepsilon} > 0$ (a bad shock) or $-\bar{\varepsilon}$ (a good shock) with equal probabilities.[2] Because the specific value of $\bar{\varepsilon}$ is irrelevant in the analysis, we normalize it to 1. Once ε is realized, the state of nature (the true unit cost) is determined and the agent observes it privately. Then, he produces the good and is compensated by P according to the terms of the contract.[3]

---

[2] The assumption that the two shock realizations are equally likely is made for simplicity and does not affect results.

[3] The assumption that the stochastic variable is additive in mean and spread and that the mean and the spread are mutually independent is also made in the previous literature, including Courty and Li [9], Eso and Szentes [14] (Example 1) and Hoffmann and Inderst [20]. Krahmer and Strausz [24] model the cost as being the sum of a signal privately known in the first stage and a random shock realized in the second stage. However, these authors mainly focus on the case where the distribution of the shock is independent of the signal.

2.1 The problem of the principal

P offers the agent an incentive contract which includes an allocation $\{y_{ij}(\varepsilon), t_{ij}(\varepsilon)\}$ for each possible distribution $ij \in \Upsilon \equiv \{LL, LH, HL, HH\}$ and each possible shock $\varepsilon \in \{-1; 1\}$. Henceforth, we refer to ij as the agent's type and to $c_{ij}(\varepsilon) = \theta_i + \varepsilon\sigma_j$ as the state of nature realized following the shock. Given the contractual allocation, an agent of type ij obtains the profit $\pi_{ij}(\varepsilon) = t_{ij}(\varepsilon) - (\theta_i + \varepsilon\sigma_j) y_{ij}(\varepsilon)$ when the shock is ε.

To elicit the agent's information about the observed shock in the second stage, conditional on the delivery of a truthful report ij in the first stage, P must design profits in such a way that they comply with *the second-stage incentive constraints*:[4]

$$ic_{ij,\varepsilon} : \pi_{ij}(\varepsilon) \geq \pi_{ij}(\widetilde{\varepsilon}) + (\widetilde{\varepsilon} - \varepsilon)\sigma_j y_{ij}(\widetilde{\varepsilon}), \ \forall ij \in \Upsilon, \ \forall \varepsilon, \widetilde{\varepsilon} \in \{-1; 1\}.$$

Satisfying these constraints requires respecting the following *monotonicity conditions*:

$$MC_{ij} : y_{ij}(-1) \geq y_{ij}(1), \ \forall ij.$$

Provided that this condition holds, it is possible to identify appropriate values of $\pi_{ij}(\varepsilon)$ for any given value of the expected profit $\Pi_{ij} = \mathbb{E}_\varepsilon [\pi_{ij}(\varepsilon)]$. This is the rent to be given up to type ij in order to elicit information in the first stage.[5] To also elicit the agent's information about the observed distribution in the first stage, P must design the expected profits in such a way that the *first-stage incentive constraints* are satisfied:[6]

$$IC_{ij}^{i'j'} : \ \Pi_{ij} \geq \frac{1}{2} \max \left\{ \sum_\varepsilon [t_{i'j'}(\varepsilon) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\varepsilon)] ; \sum_\varepsilon [t_{i'j'}(\widetilde{\varepsilon}) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\widetilde{\varepsilon})] \right\}.$$

The right-hand side of $IC_{ij}^{i'j'}$ represents what type ij would earn by announcing $i'j'$ in the first stage and either ε or $\widetilde{\varepsilon}$ in the second stage. Considering that P also needs to satisfy the *participation constraints*:

$$PC_{ij} : \Pi_{ij} \geq 0, \ \forall ij,$$

the problem of P, denoted Γ, is written as follows:

$$\underset{\{y_{ij}(\varepsilon), \Pi_{ij}\}_{ij \in \Upsilon, \varepsilon \in \{-1; 1\}}}{Max} \sum_{ij \in \Upsilon} \mathbb{E}_{ij} \left[\mathbb{E}_\varepsilon \left[S(y_{ij}(\varepsilon)) - (\theta_i + \varepsilon\sigma_j) y_{ij}(\varepsilon) \right] - \Pi_{ij} \right]$$

subject to
$$MC_{ij}, \ IC_{ij}^{i'j'} \ \text{and} \ PC_{ij}, \ \forall ij \in \Upsilon.$$

---

[4] Applying the Revelation Principle for multistage games (Myerson [32]), in the second stage the agent is required to report only the newly acquired information, namely the realized shock ε, rather than the realized unit cost $\theta_i + \varepsilon\sigma_j$, which might not be consistent with the report delivered in the first stage and would raise the overall number of incentive constraints unnecessarily. Moreover, in the second stage the agent is required to reveal the true shock conditional on revealing the true type in the first stage.

[5] We assume no discounting because all payoffs in the model depend on second-stage production levels and transfers.

[6] In our notation, $ij \neq i'j'$ indicates that either $i \neq i'$ or $j \neq j'$ or both.

Of course, absent information problems, MC_{ij} and $IC_{ij}^{i'j'}$ would be irrelevant and P would be able to decentralize the first-best allocation:

$$S'(y_{ij}^*(\varepsilon)) = \theta_i + \varepsilon\sigma_j, \ \forall ij, \ \forall\varepsilon \tag{1a}$$

$$\Pi_{ij}^* = 0, \ \forall ij. \tag{1b}$$

Although P faces the same kind of constraints as in any sequential screening problem, we will see that in our setting complications arise in the identification of the binding incentive constraints. Before turning to the analysis, it is useful to consider a few benchmarks associated with more standard information structures.

3 Benchmarks

We first propose four benchmarks of sequential screening: in the first two, denoted (1) and (2), the support of the unit cost is continuous; in the subsequent two, denoted $(1')$ and $(2')$, the support is discrete. The last benchmark, denoted (3), is one of multidimensional screening.

(1) **No private information about the spread**   Assume that the unit cost of production c takes value in the interval $[\underline{c}, \overline{c}]$ according to the cumulative distribution function $F_i(c)$, where $i \in \{1, 2, 3, 4\}$ indexes the distribution. The function $F_i(\cdot)$ is continuously differentiable and $f_i(c) = F_i'(c)$ is the density function. Assume that $F_i(\cdot)$ decreases with i, which means that any distribution i dominates the distribution $i+1$ in the sense of first-order stochastic dominance. Denoting $E[c|i] = \theta_i$ the expected unit cost conditional on the distribution being i, we have:

$$\theta_1 < \theta_2 < \theta_3 < \theta_4.$$

Let $h_i = \sum_{k=1}^{i-1} \frac{\nu_k}{\nu_i}, \ \forall i \in \{2, 3, 4\}$, and $h_1 = 0$ be the hazard rate and assume that h_i increases with i. Then, the solution to the problem of the principal is characterized as follows:

$$S'(y_1(c)) = c \ \text{ and } \ S'(y_i(c)) = c + h_i \frac{F_{i-1}(c) - F_i(c)}{f_i(c)}, \ \forall i \in \{2, 3, 4\}$$

$$\Pi_i = \sum_{k=i+1}^{4} \int_{\underline{c}}^{\overline{c}} y_k(c)(F_{k-1}(c) - F_k(c)) dc, \ \forall i \in \{1, 2, 3\}, \ \text{ and } \ \Pi_4 = 0.$$

At the solution all the incentive constraints but the local downward incentive constraints are slack. To see why, consider that under first-order stochastic dominance $F_{k-1}(c) > F_k(c)$, $\forall k$, involving that $\Pi_i - \Pi_{i+1} = \int_{\underline{c}}^{\overline{c}} y_{i+1}(c)(F_i(c) - F_{i+1}(c)) dc > 0, \ \forall i < 4$. Hence, an agent of type i has no interest in mimicking a type above $i+1$ (if any) because that lie would grant him a lower gain than would be obtained by claiming $i+1$. Moreover, provided that the quantity $y_i(c)$ is monotonic with respect to i, an agent of type $i+m$ would face a loss of $\int_{\underline{c}}^{\overline{c}} (y_{i+m}(c) - y_i(c))(F_i(c) - F_{i+m}(c)) dc$ if he were to pretend i, and hence he has no incentives

to do so. This is the formal argument that Courty and Li [9] use to prove that "upward" incentive constraints are satisfied.

(2) **No private information about the expected cost**   In the same setting as above, take four distributions indexed by $j \in \{1, 2, 3, 4\}$ with equal mean θ. Further take $\int_{\underline{c}}^{c} F_j(c) \, dc$ to be decreasing in j, which means that any distribution j is a mean-preserving spread of the distribution $j + 1$. Assuming that h_j is monotonic and that the optimal quantity profile $y_j(c)$ is monotonic in c (thus satisfying the second-period monotonicity conditions), the solution is formulated as in benchmark (1). However, resting on the property of equal mean across distributions, involving that $\int_{\underline{c}}^{\overline{c}} F_{k-1}(x) \, dx = \int_{\underline{c}}^{\overline{c}} F_k(x) \, dx, \forall k$, the information rent is now also written as:

$$\Pi_j = \sum_{k=j+1}^{4} \int_{\underline{c}}^{\overline{c}} (-y'_k(c)) \left(\int_{\underline{c}}^{c} F_{k-1}(x) \, dx - \int_{\underline{c}}^{c} F_k(x) \, dx \right) dc, \ \forall j \in \{1, 2, 3\}, \ \text{ and } \ \Pi_4 = 0.$$

Under second-order stochastic dominance, $\int_{\underline{c}}^{c} F_{k-1}(x) \, dx > \int_{\underline{c}}^{c} F_k(x) \, dx, \forall k$; furthermore, second-period monotonicity requires $y'_k(c) < 0$. It follows that $\Pi_j - \Pi_{j+1} > 0, \forall j < 4$, where $\Pi_j - \Pi_{j+1} = \int_{\underline{c}}^{\overline{c}} (-y'_{j+1}(c)) \left(\int_{\underline{c}}^{c} F_j(x) \, dx - \int_{\underline{c}}^{c} F_j(x) \, dx \right) dc$. As in benchmark (1), an agent of type j has no interest in mimicking a type above $j + 1$. Besides, an agent of type $j + m$ has no incentive to pretend j because that would yield a loss of $\int_{\underline{c}}^{\overline{c}} (-y'_j(c)) \left(\int_{\underline{c}}^{c} F_j(x) \, dx - \int_{\underline{c}}^{c} F_{j+m}(x) \, dx \right) dc$. Therefore, also in this case upward incentive constraints are satisfied, as in the proof developed by Courty and Li [9].

($1'$) **No private information about the spread**   Take $c = \theta + \varepsilon \sigma$ as in our setting. Assume that the expected costs are such that $\theta_1 < \theta_2 < \theta_3 < \theta_4$, whereas the spread σ is equal across distributions and commonly known. Also assume that $\theta_4 - \sigma \leq \theta_1 + \sigma$, implying that $\theta_{i'} - \sigma \leq \theta_i + \sigma, \forall i, i'$. Under this assumption and the monotonicity of the hazard rate h_i, the solution is characterized as follows:

$$S'(y_1(\varepsilon)) = \theta_1 + \varepsilon \sigma \ \text{ and } \ S'(y_i(\varepsilon)) = \theta_i + \varepsilon \sigma + h_i(\theta_i - \theta_{i-1}), \ \forall i \in \{2, 3, 4\}$$

$$\Pi_i = \sum_{k=i+1}^{4} (\theta_k - \theta_{k-1}) q_k, \ \forall i \in \{1, 2, 3\}, \ \text{ and } \Pi_4 = 0.$$

($2'$) **No private information about the expected cost**   Take again $c = \theta + \varepsilon \sigma$ as in our setting. Assume that the four distributions are characterized by an equal mean of θ and different spreads σ_j, where $j \in \{1, 2, 3, 4\}$ is again the agent's type, such that:

$$\sigma_1 > \sigma_2 > \sigma_3 > \sigma_4.$$

Denote μ_j the commonly known probability of the agent's type being j, the hazard rate $g_j = \sum_{k=1}^{j-1} \frac{\mu_k}{\mu_j}, \forall j \in \{2, 3, 4\}$, and $g_1 = 0$. If g_j is monotonically increasing in j and the second-period

10

monotonicity condition $y_j(-1) \geq y_j(1)$ is not binding, then the solution is given by:

$$S'(y_1(\varepsilon)) = \theta + \varepsilon\sigma_j \text{ and } S'(y_j(\varepsilon)) = \theta + \varepsilon\sigma_j - \varepsilon g_j(\sigma_{j-1} - \sigma_j), \ \forall j \in \{2,3,4\}$$

$$\Pi_j = \sum_{k=j+1}^{4} (\sigma_{k-1} - \sigma_k) r_k, \ \forall j < 4, \text{ and } \Pi_4 = 0.$$

Overall inspection of the four cases above highlights that the solution displays analogous features regardless of whether the unit cost takes values in a continuous interval that does not vary across distributions, the usual case in sequential screening models, or it takes values in a discrete set that differs across distributions. Both in benchmark (1) and (1') production levels are decreased below the efficient levels for all types but $i = 1$, higher distortions being associated with higher expected costs. Moreover, the higher the expected cost the lower the information rent accruing to the agent. Both in benchmark (2) and (2') production levels are distorted for all types but the highest-spread type, distortions being downwards for good shocks, upwards for bad shocks, and greater the lower the value of the spread. Moreover, the information rent accruing to the agent is lower the smaller the spread is. In substance, regardless of whether the agent's private information concerns the expected cost or the spread, the solution is such that greater distortions are induced for less "efficient" types (namely, types with higher expected cost and lower spread) and more efficient types are assigned higher information rents. A similar outcome is also obtained in the setting to be presented next, in which screening is multidimensional, as in our framework, but not sequential.

(3) **The two pieces of information as the unit costs of two activities** Assume that θ_i and σ_j, where $i, j \in \{1, 2, 3, 4\}$, represent the unit costs of operating two distinct activities, rather than being the two parameters of a cost distribution. As Armstrong and Rochet [3] show, when θ_i and σ_j are uncorrelated and the monotonic hazard rate property is satisfied for each activity, the solution is "twice a replica" of the standard solution à la Baron and Myerson [5].

Proposition 1 (*Riordan and Sappington [33], Courty and Li [9], Armstrong and Rochet [3]*) *The solutions in benchmarks* (1), (2) *and* (3) *are such that local incentive constraints imply global incentive constraints. Distortions in production levels are greater for less efficient types, whereas information rents are greater for more efficient types.*

This result is similar to those obtained in static screening problems à la Baron and Myerson [5]. Sequential and static problems are alike in this respect.[7] Recall now that in benchmark

---

[7] Notice that in the static screening problems we refer to the incentive mechanism is deterministic. In a recent manuscript, Krahmer and Strausz [25] show that when the incentive mechanism in the static problem is stochastic, the sequential problem à la Courty and Li [9] and the static problem are equivalent, not only "alike" in the sense of Proposition 1.

($1'$) this result was obtained under the assumption that $\theta_{i'} - \sigma \leq \theta_i + \sigma$, $\forall i, i'$. When this condition is violated, the following result holds.

Proposition 2 *If there exist $i' > i$ such that $\theta_{i'} - \sigma > \theta_i + \sigma$ in benchmark ($1'$), then type i' is assigned the same level of output regardless of the shock: $y_{i'}(\varepsilon)$ is constant $\forall \varepsilon$.*

This result is reminiscent of that derived by Krahmer and Strausz [23]. When $\theta'_i - \sigma > \theta_i + \sigma$ and type i announces $i' > i$ in the first stage, it anticipates that, following this lie, it will obtain a benefit in the second stage if it pretends a good shock when the true shock is bad. That benefit, which is equal to $(\theta_{i'} - \theta_i)(y_{i'}(-1) - y_{i'}(1))$, will then add to the potential gain from cost overstatement in the first stage. Therefore, type i will be motivated to lie again in the second stage after lying in the first stage. Specifically, it will be eager to *understate* the shock on the cost realization in the second stage after *overstating* the expected cost in the first stage. This means that it displays countervailing incentives with respect to the reports to be delivered in the two stages. It it thus not surprising that an inflexible output rule must be imposed on type i' ($y_{i'}(-1) = y_{i'}(1)$) to make cheating unattractive for type i, in the same vein as in the traditional agency models with countervailing incentives (Lewis and Sappington [28], Maggi and Rodriguez-Clare [29]). The specificity here is that countervailing incentives arise with respect to the information reported in distinct periods.

Let us now compare the screening problem we analyze with the benchmarks. Unlike in benchmarks $(1) - (2)$ and in line with benchmarks $(1') - (2')$, the set of possible cost values is discrete and depends on the first-stage distribution.[8] Moreover, the cost distributions cannot be ranked according to a specific stochastic ordering, and hence the reasoning used by Courty and Li [9] (*i.e.*, that in benchmarks $(1)-(2)$) cannot be followed with the purpose of identifying the binding incentive constraints. Lastly, the fact that θ and σ are the parameters of a cost distribution, rather than the unit costs of two distinct activities as in benchmark (3), complicates the screening problem because the incentives to misrepresent the two pieces of information are linked, provided that they both depend on the output profiles in the distribution they belong to. These divergences of our problem from the common setting in the literature will lead to a divergence of our findings from Proposition 1. In this respect, it is also useful to point out that the results in Proposition 1 are obtained regardless of the shape of the marginal surplus function $S'(\cdot)$. Actually, in Courty and Li [9] and in most recent studies the marginal valuation of the traded good is taken to be constant. By contrast, in our model $S'(\cdot)$ is decreasing in the quantity and can have any shape. Allowing for any shape of $S'(\cdot)$ will enable us to show that the specific shape of the marginal surplus has an essential impact on the determination of results. Before turning to that, we need to consider that, in line with Proposition 2, it might be optimal to bunch some of the output levels. In our framework, this entails that sequential screening is only *partial*, as we explain below.

---

[8] We show in Corollary 2 that the problem becomes trivial when the range of cost values is continuous. Moreover, cost values depend on the distribution from which they are drawn because distributions are specified by both mean and spread.

4 Partial sequential screening

We previously mentioned that the main difficulty in solving our screening problem is the identification of the binding incentive constraints. The first step consists in selecting the downward incentive constraints, whereby more efficient types are not tempted to pretend to be less efficient. The efficiency of a given type depends on its expected total cost of production. That is, the lower the expected total cost is for some given production level defined by P, the more efficient the type is. To express the expected total cost for a given production level, we let $q \equiv \frac{1}{2}(y(-1) + y(1))$ and $r \equiv \frac{1}{2}(y(-1) - y(1))$ be, respectively, the expected production level and the expected wedge between production levels when the output profile is $\{y(-1), y(1)\}$. Then, the expected total cost is given by $\theta q - \sigma r$; thus it increases with θ and, when $r > 0$, it decreases with σ. The reason why P is better off with cost distributions characterized by a high spread is that, when facing more uncertainty in the contracting stage, P benefits from an option value if she requires greater production for low cost realizations than for high cost realizations $(y(-1) > y(1))$. Therefore, a type with lower expected cost and/or higher spread is more efficient than other types. This leads to the following type ordering:

$$LH \succ LL \succ HH \succ HL \tag{2}$$

when $\Delta\theta > \Delta\sigma$, and to:

$$LH \succ HH \succ LL \succ HL \tag{3}$$

otherwise. Whereas LH and HL are unequivocally the best and the worst type, the order of intermediate types LL and HH depends on how important the degree of private information about the expected cost is relative to the degree about private information about the spread.

Considering now the real cost values, there are more than two possible orderings, depending on whether or not there exist types ij and $i'j'$ such that $\theta_i + \sigma_i < \theta_{i'} - \sigma_{j'}$. This is an important point because, when the previous inequality holds, countervailing incentives arise and P prefers to not screen type $i'j'$ in the second stage, in the same vein as in benchmark $(1')$. There are two cases of *partial* sequential screening. One concerns type HL only; the other concerns both type HL and type HH. To avoid redundancy with little additional insight, we restrict attention to the latter, more interesting case.

Proposition 3 *If $\theta_L + \sigma_j < \theta_H - \sigma_{j'} \ \forall j, j'$, then type Hj' is assigned the same production level $y_{Hj'}(\varepsilon) = y_H, \ \forall j', \ \forall \varepsilon$.*

To explain this result, it is first useful to present the solution to the problem of P. The quantities are specified as follows:

$$
\begin{aligned}
S'(y_{LH}(\varepsilon)) &= \theta_L + \varepsilon\sigma_H, \ \forall \varepsilon \\
S'(y_{LL}(\varepsilon)) &= \theta_L + \varepsilon\sigma_L - \frac{1-\mu}{\mu}\varepsilon\Delta\sigma, \ \forall \varepsilon \\
S'(y_H) &= \theta_H + \frac{\nu}{1-\nu}\Delta\theta.
\end{aligned}
$$

13

The information rents are given by $\Pi_{LH} = \Delta\sigma r_{LL} + \Delta\theta y_H$, $\Pi_{LL} = \Delta\theta y_H$ and $\Pi_{HL} = \Pi_{HH} = 0$. In line with Proposition 2, bunching of quantities is induced *within* some distributions, which are here HH and HL. In addition, Proposition 3 shows that bunching is also induced *between* those two distributions. Interestingly, whereas bunching within distributions is due to the presence of countervailing incentives, bunching between distributions is of a different nature. It is a way of limiting distortions that are not helpful for P to fine-tune rent extraction. To see this, consider that, once P decides to offer types HH and HL the output levels y_{HH} and y_{HL}, set independently of the shock realization to prevent coordinated lies between stages, the expected total costs are simply $\theta_H y_{HH}$ and $\theta_H y_{HL}$. By mimicking any of the two types HH and HL, type LL is able to benefit from information about θ only. If P decreases y_{HH} below y_{HL} to contain the rent that type LL receives for not claiming HH, then it becomes more convenient for type LL to claim HL. Similarly, if P decreases y_{HL} below y_{HH} to contain the rent that type LL obtains for not claiming HL, then it becomes more convenient for type LL to claim HH. Therefore, decreasing any of the two quantities y_{HH} and y_{HL} below the other is not a successful strategy to limit agency costs. The best for P is to require types HH and HL to deliver the same amount of the good $(y_{HH} = y_{HL} = y_H)$.

The general characteristics of the optimal contract are indicative of what one should expect in terms of cost overruns, provided that overruns can only follow an understatement of θ and/or of σ.

Corollary 1 *Under Proposition 3, no rent is conceded to prevent understatement of θ. A rent designed to prevent understatement of σ is conceded only to type LH, which has incentives to mimic type LL.*

Therefore, the incentives to lie that would lead to cost overruns are related to an understatement of the spread but not of the expected cost. This result does not look surprising, if it is considered that a σ_H-type is *efficient* in the second information dimension whereas a θ_H-type is *ine*fficient in the first information dimension. Less intuitive is probably the fact that concerns about an understatement of the spread arise only when P faces an agent with low expected cost. In light of Proposition 3, the reason for this is that type LH would have little to gain from an exaggeration of θ when θ_H-types are required to deliver the same amount of the good.

Bunching of quantities within distributions would be induced if the shock were drawn from a continuous interval rather than from a binary set.

Corollary 2 *If $\varepsilon \in [-1, 1]$, then for all $ij \neq LH$: $y_{ij}(\varepsilon) = y_{ij}$, $\forall\varepsilon$.*

This entails that, with a continuous support, screening would no longer take place in two stages and P would face a static incentive problem. Actually, this is what motivates us to focus on a discrete support for the development of our analysis.

5 Full sequential screening

We now turn to consider situations in which $\theta_L + \sigma_j \geq \theta_H - \sigma_{j'}\ \forall j, j'$. Then, the agent would have no convenience to lie again in the second stage, after lying in the first stage, so that the incentive constraints are reformulated as:

$$IC_{ij}^{i'j'} : \ \Pi_{ij} \geq \Pi_{i'j'} + (\theta_{i'} - \theta_i)\, q_{i'j'} - (\sigma_j - \sigma_{j'})r_{i'j'}, \ \forall ij, i'j' \in \Upsilon.$$

The screening problem is now fully sequential and it is more complicated to identify the binding incentive constraints. We proceed as follows. As usual in mechanism design, we first consider a reduced problem in which upward incentive constraints, whereby less efficient types are unwilling to pretend to be more efficient, are omitted. The conditions under which the solution to that reduced problem is also a solution to the general problem are stated in the following lemma.

Lemma 1 With $\theta_i + \varepsilon\sigma_j \leq \theta_{i'} + \varepsilon\sigma_{j'}, \ \forall ij, i'j', \varepsilon$, the solution to the reduced problem including the downward incentive constraints is also a solution to Γ if the production levels other than $y_{LH}(1)$ and $y_{HL}(1)$ are such that $y_{ij}(\varepsilon) \geq y_{i'j'}(\varepsilon)$, whereas $y_{LH}(1)$ and $y_{HL}(1)$ are such that:

$$q_{LH} \geq q_{HL} \quad and \quad r_{LH} \geq r_{HL}. \tag{4}$$

Proceeding as in static agency problems (Baron and Myerson [5]) and in sequential screening problems (Courty and Li [9]), to find the solution we first assume that the profile of optimal quantities respects the first-best order. The lemma shows that this quantity order involves as usual that upward incentive constraints are implied by downward incentive constraints. When we characterize the contract, we will show that the solution of the reduced problem is such that the conditions in the lemma are satisfied. The exception in the lemma that $y_{HL}(1)$ and $y_{LH}(1)$ can take any order is due to opposite effects that a change in the quantity associated with a bad shock has on incentives to misrepresent the two information dimensions. Hence, we do not require that $y_{HL}(1) \leq y_{LH}(1)$ because it has no implication for the lemma. Conditions (4) are implied by the incentive constraints included in the reduced problem.

It is essential to point out that looking at a reduced problem that includes the sole downward incentive constraints is not all. Actually, there are different possible combinations of binding incentive constraints, reflecting the fact that it might be optimal to induce bunching. Accordingly, two classes of solutions are identified (corresponding to two classes of reduced problems), depending on whether bunching or full separation entails at optimum.

Lemma 2 $\exists \delta > 1$ such that, if $\Delta\theta > \delta\Delta\sigma$, then the solution is such that $y_{HH}(\varepsilon) = y_{HL}(\varepsilon)$, $\forall \varepsilon$. If $\Delta\theta \leq \Delta\sigma$, then the solution is such that $y_{LL}(\varepsilon) = y_{HL}(\varepsilon)$, $\forall \varepsilon$.

Bunching of quantities between distributions has a similar explanation as in Proposition 3. Indeed, it arises as a way of limiting distortions that do not help P contain information rents,

rather than reflecting binding monotonicity conditions as in the presence of countervailing incentives. When $\Delta\theta > \Delta\sigma$, types HH and HL are treated as being one type to discourage type LL from exaggerating θ. When $\Delta\theta < \Delta\sigma$, instead, types LL and HL are treated as being one type to discourage type HH from understating σ. In either case, the agent's incentive to cheat is eliminated by bunching quantities with respect to the information dimension that is less important, namely σ in the former case and θ in the latter. However, unlike in Proposition 3, P screens the agent in two stages. Therefore, quantities are equal two by two only, namely $y_{HH}(-1) = y_{HL}(-1)$ and $y_{HH}(1) = y_{HL}(1)$, and the solution is separating within the distributions. It is noteworthy that bunching of quantities as a way of limiting unnecessary distortions is unusual not only in sequential screening problems (as previously mentioned), but also in multidimensional settings. As an illustration, Armstrong [2] considers a setting with private information about both the cost and the demand faced by the agent and shows that, because there is only one instrument (the output level) to screen two information dimensions, bunching is induced to ensure the monotonicity of the output profile, in the same vein as in the literature on countervailing incentives.

Once it is understood that there are distinct classes of solutions to the reduced problem, depending on whether bunching or full separation entails at optimum, it might look natural to turn to the characterization of the various admissible solutions. However, identifying the combination of binding incentive constraints for each of the reduced problems is a complex task in that, unlike in sequential screening models, neither (2) nor (3) respects a specific stochastic order. Moreover, at odds with Armstrong and Rochet [3], q_{ij} and r_{ij} cannot be viewed as the volumes of two distinct activities in that they rather represent two interdependent decision variables pertaining to one distribution. In this context, each reduced problem may admit multiple combinations of binding incentive constraints.

In light of the foregoing, instead of plunging directly into a tedious characterization of the solution, it is more instructive to identify what exactly determines the agent's incentives that are relevant for P and, more specifically, to which information dimension they are related. As will soon become apparent, this will be key to conveniently narrowing the set of incentive constraints to be considered when looking at specific applications. To that end, the properties of concave/convex functions can be used to pinpoint the relationship between the preferences of the principal for the good and the relevant incentives of the agent to misrepresent information.

Denote the inverse function of $S'(\cdot)$ as $f(\cdot)$. Also denote $S'(y_{ij}(-1)) \equiv \underline{a}_{ij}$ and $S'(y_{ij}(1)) \equiv \overline{a}_{ij}$, $\forall ij$, with reference to the solution to the reduced problem. Then, the following equivalences hold for $x = \underline{a}_{i'j'} - \underline{a}_{ij}$:

$$
\begin{aligned}
f(\underline{a}_{ij}) - f(\underline{a}_{i'j'}) &\geq f(\overline{a}_{ij}) - f(\overline{a}_{i'j'}) \\
&\Leftrightarrow (f(\underline{a}_{ij}) - f(\underline{a}_{ij} + x)) - (f(\overline{a}_{ij}) - f(\overline{a}_{ij} + x)) \geq f(\overline{a}_{ij} + x) - f(\overline{a}_{i'j'})
\end{aligned}
$$

and

$$f(\underline{a}_{ij}) - f(\underline{a}_{i'j'}) \geq f(\overline{a}_{i'j'}) - f(\overline{a}_{ij})$$
$$\Leftrightarrow (f(\underline{a}_{ij}) - f(\underline{a}_{ij} + x)) - ((f(\overline{a}_{i'j'}) - f(\overline{a}_{i'j'} + x)) \geq f(\overline{a}_{i'j'} + x) - f(\overline{a}_{ij}).$$

These inequalities are just equivalent formulations of the inequalities $r_{ij} \geq r_{i'j'}$ and $q_{ij} \geq q_{i'j'}$ at the solution to the problem. By considering the relationship between each of the differences $(f(\underline{a}_{ij}) - f(\underline{a}_{ij} + x)) - (f(\overline{a}_{ij}) - f(\overline{a}_{ij} + x))$ and $(f(\underline{a}_{ij}) - f(\underline{a}_{ij} + x)) - ((f(\overline{a}_{i'j'}) - f(\overline{a}_{i'j'} + x))$ on one side, and the degree of concavity/convexity of $S'(\cdot)$ on the other, one can verify whether those inequalities hold or they are violated for different values of $S'''(\cdot)$. To that end, we define:

$$\zeta(z_1, z_2, x) \equiv (f(z_1) - f(z_1 + x)) - (f(z_2) - f(z_2 + x)), \ \forall z_1, z_2, x, \ z_1 < z_2, \ x > 0$$

so that we can reformulate the equivalences listed above as:

$$f(\underline{a}_{ij}) - f(\underline{a}_{i'j'}) \geq f(\overline{a}_{ij}) - f(\overline{a}_{i'j'}) \tag{5}$$
$$\Leftrightarrow \zeta(\underline{a}_{ij}, \overline{a}_{ij}, \underline{a}_{i'j'} - \underline{a}_{ij}) \geq f(\overline{a}_{ij} + \underline{a}_{i'j'} - \underline{a}_{ij}) - f(\overline{a}_{i'j'})$$

and

$$f(\underline{a}_{ij}) - f(\underline{a}_{i'j'}) \geq f(\overline{a}_{i'j'}) - f(\overline{a}_{ij}) \tag{6}$$
$$\Leftrightarrow \zeta(\underline{a}_{ij}, \overline{a}_{i'j'}, \underline{a}_{i'j'} - \underline{a}_{ij}) \geq f(\overline{a}_{i'j'} + \underline{a}_{i'j'} - \underline{a}_{ij}) - f(\overline{a}_{ij})$$

and then state the following lemma.

Lemma 3 $\zeta(z_1, z_2, x) > 0$ if $S'''(\cdot) > 0$; $\zeta(z_1, z_2, x) = 0$ if $S'''(\cdot) = 0$; $\zeta(z_1, z_2, x) < 0$ if $S'''(\cdot) < 0$.

This result is useful to verify how tight the different incentive constraints are, and hence to identify those that are slack in the problem, depending on the curvature of $S'(\cdot)$, which is expressed by the value of $\zeta(\cdot, \cdot, \cdot)$. By comparing how tight the incentive constraints are for different degrees of concavity/convexity of $S(\cdot)$, we draw the following result.

Proposition 4 *Assume that the conditions in Lemma 1 hold. At the solution to Γ, as $\zeta(\cdot, \cdot, \cdot)$ increases, the information rent accruing to the agent is increasingly more related to the possibility of misrepresenting higher-order types relative to lower-order types.*

Unlike in the benchmarks, the combination of binding incentive constraints depends on the preferences of the principal for the good. This result is unusual in sequential screening problems but similar in nature to the finding of the consumption theory under uncertainty. Specifically, in consumption models, the trade-off between consumption and savings faced by an individual who takes the decision in conditions of uncertainty depends on the shape of the individual's marginal utility function. In our screening model, the characteristics of the optimal

delegation contract depend on the shape of the principal's marginal surplus function, provided that the unit production cost is unknown at the time when the contract is drawn up. The exact implications of this result are stated in the corollaries below, whereas further considerations in light of the consumption models are left for a later discussion.

Corollary 3 *Under Proposition* 4, *as* $\zeta(\cdot, \cdot, \cdot)$ *increases, the principal's concerns are increasingly more related with the possibility of the agent misrepresenting* σ *rather than* θ *when* $\Delta\theta > \Delta\sigma$, *and with the possibility of the agent misrepresenting* θ *rather than* σ *when* $\Delta\theta \leq \Delta\sigma$.

There is no solution where screening reduces to a unidimensional screening problem. Not surprisingly, both θ and σ should be reported by the agent. What one can say is whether the principal should be more concerned with the possibility that the mean value is misrepresented, or the spread. The corollary shows that the "concerns" of the principal actually depend on her own preferences. Specifically, as the proof in Appendix C.4 shows, when $\Delta\theta > \Delta\sigma$, as $\zeta(\cdot, \cdot, \cdot)$ increases, IC_{LL}^{HH} becomes tighter than IC_{LL}^{HL}, IC_{LH}^{HH} and IC_{LH}^{LL} both become tighter than IC_{LH}^{HL}, and IC_{LH}^{LL} becomes tighter than IC_{LH}^{HH}. This transition reflects an increasing concern of P with the possibility of the agent misrepresenting σ rather than θ. When $\Delta\theta \leq \Delta\sigma$, as $\zeta(\cdot, \cdot, \cdot)$ takes higher values, IC_{LH}^{HH} becomes tighter relative to both IC_{LH}^{LL} and IC_{LH}^{HL}, which reflects an increasing concern with the possibility that the agent is misrepresenting θ rather than σ. This is tantamount to saying that when $S'(\cdot)$ shifts from concave to convex the incentive constraints related to adjacent types are more likely to be binding.

Corollary 4 *Under Proposition* 4, *local incentive constraints imply global incentive constraints if and only if* $\zeta(\cdot, \cdot, \cdot)$ *is positive and sufficiently high.*

It is interesting to compare this result with the studies on sequential screening. When $S'(\cdot)$ is sufficiently convex, each of the two type orderings plays a similar role in our setting to the stochastic ordering in those studies. Then, the can focus on the local incentive constraints only. This is not true otherwise, particularly when the marginal surplus is linear, as is usually assumed in sequential screening problems.

Let us now return to the possibility of observing cost overruns and to the incentives in which they are rooted. As in the case of partial sequential screening, it turns out that cost overruns only mirror incentives to understate σ, whereas concerns about misrepresentation of θ are negligible.

Corollary 5 *Under Proposition* 4, *a rent designed to prevent understatement of* θ *is conceded only to type* HH, *which has incentives to understate both* θ *and* σ *(to pretend* LL). *That rent accrues only when* $\Delta\theta \leq \Delta\sigma$ *and it is equal to that conceded to type* HH *to not understate* σ *only (to not pretend* HL).

6 A characterization of the optimal contract under full sequential screening

We now illustrate what the optimal contract looks like in the trickier case to solve, that in which $\Delta\theta > \Delta\sigma$. Albeit the characterization of the solution is rather lengthy, it helps us clarify how useful it is to refer to the relationship between the principal's preferences and the agent's incentives to identify the constraints that are binding at optimum, as previously stated in Proposition 4. The mathematical development is reported in Appendix D.

When the type order in (2) applies, the information rents are expressed as follows:

$$\Pi_{HL} = 0 \tag{7a}$$

$$\Pi_{HH} = \Delta\sigma r_{HL} \tag{7b}$$

$$\Pi_{LL} = \beta\Pi_{LL,1} + (1-\beta)\Pi_{LL,2} \tag{7c}$$

$$\Pi_{LH} = \gamma_1\Pi_{LH,1} + \gamma_2\left[\beta\Pi_{LH,2} + (1-\beta)\Pi_{LH,3}\right] + \gamma_3\Pi_{LH,4}, \tag{7d}$$

where:

$$\Pi_{LL,1} = \Delta\theta q_{HL} \quad\text{and}\quad \Pi_{LL,2} = \Delta\theta q_{HH} - \Delta\sigma\left(r_{HH} - r_{HL}\right) \tag{8}$$

together with:

$$\Pi_{LH,1} = \Delta\theta q_{HH} + \Delta\sigma r_{HL}, \quad \Pi_{LH,2} = \Delta\theta q_{HL} + \Delta\sigma r_{LL}, \tag{9}$$

$$\Pi_{LH,3} = \Delta\theta q_{HH} + \Delta\sigma r_{HL} - \Delta\sigma\left(r_{HH} - r_{LL}\right) \quad\text{and}\quad \Pi_{LH,4} = \Delta\theta q_{HL} + \Delta\sigma r_{HL},$$

where $\beta \in [0,1]$, $\gamma_z \in [0,1]$ and $\sum_{z=1}^{3}\gamma_z = 1$.

Identifying the exact downward incentive constraints that are relevant for the solution to Γ - thus, the exact information rents to be conceded - is tantamount to identifying the values of the parameters β and γ_z, $\forall z \in \{1,2,3\}$. If only local incentive constraints were relevant, then only the case where $\beta = 0$ and $\gamma_2 = 1$ would have to be considered. Using the expressions of the information rents listed above and focusing on the case in which the optimal output profile is fully separating, we identify the solution to a first reduced problem, denoted Γ'.

Quantity solution to the reduced problem Γ'. For the solution to Γ', for all ε, type LH

produces the first-best quantity $y_{LH}^*(\varepsilon)$; for the other types quantities are such that:

$$S'(y_{LL}(\varepsilon)) = \theta_L + \varepsilon\sigma_L - \gamma_2\frac{1-\mu}{\mu}\varepsilon\Delta\sigma \tag{10}$$

$$S'(y_{HH}(\varepsilon)) = \theta_H + \varepsilon\sigma_H \tag{11}$$
$$+ \frac{\nu}{1-\nu}\left\{\left[\gamma_1 + (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\right]\Delta\theta + (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\varepsilon\Delta\sigma\right\}$$

$$S'(y_{HL}(\varepsilon)) = \theta_H + \varepsilon\sigma_L \tag{12}$$
$$+ \frac{\nu}{1-\nu}\left\{\left[\beta + (\gamma_2\beta + \gamma_3)\frac{1-\mu}{\mu}\right]\Delta\theta - \left[1-\beta + (1-\nu\gamma_2\beta)\frac{1-\mu}{\nu\mu}\right]\varepsilon\Delta\sigma\right\}.$$

The quantity solution above does not satisfy the incentive constraints in Γ when $\Delta\theta \geq \delta\Delta\sigma$, where:

$$\delta = \frac{\frac{1}{\nu} + (\mu - \beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)}{\left|1 - \gamma_1 - (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\right|}. \tag{13}$$

A second reduced problem, denoted Γ'', must be considered, which is solved by requiring types HL and HH to produce the same output: $y_{HL}(\varepsilon) = y_{HH}(\varepsilon), \forall\varepsilon$.

Quantity solution to the reduced problem Γ''. For the solution to Γ'', for all ε, types LH and LL produce the first-best quantities $y_{LH}^*(\varepsilon)$ and $y_{LL}^*(\varepsilon)$; for the other types, quantities are such that:

$$y_{HL}(\varepsilon) = y_{HH}(\varepsilon)$$
$$S'(y_{HH}(\varepsilon)) = \theta_H + \varepsilon\sigma_H + \frac{\nu}{(1-\nu)(1-\mu)}\Delta\theta - \frac{1-\gamma_2\nu}{1-\nu}\varepsilon\Delta\sigma. \tag{14}$$

Resting on these solutions, and defining:

$$\zeta_1 \equiv \frac{\Delta\theta + \Delta\sigma}{\Delta\theta - \Delta\sigma}(f(\bar{a}_{HL}) - f(\bar{a}_{HH})) - [f(\bar{a}_{HL}) - f(\bar{a}_{HL} + \underline{a}_{HL} - \underline{a}_{HH})] \tag{15}$$

$$\varepsilon_H \equiv f(\bar{a}_{HL} + \underline{a}_{HL} - \underline{a}_{HH}) - f(\bar{a}_{HH}) \tag{16}$$

$$\varepsilon_L \equiv f(\bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL}) - f(\bar{a}_{HH}), \tag{17}$$

we can identify the essential features of the optimal contract.

Proposition 5 *Assume that $y_{ij}(-1) > y_{ij}(1), \forall ij$, and that $\Delta\theta > \Delta\sigma$.*
(i) If $\Delta\theta < \delta\Delta\sigma$, then the solution to Γ satisfying the conditions in Lemma 1 is such that:

- *$\beta = 1$ when $\zeta(\underline{a}_{HH}, \bar{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) < \zeta_1$;*

- $\gamma_3 = 1$ and $\gamma_1 = \gamma_2 = 0$ when $\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) < \varepsilon_H$, where $\varepsilon_H < 0$;

- $\gamma_3 = 0$, $\gamma_1 \geq 0$, $\gamma_2 > 0$ when $\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) \geq \varepsilon_H$;

- $\gamma_3 = \gamma_1 = 0$, $\gamma_2 = 1$ when $\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{HH} - \underline{a}_{LL}\right) \geq \varepsilon_L$, where $\varepsilon_L > 0$.

(*ii*) *If* $\Delta\theta \geq \delta\Delta\sigma$, *then the solution to* Γ *that satisfies the conditions in Lemma 1 is such that:*

- $\gamma_1 = 1$ when $\zeta\left(\underline{a}_{LL}, \overline{a}_{HH}, a_{HH} - \underline{a}_{LL}\right) < \varepsilon_L$, where $\varepsilon_L < 0$;

- $\gamma_2 = 1$ when $\zeta\left(\underline{a}_{LL}, \overline{a}_{HH}, a_{HH} - \underline{a}_{LL}\right) > \varepsilon_L$, where $\varepsilon_L > 0$.

Proposition 5 confirms that the general problem can be replaced by an appropriate relaxed problem, which indicates that the methodology adopted to solve multidimensional screening problems with a discrete number of types also applies to sequential screening problems.

Corollary 6 *Information rents and production levels in Proposition 5 reflect:*
(*a*) *incentives of type LL to overstate* θ $\forall\zeta\left(\cdot,\cdot,\cdot\right)$, *and to overstate* σ *if and only if:*

$$\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) \geq \zeta_1 \tag{18}$$

when (13) *holds, and* $\forall\zeta\left(\cdot,\cdot,\cdot\right)$ *otherwise;*
(*b*) *incentives of type LH to overstate* θ *if and only if:*

$$\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) \leq \varepsilon_L, \tag{19}$$

and to understate σ *if and only if:*

$$\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) \geq \varepsilon_H, \tag{20}$$

when (13) *holds, and if and only if:*

$$\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) \geq \varepsilon_L$$

otherwise.

This corollary is an application of the result contained in Proposition 4 that as $\zeta\left(\cdot,\cdot,\cdot\right)$ takes higher values, the information rents are more closely related to the possibility of σ being misrepresented instead of θ. In particular, exaggeration of the expected cost by a low-expected-cost high-spread agent is not an issue for P when $S'\left(\cdot\right)$ is sufficiently convex ((19) violated). In that situation, indeed, P prefers to set more dispersed production levels for each cost distribution. With this output profile, an understatement of σ becomes more worrisome than an overstatement of θ.

In Proposition 5, it is also interesting to identify the ranges of values of $\zeta(\cdot, \cdot, \cdot)$ over which local incentive constraints imply global incentive constraints, as is usually the case in sequential screening frameworks.

Corollary 7 *The optimal contract reflects the circumstance that local incentive constraints imply global incentive constraints if and only if:*

$$\zeta(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) \geq \zeta_1$$

together with:

$$\zeta(\underline{a}_{LL}, \overline{a}_{HH}, a_{HH} - \underline{a}_{LL}) \geq \varepsilon_L$$

for $\beta = 0$ and $\gamma_2 = 1$.

The solution does not reduce to that of a sequential screening problem with a single piece of information (Riordan and Sappington [33] and Courty and Li [9]) even when $S'(\cdot)$ is sufficiently convex. However, some similarities are detected. When $S'(\cdot)$ is sufficiently convex, only the highest-order type is assigned the first-best trade volume for each possible realization of the unknown variable. The middle-order types LL and HH represent an attractive lie for some higher-order type and are both assigned distorted productions. When $S'(\cdot)$ is little convex or concave instead, this occurs for one such type only (either LL or HH) because in that case type LH is more attracted by a nonadjacent type.

To complete the characterization of the contract, we have to verify that the quantity solution to Γ' or to Γ'' satisfies the conditions in Lemma 1, so that it also solves the general problem. We will use the following definitions:

$$
\begin{aligned}
\varepsilon_{LH} &\equiv f(\overline{a}_{LH} + \underline{a}_{HL} - \underline{a}_{LH}) - f(\overline{a}_{HL}) \\
\varepsilon'_{LH} &\equiv f(\overline{a}_{HL} + \underline{a}_{HL} - \underline{a}_{LH}) - f(\overline{a}_{LH})
\end{aligned}
$$

Proposition 6 *For the solution to Γ' or to Γ'', $r_{ij} > 0$ $\forall ij$, if:*

$$\frac{\sigma_H}{\sigma_L} < \frac{1}{1 - \mu(1 - \nu)}. \tag{21}$$

The solution to Γ' or to Γ'' satisfies the quantity ranking in Lemma 1. The conditions in (4) are satisfied in both Γ' and Γ'' if and only if

$$\zeta(\underline{a}_{LH}, \overline{a}_{LH}, \underline{a}_{HL} - \underline{a}_{LH}) \geq \varepsilon_{LH} \tag{22}$$

and

$$\zeta(\underline{a}_{LH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{LH}) \geq \varepsilon'_{LH} \tag{23}$$

for $\beta = \gamma_3 = 1$.

When (21) is violated, the output levels are bunched within at least one distribution and

sequential screening only concerns the other distributions. Under (22) or (23), the general problem is replaced by Γ' or Γ'', respectively, unless the marginal surplus is very concave. If such relevant condition were violated, then it would be necessary to consider another reduced problem for the situation in which $S'(\cdot)$ is very concave. In the new reduced problem, one would need to impose that $q_{LH} = q_{HL}$ or $r_{LH} = r_{HL}$ so as to satisfy the first or the second condition in (4) violated in Γ' or Γ''. Provided that $\varepsilon_{LH} < \varepsilon_H$ and $\varepsilon'_{LH} < \varepsilon_H$, this case only applies when $S'(\cdot)$ is very concave. Furthermore, it has no implications on the results in Proposition 5.

7 More than four distributions

We now consider a setting in which the principal faces more than four possible distributions of the unit cost and investigate whether there is any change in the agent's incentives, and hence in the contractual choice, in this more general setting. For instance, we would like to figure out whether type LH is more attracted by type LL or, rather, by a type with lower θ and/or lower σ; by type HH or a type with higher θ and/or higher σ; by type HL or a type with higher θ and/or lower σ. We saw that in the four-type setting complications follow from the possibility of each of the two information dimensions being misrepresented in either direction. We will now explore whether complications might also be determined by the *extent* to which each information dimension can be misrepresented. We provide monotonicity conditions under which the additional incentive constraints whereby "more distant" types are unattractive reports are not binding in the generalized setting.

Assume that there are more than four types and that, for any triplet of types indexed by $h \in \{1, 2, 3\}$, it is:

$$\theta_1 \leq \theta_2 \leq \theta_3$$

and either

$$\sigma_1 \geq \sigma_2 \geq \sigma_3$$

or

$$\sigma_1 \leq \sigma_2 \leq \sigma_3,$$

together with either $\theta_h - \theta_{h-1} \geq \sigma_h - \sigma_{h-1}$ or $\theta_h - \theta_{h-1} \leq \sigma_h - \sigma_{h-1}$, $\forall h \in \{2, 3\}$. In either case, it is possible to have type $h - 1$ more efficient than type h. The incentives of type 1 to mimic type 2 are weakly stronger than its incentives to mimic type 3. When $\theta_h - \theta_{h-1} \leq \sigma_h - \sigma_{h-1}$ and type $h + 1$ is potentially more efficient than type h, the incentives of type 3 to mimic type 2 are weakly stronger than its incentives to mimic type 1.

Resting on this result, one can say that, in a richer framework with more than four possible distributions, if the extent of private information about one dimension is monotonic relative to that about the other, then for type LH the report LL is more attractive than any report including higher θ and/or lower σ, the report HH is more attractive than any report including higher θ and/or higher σ, and so on. This reasoning applies to all possible "downward" lies

identified according to type orderings (2) and (3).

As an illustration, assume that there are four possible groups of four distributions ij, where $i, j \in \{1, 2, 3, 4\}$:

$$\{14, 13, \ 24, \ 23\}, \ \{12, \ 11, \ 22, \ 21\}, \ \{34, \ 33, \ 44, \ 43\}, \ \{32, \ 31, \ 42, \ 41\}.$$

Denote θ_i and σ_j the parameters of distribution ij, such that $\theta_i > \theta_{i-1}$ and $\sigma_j > \sigma_{j-1}$ for $i, j \in \{2, 3, 4\}$. The distributions in each group display the characteristics of the four distributions considered in our model. As compared to the distributions in the first group, those in the other three groups are such that either the spread takes a lower value (namely, σ_1 and σ_2 rather than σ_3 and σ_4) or the expected cost takes a higher value (namely, θ_3 and θ_4 rather than θ_1 and θ_2), or both. With these characteristics of the distributions, not only can the agent misrepresent either one or both parameters. He can also misrepresent the two parameters to different extents. However, according to the result previously obtained, if the differences in $\theta's$ are monotonic relative to the differences in $\sigma's$, then only "local" lies about either parameter may be an issue. It follows that for the first three types in each group the incentives to mimic types belonging to other groups are weaker than the incentives to mimic types belonging to their own group. For instance, for type 14 the report 13 is more attractive than 12 or 11, the report 24 is more attractive than 34 or 44, the report 23 is more attractive than any of the reports 22, 21, 33, 43, 32, 31, 42 and 41.

As a conclusion, when more than four distributions are possible and the extent of private information about one parameter is monotonic relative to that about the other, the core difficulty in the identification of the binding incentive constraints is still related to the possibility of the agent camouflaging each of the two parameters in either direction. The exact extent to which camouflage can be done has no bite in that respect.

8   On the relationship between preferences and incentives

A fundamental prediction of our analysis is that in the agency relationships of the kind here considered there is a strong nexus between the characteristics of the optimal contract and the principal's preferences for the good delivered by the agent. To understand the reach of this result, it is useful to consider two alternative interpretations.

8.1   The principal as a consumer taking decisions under uncertainty

The first interpretation, which we previously mentioned, rests on the theoretical studies about consumption decisions under uncertainty (see Leland [27] for a by-now classical contribution; for more recent work, see Menegatti [30], for instance). According to those studies, an individual whose marginal utility is (weakly) convex with respect to consumption is prudent

and engages in precautionary saving in order to avoid facing a low consumption in the future. As an alternative definition, used by Eeckhoudt and Schlesinger [13], a prudent individual prefers a lottery yielding higher possible values of the random outcome, $\widetilde{\varepsilon}$, at a higher certain cost of k over a lottery yielding lower possible values of the random outcome, $\widetilde{\varepsilon} - k$, at no cost (see Definition 1, p.282).[9] This entails that the certain negative event $(-k)$ is separated from the random outcome $(\widetilde{\varepsilon})$ so as to reduce the risk of facing undesirable outcomes.

In light of our results, the preference for the first of the two lotteries aforementioned admits a broader interpretation. Instead of a preference for separating the certain negative event from the random outcome, one can read it as a preference for incurring a higher certain cost against the possibility of facing a better outcome in any possible event. To see this, suppose that in our setting P can choose between two lotteries under incomplete information about the *expected cost*. One such lottery yields the outcome $(\underline{y}_1, \overline{y}_1)$ - an output profile - and entails an expected cost of R_1 - an information rent -, the other lottery yields $(\underline{y}_2, \overline{y}_2)$ with expected cost of R_2, and the two lotteries are such that $\underline{y}_1 > \underline{y}_2$ and $\overline{y}_1 > \overline{y}_2$ together with $R_1 > R_2$. This involves that P must choose between a lottery that grants higher possible outcomes $(\underline{y}_1$ and $\overline{y}_1)$ at a higher certain cost (R_1) and a lottery that grants lower possible outcomes $(\underline{y}_2$ and $\overline{y}_2)$ at a lower certain cost (R_2). When S' is sufficiently convex, P prefers the former lottery to the latter. In other words, she is eager to incur a higher certain cost against the possibility of facing a better outcome in any possible event, consistent with the re-interpretation of the lottery preference in Eeckhoudt and Schlesinger [13] suggested above.

Next suppose that P can choose between two lotteries under incomplete information about the *spread* of the cost. Using the same notation as above, the outcomes of the lotteries are now such that $\underline{y}_1 > \underline{y}_2$ and $\overline{y}_1 < \overline{y}_2$, whereas the costs are such that $R_1 > R_2$. Reading our results as the preference for a certain lottery against some given alternative, we can say that the lottery $\left\{(\underline{y}_1, \overline{y}_1); R_1\right\}$ becomes more attractive to P than the lottery $\left\{(\underline{y}_2, \overline{y}_2); R_2\right\}$ as the marginal surplus becomes more convex. That is, P becomes more eager to incur a higher certain cost against the possibility of facing a better outcome when a good event occurs (*i.e.*, consuming more when the shock is good), although this also involves facing a worse outcome when the bad event occurs (*i.e.*, consuming less when the shock is bad), rather than incurring a lower certain cost associated with more similar outcomes. Therefore, not only does a more convex marginal surplus mirror a stronger preference for a higher expected outcome. It also involves a preference for enjoying high outcomes when favorable events are realized, although this is associated with accepting low outcomes when events are unfavorable.

Admitting now private information about *both* the expected cost and the spread of the cost, as in our screening problem, the lottery preferences of P display both of the patterns just described as S''' increases. Therefore, the contractual solution reflects a double choice, namely a choice between a "less risky" lottery and "riskier" lottery, in the same vein as in Consumption theory, and a choice between two lotteries that are both "less risky" than the "riskier" one.

---

[9]The authors acknowledge that the equivalence between the lottery preference they refer to and the (weak) convexity of the marginal utility is shown by Bigelow and Menezes (1995).

The preference for one of the less risky lotteries over the other is determined by how great the extent of information about the expected cost is relative to extent of information about the spread. This is what Corollary 6 above determines.

Interestingly, the comparison with Consumption theory serves to highlight the benefit of screening the agent in two subsequent stages rather than in the contracting stage only. Suppose for a moment that P does not screen the agent sequentially and sets the same output regardless of the shock for each cost distribution. This can be regarded as a "certain outcome," and the benefit of sequential screening can be derived as the gain that P obtains by moving away from that outcome and conditioning the output level on the shock that will determine the final cost. Let $\widehat{y}_{ij}$ be the inflexible output rule and R' the expected rent under static screening. Also define $y_{ij}(\varepsilon) = \widehat{y}_{ij} + \eta_{ij}(\varepsilon)$ the shock-contingent output under sequential screening. The "flexibility gain" to P is given by:

$$\sum_{ij} \mathbb{E}_{ij} \left[\mathbb{E}_{\varepsilon} \left[S(\widehat{y}_{ij} + \eta_{ij}(\varepsilon)) - (\theta_i + \varepsilon\sigma_j)(\widehat{y}_{ij} + \eta_{ij}(\varepsilon)) \right] - (S(\widehat{y}_{ij}) - \theta_i\widehat{y}_{ij}) \right] - (R - R').$$

The fact that it is optimal for P to screen in two stages involves that this expression is positive. Therefore, sequential screening grants an option value to P, which consists in the possibility of adjusting the future production according to the cost realization. The derivative of the above expression with respect to $\widehat{y}_{ij}$, namely $\mathbb{E}_{ij} \left\{ \mathbb{E}_{\varepsilon} \left[S'(\widehat{y}_{ij} + \eta_{ij}(\varepsilon)) \right] - S'(\widehat{y}_{ij}) \right\} - \frac{d}{d\widehat{y}}(R - R')$, is higher the higher S'''. Under this condition, adding the state-contingent component $\eta_{ij}(\varepsilon)$ to the non-sequential outcome $\widehat{y}_{ij}$ increases the expected marginal surplus of P, and hence the flexibility gain she obtains by screening the agent in two stages. In substance, sequential screening is more beneficial to P the higher S''' is.

8.2 The principal as a regulator taking decisions according to the price-elasticity of demand

The relationship between P and the agent can also be regarded as one between a regulator and a privately informed monopolist supplying a good (or service) to society. The regulator collects revenues in the market and pay transfers to the firm as a compensation for its production. The objective of the regulator is to maximize the expected consumer surplus net of the compensation to the firm. In this context, the marginal surplus function is the inverse demand function, which expresses the consumer willingness to pay for any given consumption level. A shift from concavity to convexity of $S'(\cdot)$ corresponds to a shift from less to more elastic demand. When the demand is little elastic, an increase in the price triggers a limited reduction in the demand for the good. The regulator can afford significant quantity distortions for rent-extraction purposes. When the demand is more elastic, an increase in the price triggers greater demand reductions. The regulator can afford smaller quantity distortions and is amenable to concede higher rents to eliminate the firm's incentives to camouflage information.

As a simple illustration, consider the surplus function $S(y) = ky - y^{e+1}/(e+1)$, where

$k, e > 0$. The marginal surplus $S'(y) = k - y^e$ is strictly concave in y if $e > 1$, linear if $e = 1$, and strictly convex if $e < 1$. $S'(y) \equiv p(y)$ is the market demand and e measures the elasticity of the latter to price. A shift from concavity to convexity of S' represents a shift from less to more price-elastic demand. Using the expression of the first-best quantity of type $i'j'$, namely $y^*_{i'j'}(\varepsilon) = (k - \theta_{i'} + \varepsilon\sigma_{j'})^{\frac{1}{e}}$, $\forall \varepsilon \in \{-1; 1\}$, we see that the expected quantity $q^*_{i'j'} = \frac{1}{2}\left[(k - \theta_{i'} + \sigma_{j'})^{\frac{1}{e}} + (k - \theta_{i'} - \sigma_{j'})^{\frac{1}{e}}\right]$ and the expected quantity difference $r^*_{i'j'} = \frac{1}{2}\left[(k - \theta_{i'} + \sigma_{j'})^{\frac{1}{e}} - (k - \theta_{i'} - \sigma_{j'})^{\frac{1}{e}}\right]$ are both inversely related to e. Indeed:

$$\frac{dq^*_{i'j'}}{de} = -\left(\frac{1}{4e}\right)^2 \left[(k - \theta_{i'} + \sigma_{j'})^{\frac{1}{e}}\ln(k - \theta_{i'} + \sigma_{j'}) + (k - \theta_{i'} - \sigma_{j'})^{\frac{1}{e}}\ln(k - \theta_{i'} - \sigma_{j'})\right] < 0$$

$$\frac{dr^*_{i'j'}}{de} = -\left(\frac{1}{4e}\right)^2 \left[(k - \theta_{i'} + \sigma_{j'})^{\frac{1}{e}}\ln(k - \theta_{i'} + \sigma_{j'}) - (k - \theta_{i'} - \sigma_{j'})^{\frac{1}{e}}\ln(k - \theta_{i'} - \sigma_{j'})\right] < 0.$$

Knowing from the formulation of $IC^{i'j'}_{ij}$ that the incentives of type ij to announce $i'j'$ depend on $\Delta\theta q_{i'j'}$ and $\Delta\sigma r_{i'j'}$, this is suggestive of the decisions that P will make for the optimal incentive provision under incomplete information. When the market demand is more elastic to price, she will want to secure more consumption in each state and, in particular, in the good state relative to the bad state. Hence, she will be more reticent to distort $q^*_{i'j'}$ and/or $r^*_{i'j'}$ to contain the information rents to be given up to prevent a false claim $i'j'$.

The correspondence between the curvature of the marginal surplus function and the price-elasticity of market demand suggests a way of making functional use of the insights of our study along the current regulatory practice. It is well known that, when regulators have only a vague knowledge of the market demand, they rest widely on elasticity estimates, which can be formed with reasonable accuracy in spite of the poor available information on demand conditions. In regulatory contexts in which the information structure is akin to that here considered, the regulator could use elasticity estimates to identify the relevant information rents and set output accordingly.

9   Conclusion

We studied a multi-dimensional and sequential screening problem in which in the contracting stage the principal faces four possible distributions of cost values, two with equal expected costs but different spreads, the other two with different expected costs but equal spreads. We found that the features of the optimal contract are driven essentially by the relationship between the principal's preferences for consumption of the good and the agent's incentive constraints that are binding at optimum. When looking at specific applications, a clear understanding of that relationship is extremely useful in that it enables one to narrow the number of constraints to be accounted for in the identification of the solution. The characterization of the optimal contract we provided for one possible type ordering illustrates how lengthy and intricate the exercise is if that relationship is not considered.

Our analysis highlights that the only situation in which local incentive constraints imply global incentive constraints, as in sequential screening models with one information dimension, is that in which the principal's marginal surplus is sufficiently convex in quantity. That is the situation in which the principal obtains the highest "flexibility gain" from adjusting output levels to future cost realizations rather than adopting an inflexible output rule. Re-interpreting the contractual choice as a preference over lotteries as in Consumption theory, one can think of the principal as preferring a lottery yielding more efficient outcomes at a higher agency cost over a lottery yielding less efficient outcomes at a lower agency cost. Whereas this re-interpretation appears to fit in with any screening problem, the interesting aspect here is that the preference for a particular lottery corresponds to a particular combination of incentive constraints binding at optimum.

Extending the analysis to allow for a greater number of cost distributions, we found that the relevant incentives to misrepresent information are still of the same nature as in the four-distribution scenario. This is because what really matters in the determination of the solution to our two-dimensional screening problem is the possibility of the agent lying on either piece of information and in either direction, rather than the extent to which the reported values can diverge from the real values. Therefore, our results carry over naturally in the more general setting.

Our investigation was inspired by situations in which public activities are delegated to firms that might have incentives to manipulate forecasts of initially unknown variables *vis-à-vis* public authorities. In practice, such activities are now typically awarded to firms by means of tendering procedures. To account for this, we might consider an auction mechanism rather than looking at an incentive contract. However, our choice was not reductive. The insights of our work would not change in that environment because in our model, as in Riordan and Sappington [33], there would be separability between the number of bidders and contractual allocation. Moreover, while we focused on a full-commitment framework, the delegation of public activities sometimes occurs in limited-commitment environments in which firms might camouflage forecasts in the contracting stage in view of a later renegotiation. To eliminate the perspective of contractual renegotiation, hence incentives to strategic misrepresentation related to that, one can think of the principal as being able to credibly engage in the future enforcement in the presence of credible institutions, and of the agent as being motivated to comply with the contract in the presence of cancellation fees or *ex post* participation constraints.[10] Contract design under *ex post* participation constraints is analyzed by Spulber [34] and Chen and Smith [7] in a setting in which the firm has private information on distribution parameters involving the cost in the contracting stage, but the cost realization is publicly observed in a later stage. Because it is natural that the firm observes privately not only the distribution parameters but also the cost realization, our findings suggest that the impact of *ex post* participation constraints on contractual design cannot be well understood without considering the relationship between the principal's preferences and relevant incentives. Exploring this problem is on our research

---

[10]For a discussion of commitment in continuing relationships, see Baron and Besanko [4].

agenda.

References

[1] Armstrong, M. (1996), "Multiproduct nonlinear pricing," *Econometrica*, 64, 51-75

[2] Armstrong, M. (1999), "Optimal Regulation with Unknown Demand and Cost Functions," *Journal of Economic Theory*, 84, 196-215

[3] Armstrong, M., and J.C. Rochet (1999), "Multi-dimensional screening: A user's guide," *European Economic Review*, 43, 959-979

[4] Baron, D.P., and D. Besanko (1984), "Regulation and information in a continuing relationship," *Information Economics and Policy*, 1, 267-302

[5] Baron, D.P., and R.B. Myerson (1982), "Regulating a Monopolist with Unknown Costs," *Econometrica*, 50(4), 911-930

[6] Board, S. (2007), "Selling options," *Journal of Economic Theory*, 136, 324-340

[7] Chen, Y., and R. Smith (2001), "Equilibrium Cost Overruns," *Annals of Economics and Finance*, 2(2), 401-414

[8] Choné, P., and J.C. Rochet (1998), "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66(4), 783-826

[9] Courty, P., and H. Li (2000), "Sequential Screening," *Review of Economic Studies*, 67, 697-717

[10] Dai, C., T.R. Lewis and G. Lopomo (2006), "Delegating Management to Experts," *RAND Journal of Economics*, 37(3), 503-520

[11] Danau, D., and A. Vinella (2015), "Public-private contracting under limited commitment," *Journal of Public Economic Theory*, 17(1), 78-110

[12] Dixit, A. and Pindyck R. S. (1994), *Investment under uncertainty*, Princeton University Press

[13] Eeckhoudt, L., and H. Schlesinger (2006), "Putting Risk in Its Proper Place," *The American Economic Review*, 96(1), 280-289

[14] Eso, P., and B. Szentes (2007), "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, 74, 705–731

[15] Flyvbjerg, B. (2008), "Curbing Optimism Bias and Strategic Misrepresentation in Planning: Reference Class Forecasting in Practice," *European Planning Studies*, 16(1)

[16] Flyvbjerg, B., M.K. Skamris Holm, and S.L. Buhl (2002), "Underestimating costs in public works projects: Error or lie?," *Journal of the American Planning Association*, 68(3), 279-295

[17] Flyvbjerg, B., M.K. Skamris Holm and S.L. Buhl (2005), "How (in)accurate are demand forecasts in public works projects? The case of transportation," *Journal of the American Planning Association*, 71(2), 131-146

[18] Guasch, J. L., J.J. Laffont and S. Straub (2006), "Renegotiation of concession contracts: A theoretical approach," *Review of Industrial Organization*, 29, 55-73

[19] Guasch, J. L., J.J. Laffont and S. Straub (2008), "Renegotiation of Concession Contracts in Latin America, Evidence from the Water and Transport Sectors," *International Journal of Industrial Organization*, 26(2), 421-442

[20] Hoffmann, F., and R. Inderst (2011), "Pre-sale information," *Journal of Economic Theory*, 146, 2333–2355

[21] Iossa, E., and D. Martimort (2012), "Risk allocation and the costs and benefits of public–private partnerships," *RAND Journal of Economics*, 43(3), 442-474

[22] Kjerstad, E., and S. Vagstad (2000), "Procurement auctions with entry of bidders," *International Journal of Industrial Organization*, 18(8), 1243-1257

[23] Krahmer, D., and R. Strausz (2008), "Ex Post Private Information and Monopolistic Screening," *The B.E. Journal of Theoretical Economics - Topics*, 8(1), Article 25

[24] Krahmer, D., and R. Strausz (2011), "Optimal Procurement Contracts with Pre-Project Planning," *Review of Economic Studies*, 78, 1015-1041

[25] Krahmer, D., and R. Strausz (2015), "Sequential versus Static Screening: An equivalence result," *mimeo*

[26] Laffont, J.J. (2003), "Enforcement, Regulation and Development," *Journal of African Economies*, 12, 193-211

[27] Leland, H., (1968) "Saving and uncertainty: The precautionary demand for saving," *Quarterly Journal of Economics*, 82, 465-473

[28] Lewis, T.R., and D.E.M. Sappington (1989), "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294-313

[29] Maggi, G., and A. Rodriguez-Clare (1995), "On Countervailing Incentives," *Journal of Economic Theory*, 66, 238-263

[30] Menegatti, M. (2001), "On the Conditions for Precautionary Saving," *Journal of Economic Theory*, 98, 189-193

[31] Mougeot, M., and F. Naegelen (2011), "Franchise bidding, regulation and investment costs," *Review of Economic Design*, 15(1), 37-58

[32] Myerson, R. (1986), "Multistage games with communication," *Econometrica*, 54(2), 323-358

[33] Riordan, M.H., and D.E.M. Sappington (1987), "Awarding Monopoly Franchises," *American Economic Review*, 77(3), 375-387

[34] Spulber, D.F. (1990), "Auctions and contract enforcement," *Journal of Law, Economics & Organization*, 6(2), 325-344

[35] Wilson, R. (1993), *Nonlinear Pricing*, Oxford University Press, Oxford

A   Benchmarks

A.1   Proof of Proposition 2

This proof is similar to Proof B.1.1 of Proposition 3 here below, and hence it is omitted.

B   Partial sequential screening and cost overruns

B.1   Proof of Proposition 3

We first prove that type $i'j'$ is assigned $y_{i'j'}(-1) = y_{i'j'}(1)$. We then prove that for types HH and HL it is also $y_{HH} = y_{HL}$.

B.1.1 Proof of $y_{Hj'}(\varepsilon) = y_{Hj'}(\widetilde{\varepsilon})$

The incentive constraint whereby an agent of type ij is unwilling to claim $i'j'$, when he anticipates that in the second stage he will report either the true shock ε or the false shock $\widetilde{\varepsilon}$, is written as:

$$\Pi_{ij} \geq \frac{1}{2} \max\left\{\sum_\varepsilon \left[t_{i'j'}(\varepsilon) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\varepsilon)\right]; \sum_\varepsilon \left[t_{i'j'}(\widetilde{\varepsilon}) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\widetilde{\varepsilon})\right]\right\}. \quad (24)$$

We have:

$$\frac{1}{2}\sum_\varepsilon \left[t_{i'j'}(\varepsilon) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\varepsilon)\right] = \frac{1}{2}\sum_\varepsilon \left\{\pi_{i'j'}(\varepsilon) + \left[\theta_{i'} - \theta_i + (\sigma_{j'} - \sigma_j)\varepsilon\right] y_{i'j'}(\varepsilon)\right\}$$

together with:

$$\frac{1}{2}\sum_\varepsilon \left[t_{i'j'}(\widetilde{\varepsilon}) - (\theta_i + \varepsilon\sigma_j) y_{i'j'}(\widetilde{\varepsilon})\right] = \frac{1}{2}\sum_\varepsilon \left\{\pi_{i'j'}(\widetilde{\varepsilon}) + \left[\theta_{i'} + \widetilde{\varepsilon}\sigma_{j'} - (\theta_i + \varepsilon\sigma_j)\right] y_{i'j'}(\widetilde{\varepsilon})\right\}.$$

If the former amount is at least as great as the latter, then $IC_{ij}^{i'j'}$ implies (24). This is the case if and only if:

$$\frac{1}{2}\sum_\varepsilon \left(\pi_{i'j'}(\varepsilon) - \pi_{i'j'}(\widetilde{\varepsilon})\right) \quad (25)$$

$$\geq -\frac{1}{2}\sum_\varepsilon \left[(\theta_{i'} - \theta_i)(y_{i'j'}(\varepsilon) - y_{i'j'}(\widetilde{\varepsilon})) + (\sigma_{j'} - \sigma_j)\varepsilon y_{i'j'}(\varepsilon) + \varepsilon\sigma_j y_{i'j'}(\widetilde{\varepsilon}) - \widetilde{\varepsilon}\sigma_{j'} y_{i'j'}(\widetilde{\varepsilon})\right].$$

The constraints $ic_{ij,\widetilde{\varepsilon}}$ and $ic_{ij,\varepsilon}$ hold jointly only if:

$$(\widetilde{\varepsilon} - \varepsilon)\sigma_{j'} y_{i'j'}(\widetilde{\varepsilon}) \leq \pi_{i'j'}(\varepsilon) - \pi_{i'j'}(\widetilde{\varepsilon}) \leq (\widetilde{\varepsilon} - \varepsilon)\sigma_{j'} y_{i'j'}(\varepsilon),$$

which imply the following conditions:

$$\sum_\varepsilon (\widetilde{\varepsilon} - \varepsilon)\sigma_{j'} y_{i'j'}(\widetilde{\varepsilon}) \leq \sum_\varepsilon \left[\pi_{i'j'}(\varepsilon) - \pi_{i'j'}(\widetilde{\varepsilon})\right] \leq \sum_\varepsilon (\widetilde{\varepsilon} - \varepsilon)\sigma_{j'} y_{i'j'}(\varepsilon).$$

$\exists \pi_{i'j'}(\varepsilon), \pi_{i'j'}(\widetilde{\varepsilon})$ satisfying these conditions together with (25) if and only if:

$$\sum_\varepsilon \left[\widetilde{\varepsilon}\sigma_{j'} - \varepsilon\sigma_j + (\theta_{i'} - \theta_i)\right](y_{i'j'}(\varepsilon) - y_{i'j'}(\widetilde{\varepsilon})) \geq 0. \quad (26)$$

Provided that $\theta_L + \sigma_j < \theta_H - \sigma_{j'}$, if $\varepsilon < \widetilde{\varepsilon}$, then it is necessary that $y_{Hj'}(\varepsilon) \leq y_{Hj'}(\widetilde{\varepsilon})$. However, the second-stage incentive constraints require that $y_{Hj'}(\varepsilon) \geq y_{Hj'}(\widetilde{\varepsilon})$. Hence $y_{Hj'}(\varepsilon) = y_{Hj'}(\widetilde{\varepsilon})$. In the next proof we use the notation: $y_{Hj'}(\varepsilon) = y_{Hj'}, \forall\varepsilon$.

B.1.2 Proof of $y_{HH} = y_{HL}$

Knowing that $\theta_L + \sigma_j < \theta_H - \sigma_{j'}$, hence $\Delta\theta > \Delta\sigma$, the type ordering is (2) and information rents are written as (7a) - (7d). Using $y_{Hj'}(\varepsilon) = y_{Hj'}, \forall\varepsilon$, $r_{HL} = r_{HH} = 0$ and it is easy to deduce that $\gamma_2 = 1$. Replacing this in the expressions of the rents, then replacing the expressions of the rents in the objective function together with $\mathbb{E}_\varepsilon\left[(\theta_i + \varepsilon\sigma_j) y_{ij}\right] = \theta_i y_{ij}$, the

objective function of P is rewritten as:

$$\sum_{ij \in \{LL, LH\}} \mathbb{E}_{ij} \left[\mathbb{E}_{\varepsilon} \left[S\left(y_{ij}\left(\varepsilon\right)\right) - \left(\theta_i + \varepsilon\sigma_j\right) y_{ij}\left(\varepsilon\right) \right] \right] + \sum_{ij \in \{HH, HL\}} \mathbb{E}_{ij} \left[S\left(y_{ij}\right) - \theta_i y_{ij} \right]$$
$$- \nu\mu\Delta\theta \left[\beta y_{HL} + \left(1 - \beta\right) y_{HH} \right] - \nu(1 - \mu) \left[\beta\Delta\theta y_{HL} + \left(1 - \beta\right)\Delta\theta y_{HH} + \Delta\sigma r_{LL} \right].$$

The first-order conditions with respect to $y_{LL}\left(\varepsilon\right)$, y_{HH} and y_{HL} are given by:

$$S'(y_{LL}\left(\varepsilon\right)) = \theta_L + \varepsilon\sigma_L - \frac{1 - \mu}{\mu}\varepsilon\Delta\sigma$$

$$S'(y_{HH}) = \theta_H + \frac{\nu}{1 - \nu}\left(1 - \beta\right)\left(1 + \frac{\mu}{1 - \mu}\right)\Delta\theta$$

$$S'(y_{HL}) = \theta_H + \frac{\nu}{1 - \nu}\beta\left(1 + \frac{1 - \mu}{\mu}\right)\Delta\theta.$$

Hence, we can compute:

$$S'(y_{HH}) - S'(y_{HL}) = \frac{\nu}{1 - \nu}\frac{\mu - \beta}{\mu\left(1 - \mu\right)}\Delta\theta.$$

First suppose that $\beta > \mu$. Then, $S'(y_{HH}) < S'(y_{HL})$ and $y_{HH} > y_{HL}$. However, this implies that $\beta = 0$, which leads to a contradiction. Next suppose that $\beta < \mu$. Then, $S'(y_{HH}) > S'(y_{HL})$ and $y_{HH} < y_{HL}$. However, this implies that $\beta = 1$, which leads to a contradiction. We conclude that $\beta = \mu$ and $y_{HH} = y_{HL}$.

B.2   Proof of Corollary 1

The fact that $\gamma_2 = 1$ in the proof of Proposition 3 signifies that type LH is assigned a rent for not pretending LL. Types HH and HL are bunched, hence no rent is assigned to type HH for not pretending HL.

B.3   Proof of Corollary 2

In Proof B.1.1, for some ij and $i'j'$, it is impossible to have (26) satisfied as $\widetilde{\varepsilon} \to \varepsilon$.

C Full sequential screening and cost overruns

The first-stage incentive constraints are written as follows:

$$\Pi_{LL} \geq \Pi_{HL} + \Delta\theta q_{HL} \tag{IC1}$$
$$\Pi_{LL} \geq \Pi_{LH} - \Delta\sigma r_{LH} \tag{IC2}$$
$$\Pi_{LL} \geq \Pi_{HH} + \Delta\theta q_{HH} - \Delta\sigma r_{HH} \tag{IC3}$$
$$\Pi_{HL} \geq \Pi_{LL} - \Delta\theta q_{LL} \tag{IC4}$$
$$\Pi_{HL} \geq \Pi_{HH} - \Delta\sigma r_{HH} \tag{IC5}$$
$$\Pi_{HL} \geq \Pi_{LH} - \Delta\theta q_{LH} - \Delta\sigma r_{LH} \tag{IC6}$$
$$\Pi_{LH} \geq \Pi_{HH} + \Delta\theta q_{HH} \tag{IC7}$$
$$\Pi_{LH} \geq \Pi_{LL} + \Delta\sigma r_{LL} \tag{IC8}$$
$$\Pi_{LH} \geq \Pi_{HL} + \Delta\theta q_{HL} + \Delta\sigma r_{HL} \tag{IC9}$$
$$\Pi_{HH} \geq \Pi_{LH} - \Delta\theta q_{LH} \tag{IC10}$$
$$\Pi_{HH} \geq \Pi_{HL} + \Delta\sigma r_{HL} \tag{IC11}$$
$$\Pi_{HH} \geq \Pi_{LL} - \Delta\theta q_{LL} + \Delta\sigma r_{LL}. \tag{IC12}$$

The downward constraints are (IC1), (IC7), (IC8), (IC9) and (IC11), together with (IC3) when $\Delta\theta > \Delta\sigma$, and with (IC12) otherwise.

C.1 Proof of Lemma 1

C.1.1 The case of $\Delta\theta > \Delta\sigma$

We write all the information rents resulting if one downward incentive constraint of each type except HL is binding, as a function of Π_{HL}. For type HH:

$$\Pi_{HH} = \Pi_{HL} + \Delta\sigma r_{HL}. \tag{27}$$

For type LL either:

$$\Pi_{LL,1} = \Pi_{HL} + \Delta\theta q_{HL} \tag{28}$$

or

$$\begin{aligned}\Pi_{LL,2} &= \Pi_{HH} + \Delta\theta q_{HH} - \Delta\sigma r_{HH} \\ &= \Pi_{HL} + \Delta\theta q_{HL} + \Delta\theta \left(q_{HH} - q_{HL}\right) - \Delta\sigma \left(r_{HH} - r_{HL}\right).\end{aligned} \tag{29}$$

For type LH either:

$$\begin{aligned}\Pi_{LH,1} &= \Pi_{HH} + \Delta\theta q_{HH} \\ &= \Pi_{HL} + \Delta\sigma r_{HL} + \Delta\theta q_{HL} + \Delta\theta \left(q_{HH} - q_{HL}\right)\end{aligned} \tag{30}$$

or:

$$\begin{aligned}\Pi_{LH,2} &= \Pi_{LL} + \Delta\sigma r_{LL} \\ &= \Pi_{HL} + \Delta\theta q_{HL} + \Delta\sigma r_{HL} + \Delta\sigma \left(r_{LL} - r_{HL}\right)\end{aligned} \tag{31}$$

or:

$$\begin{aligned}\Pi_{LH,3} &= \Pi_{LL} + \Delta\sigma r_{LL} \\ &= \Pi_{HL} + \Delta\theta q_{HL} + \Delta\sigma r_{HL} + \Delta\sigma \left(r_{LL} - r_{HL}\right) + \Delta\theta \left(q_{HH} - q_{HL}\right) - \Delta\sigma \left(r_{HH} - r_{HL}\right).\end{aligned} \tag{32}$$

or:
$$\Pi_{LH,4} = \Pi_{HL} + \Delta\sigma r_{HL} + \Delta\theta q_{HL} \tag{33}$$

At various stages, we will use the implication:

$$y_{ij}\left(\varepsilon\right) \geq y_{i'j'}\left(\varepsilon\right), \ \forall\varepsilon \Rightarrow \Delta\theta(q_{ij} - q_{i'j'}) \geq \Delta\sigma(r_{ij} - r_{i'j'}), \tag{34}$$

which is due to the fact that the latter inequality in (34) is equivalent to:

$$\left(\Delta\theta - \Delta\sigma\right)\left(y_{ij}\left(-1\right) - y_{i'j'}\left(-1\right)\right) \geq \left(\Delta\theta + \Delta\sigma\right)\left(y_{i'j'}\left(1\right) - y_{ij}\left(1\right)\right).$$

Take HH as possible report.

Given (27), (IC5) holds if $r_{HH} \geq r_{HL}$, which is satisfied because $y_{HH}\left(-1\right) \geq y_{HL}\left(-1\right)$ and $y_{HH}\left(1\right) \leq y_{HL}\left(1\right)$.

Take LL as possible report.

Given (28), (IC4) holds if $q_{LL} \geq q_{HL}$, which is satisfied because $y_{LL}\left(\varepsilon\right) \geq y_{HL}\left(\varepsilon\right) \ \forall\varepsilon$; given (29), it holds if:

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

This condition holds if $r_{HH} \geq r_{HL}$, which is satisfied, and $q_{LL} \geq q_{HH}$, which holds, in turn, because $y_{LL}\left(\varepsilon\right) \geq y_{HH}\left(\varepsilon\right) \ \forall\varepsilon$. Given (28), (IC12) holds if:

$$\Delta\theta\left(q_{LL} - q_{HL}\right) \geq \Delta\sigma\left(r_{LL} - r_{HL}\right),$$

which is satisfied because $y_{LL}\left(\varepsilon\right) \geq y_{HL}\left(\varepsilon\right) \ \forall\varepsilon$; given (29), it holds if:

$$\Delta\theta\left(q_{LL} - q_{HH}\right) \geq \Delta\sigma\left(r_{LL} - r_{HH}\right),$$

which is satisfied because $y_{LL}\left(\varepsilon\right) \geq y_{HH}\left(\varepsilon\right) \ \forall\varepsilon$.

Take LH as possible report.

Given (33), (IC6) holds if:

$$\Delta\theta\left(q_{LH} - q_{HL}\right) \geq \Delta\sigma\left(r_{HL} - r_{LH}\right),$$

which is satisfied when $q_{LH} \geq q_{HL}$ together with $r_{LH} \geq r_{HL}$; given (30), it holds if:

$$\Delta\theta\left(q_{LH} - q_{HH}\right) \geq \Delta\sigma\left(r_{HL} - r_{LH}\right),$$

which is satisfied when $q_{LH} \geq q_{HH}$ (which holds because $y_{LH}\left(\varepsilon\right) \geq y_{HH}\left(\varepsilon\right) \ \forall\varepsilon$) together with $r_{LH} \geq r_{HL}$; given (31), it holds if:

$$\Delta\theta\left(q_{LH} - q_{HL}\right) \geq \Delta\sigma\left(r_{LL} - r_{LH}\right),$$

which is satisfied when $q_{LH} \geq q_{HL}$ together with $r_{LH} \geq r_{LL}$, which holds because $y_{LH}\left(-1\right) \geq y_{LL}\left(-1\right)$ and $y_{LH}\left(1\right) \leq y_{LL}\left(1\right)$; given (32), it holds if:

$$\Delta\theta\left(q_{LH} - q_{HH}\right) \geq -\Delta\sigma\left(r_{LH} - r_{LL}\right) - \Delta\sigma\left(r_{HH} - r_{HL}\right),$$

which is satisfied if $q_{LH} \geq q_{HH}$ together with $r_{LH} \geq r_{LL}$ and with $r_{HH} \geq r_{HL}$, which are all satisfied.

Given (33), (IC10) holds if $q_{LH} \geq q_{HL}$; given (30), it holds if $q_{LH} \geq q_{HH}$, which is satisfied; given (31), it holds if:

$$\Delta\theta\left(q_{LH} - q_{HL}\right) \geq \Delta\sigma\left(r_{LL} - r_{HL}\right), \tag{35}$$

which is ambiguous; and if

$$\Delta\theta\left(q_{LH}-q_{HH}\right)\geq\Delta\sigma\left(r_{LL}-r_{HH}\right),$$

which is satisfied because $r_{LH}\geq r_{LL}$ holds and $y_{LH}\left(\varepsilon\right)\geq y_{HH}\left(\varepsilon\right)\,\forall\varepsilon$. (35) was obtained from (31) and (IC10). Recall that (31) corresponds to the situation where (IC3) is binding. Using the binding (IC3) in (IC2), the latter implies (IC10) if:

$$\Delta\theta\left(q_{LH}-q_{HH}\right)\geq\Delta\sigma\left(r_{LH}-r_{HH}\right),$$

which is satisfied because $y_{LH}\left(\varepsilon\right)\geq y_{HH}\left(\varepsilon\right)\,\forall\varepsilon$.

Given (33) and using (IC1), (IC2) holds if $r_{LH}\geq r_{HL}$; given (30) and using (IC1), it holds if $r_{LH}\geq r_{LL}$, which is satisfied; given either (31) or (32) formulated as $\Pi_{LH,3}=\Pi_{LL}+\Delta\sigma r_{LL}$ and using (IC1), it holds if $r_{LH}\geq r_{LL}$.

The conditions that are not implied by the quantity ranking in the lemma are $r_{LH}\geq r_{HL}$ and $q_{LH}\geq q_{HL}$, as reported in (4).

C.1.2 The case of $\Delta\theta\leq\Delta\sigma$

Because $y_{LL}\left(\varepsilon\right)=y_{HL}\left(\varepsilon\right),\,\forall\varepsilon$ (Lemma 2), the information rents change as follows: $\Pi_{LL}=\Pi_{HL}+\Delta\theta q_{HL}$ for type LL; $\Pi_{HH}=\Pi_{HL}+\Delta\sigma r_{HL}$ for type HH; either $\Pi_{LH,1}$ or $\Pi_{LH,2}=\Pi_{LH,3}=\Pi_{LH,4}$ for type LH.

Take HH as possible report.

Given (27), (IC5) is satisfied if $r_{HH}\geq r_{HL}$, which holds because $y_{HH}\left(-1\right)\geq y_{HL}\left(-1\right)$ and $y_{HH}\left(1\right)\leq y_{HL}\left(1\right)$. (IC3) is satisfied if:

$$\Delta\sigma r_{HH}\geq\Delta\theta q_{HH}+\Delta\sigma r_{HL}$$

$$\Leftrightarrow$$

$$\left(\Delta\sigma-\Delta\theta\right)\left(y_{HH}\left(-1\right)-y_{HL}\left(-1\right)\right)\geq\left(\Delta\sigma+\Delta\theta\right)\left(y_{HH}\left(1\right)-y_{HL}\left(1\right)\right)-\Delta\theta\left(y_{HL}\left(-1\right)-y_{HL}\left(1\right)\right),$$

which is satisfied as well because $y_{HH}\left(-1\right)\geq y_{HL}\left(-1\right)$ and $y_{HH}\left(1\right)\leq y_{HL}\left(1\right)$.

Take LH as possible report.

Given (30), (IC2) is satisfied if:

$$\Delta\sigma\left(r_{LH}-r_{HL}\right)\geq\Delta\theta\left(q_{HH}-q_{HL}\right).$$

Using $r_{HL}=r_{LL}$ and $q_{HL}=q_{LL}$, this is rewritten as:

$$\Delta\sigma\left(r_{LH}-r_{LL}\right)\geq\Delta\theta\left(q_{HH}-q_{LL}\right).$$

Provided $q_{LH}\geq q_{HH}$, because $y_{LH}\left(\varepsilon\right)\geq y_{HH}\left(\varepsilon\right)\,\forall\varepsilon$, the condition above is implied by:

$$\Delta\sigma\left(r_{LH}-r_{LL}\right)\quad\geq\quad\Delta\theta\left(q_{LH}-q_{LL}\right)$$
$$\Leftrightarrow$$
$$\left(\Delta\sigma-\Delta\theta\right)\left(y_{LH}\left(-1\right)-y_{LL}\left(-1\right)\right)\quad\geq\quad\left(\Delta\theta+\Delta\sigma\right)\left(y_{LH}\left(1\right)-y_{LL}\left(1\right)\right),$$

which is satisfied because $y_{LH}\left(-1\right)\geq y_{LL}\left(-1\right)$, $y_{LH}\left(1\right)\leq y_{LL}\left(1\right)$, and $\Delta\theta\leq\Delta\sigma$. Given (31), (IC2) is satisfied if $r_{LH}\geq r_{HL}$. Given (30), (IC6) is satisfied if:

$$\Delta\theta\left(q_{LH}-q_{HH}\right)\geq\Delta\sigma\left(r_{HL}-r_{LH}\right),$$

which is satisfied if $r_{LH}\geq r_{HL}$, provided $q_{LH}\geq q_{HH}$ because $y_{LH}\left(\varepsilon\right)\geq y_{HH}\left(\varepsilon\right)\,\forall\varepsilon$; given (31),

(IC6) is satisfied if:
$$\Delta\theta\left(q_{LH}-q_{HL}\right)\geq\Delta\sigma\left(r_{HL}-r_{LH}\right),$$
which is satisfied if $q_{LH}\geq q_{HL}$ and $r_{LH}\geq r_{HL}$.

Given (30), (IC10) is satisfied if $q_{LH}\geq q_{HH}$, which holds; given (31), (IC10) holds if $q_{LH}\geq q_{HL}$.

The remaining conditions are $r_{LH}\geq r_{HL}$ and $q_{LH}\geq q_{HL}$, as reported in (4).

C.2 Proof of Lemma 2

C.2.1 Case $\Delta\theta>\Delta\sigma$ and reduced problem Γ'

Let Γ' be the reduced problem when $\Delta\theta>\Delta\sigma$ and the quantity solution is fully separating. Replacing $\Pi_{HL}=0$, the rents in (7a) - (7d) are obtained. Using (7a) - (7d), the objective function in Γ' is written as:

$$\sum_{ij\in\Upsilon}\mathbb{E}_{ij}\left[\mathbb{E}_{\varepsilon}\left[S\left(y_{ij}\left(\varepsilon\right)\right)-\left(\theta_{i}+\varepsilon\sigma_{j}\right)y_{ij}\left(\varepsilon\right)\right]\right]\tag{36}$$
$$-\nu\mu\left[\beta\Delta\theta q_{HL}+\left(1-\beta\right)\left(\Delta\sigma r_{HL}+\Delta\theta q_{HH}-\Delta\sigma r_{HH}\right)\right]$$
$$-\left(1-\nu\right)\left(1-\mu\right)\Delta\sigma r_{HL}$$
$$-\nu(1-\mu)\left\{\gamma_{1}\left(\Delta\sigma r_{HL}+\Delta\theta q_{HH}\right)\right.$$
$$+\gamma_{2}\left[\beta\Delta\theta q_{HL}+\left(1-\beta\right)\left(\Delta\sigma r_{HL}+\Delta\theta q_{HH}-\Delta\sigma r_{HH}\right)+\Delta\sigma r_{LL}\right]$$
$$\left.+\gamma_{3}\left(\Delta\theta q_{HL}+\Delta\sigma r_{HL}\right)\right\}.$$

From (36) the first-order conditions (10) - (12) are derived. From (11) and (12), we compute:

$$S'(y_{HH}\left(\varepsilon\right))-S'(y_{HL}\left(\varepsilon\right))$$
$$=\varepsilon\Delta\sigma\left[\frac{1}{\nu}-\left(\beta-\mu\right)\left(\gamma_{2}+\frac{\mu}{1-\mu}\right)\right]-\Delta\theta\left[1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{\left(1-\beta\right)\mu}{1-\mu}\right].$$

First suppose that $1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{(1-\beta)\mu}{1-\mu}>0$. Define:

$$\phi\equiv\frac{\frac{1}{\nu}-\left(\beta-\mu\right)\left(\gamma_{2}+\frac{\mu}{1-\mu}\right)}{1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{(1-\beta)\mu}{1-\mu}}.\tag{37}$$

We see that $\phi>0$ and that, if $\Delta\theta>\phi\Delta\sigma$, then $S'(y_{HH}\left(\varepsilon\right))<S'(y_{HL}\left(\varepsilon\right))$, $\forall\varepsilon\Rightarrow y_{HH}\left(\varepsilon\right)>y_{HL}\left(\varepsilon\right)$, $\forall\varepsilon$. Recall that $\beta=1$ if and only if:

$$\Delta\theta(q_{HH}-q_{HL})<\Delta\sigma(r_{HH}-r_{HL}),$$

which is equivalent to

$$\left(\Delta\theta+\Delta\sigma\right)\left(y_{HL}\left(1\right)-y_{HH}\left(1\right)\right)\geq\left(\Delta\theta-\Delta\sigma\right)\left(y_{HH}\left(-1\right)-y_{HL}\left(-1\right)\right)\tag{38}$$

As $y_{HH}\left(\varepsilon\right)>y_{HL}\left(\varepsilon\right)$, $\forall\varepsilon$, this condition is violated so that $\beta=0$. However, provided it is also $\gamma_{3}=0$, this contradicts the hypothesis that $1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{(1-\beta)\mu}{1-\mu}>0$ because we then have $\gamma_{3}-\frac{\mu}{1-\mu}=-\frac{\mu}{1-\mu}<0$.

Next suppose that $1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{(1-\beta)\mu}{1-\mu}<0$ so that $\phi<0$. If $\Delta\theta>-\Delta\sigma\phi$, then $S'(y_{HH}\left(\varepsilon\right))>S'(y_{HL}\left(\varepsilon\right))$ and so $y_{HH}\left(\varepsilon\right)<y_{HL}\left(\varepsilon\right)$, $\forall\varepsilon$. It follows that $\beta=1$. However, this contradicts the hypothesis that $1-\left(\gamma_{1}+\gamma_{2}\right)+\gamma_{2}\beta-\frac{(1-\beta)\mu}{1-\mu}<0$ because then we have

$1 - (\gamma_1 + \gamma_2) + \gamma_2\beta - \frac{(1-\beta)\mu}{1-\mu} < 0 \Leftrightarrow \gamma_1 > 1$.

Overall, if $\Delta\theta > |\phi|\,\Delta\sigma$, then $\forall\varepsilon$ the hypothesis that $y_{HH}(\varepsilon) \lessgtr y_{HL}(\varepsilon)$ at the solution to Γ leads to a contradiction.

C.2.2   Case $\Delta\theta \le \Delta\sigma$ and reduced problem Γ'''

Let Γ''' be the reduced problem when $\Delta\theta \le \Delta\sigma$. Using the downward constraints, the information rents are written as:

$$
\begin{aligned}
\Pi_{HL} &= 0 \\
\Pi_{LL} &= \Delta\theta q_{HL} \\
\Pi_{HH} &= \beta\Delta\sigma r_{HL} + (1-\beta)(\Delta\theta q_{HL} - \Delta\theta q_{LL} + \Delta\sigma r_{LL}) \\
\Pi_{LH} &= \gamma_1 [\beta\Delta\sigma r_{HL} + (1-\beta)(\Delta\theta q_{HL} - \Delta\theta q_{LL} + \Delta\sigma r_{LL}) + \Delta\theta q_{HH}] \\
&\quad + \gamma_2 (\Delta\theta q_{HL} + \Delta\sigma r_{LL}) + \gamma_3 (\Delta\theta q_{HL} + \Delta\sigma r_{HL}).
\end{aligned}
$$

Replacing these expressions into the objective function, quantities are such that:

$$
\begin{aligned}
S'(y_{LL}(\varepsilon)) &= \theta_L + \varepsilon\sigma_L + \frac{1-\mu}{\mu}\frac{(1-\nu)(1-\beta)\Delta\theta + \nu[\gamma_1\Delta\theta - (\gamma_1 + \gamma_2)\varepsilon\Delta\sigma]}{\nu} \\
S'(y_{HL}(\varepsilon)) &= \theta_H + \varepsilon\sigma_L - \Delta\theta \\
&\quad - \frac{1-\mu}{\mu}\frac{(1-\nu)[\beta\varepsilon\Delta\sigma + (1-\beta)\Delta\theta] + \nu[(1-\beta\gamma_1)\Delta\theta + (\beta\gamma_1 + \gamma_3)\varepsilon\Delta\sigma]}{1-\nu}.
\end{aligned}
$$

We can compute:

$$
\begin{aligned}
S'(y_{LL}(\varepsilon)) - S'(y_{HL}(\varepsilon)) &= \frac{1-\mu}{\mu}\left[1 + \gamma_1 - \beta + \frac{1-\nu}{\nu}(1-\beta) + \frac{\nu(1-\beta\gamma_1)}{1-\nu}\right]\Delta\theta \\
&\quad - \frac{1-\mu}{\mu}\left[\gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1 + \gamma_3)}{1-\nu}\right]\varepsilon\Delta\sigma.
\end{aligned}
$$

Clearly $1 + \gamma_1 - \beta + \frac{1-\nu}{\nu}(1-\beta) + \frac{\nu(1-\beta\gamma_1)}{1-\nu} = \gamma_1 + \frac{1-\beta}{\nu} + \frac{\nu(1-\beta\gamma_1)}{1-\nu} > 0$. Define:

$$
\phi' \equiv \gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1 + \gamma_3)}{1-\nu}.
$$

First suppose that $\gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1 + \gamma_3)}{1-\nu} < 0$ so that $\phi' < 0$. If $\Delta\theta < -\Delta\sigma\phi'$, then $S'(y_{LL}(1)) > S'(y_{HL}(1))$ so that $y_{HL}(1) > y_{LL}(1)$, and $S'(y_{LL}(-1)) < S'(y_{HL}(-1))$ so that $y_{LL}(-1) > y_{HL}(-1)$. It is $\beta = 0$ if and only if:

$$
\Delta\sigma(r_{HL} - r_{LL}) < \Delta\theta(q_{HL} - q_{LL}),
$$

which is equivalent to:

$$
(\Delta\sigma + \Delta\theta)(y_{HL}(1) - y_{LL}(1)) > (\Delta\sigma - \Delta\theta)(y_{HL}(-1) - y_{LL}(-1)).
$$

Moreover, it is $\gamma_3 > 0$ only if:

$$
r_{HL} \ge r_{LL} \Leftrightarrow y_{HL}(-1) - y_{LL}(-1) \ge y_{HL}(1) - y_{LL}(1).
$$

Hence, we deduce that $\beta = \gamma_3 = 0$. Then, $\gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1 + \gamma_3)}{1-\nu} = \gamma_1 + \gamma_2 = 1 > 0$, which

contradicts the hypothesis that this expression is negative.

Next suppose that $\gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1+\gamma_3)}{1-\nu} > 0$ so that $\phi' > 0$. If $\Delta\theta < \Delta\sigma\phi'$, then $S'(y_{LL}(1)) < S'(y_{HL}(1))$ so that $y_{HL}(1) < y_{LL}(1)$, and $S'(y_{LL}(-1)) > S'(y_{HL}(-1))$ so that $y_{HL}(-1) > y_{LL}(-1)$. Since $\beta = 1$ if and only if:

$$(\Delta\sigma - \Delta\theta)(y_{HL}(-1) - y_{LL}(-1)) > (\Delta\sigma + \Delta\theta)(y_{HL}(1) - y_{LL}(1)),$$

it follows that $\beta = 1$. Then, $\gamma_1 + \gamma_2 - \beta - \frac{\nu(\beta\gamma_1+\gamma_3)}{1-\nu} = -[1 - (\gamma_1 + \gamma_2)] - \frac{\nu(\gamma_1+\gamma_3)}{1-\nu} < 0$, which contradicts the hypothesis that this expression is positive.

Overall, if $\Delta\theta < \Delta\sigma|\phi'|$, then the hypothesis that $y_{HH}(\varepsilon) \lessgtr y_{HL}(\varepsilon)$ at the solution to Γ leads to a contradiction $\forall\varepsilon$. Moreover, if $\Delta\theta > \Delta\sigma|\phi'|$, then $S'(y_{LL}(\varepsilon)) > S'(y_{HL}(\varepsilon))$ so that $y_{HL}(\varepsilon) > y_{LL}(\varepsilon)$, $\forall\varepsilon$. It involves that $q_{HL} > q_{LL}$, which contradicts the monotonicity condition implied by (IC1) and (IC4). Hence, when $\Delta\theta \leq \Delta\sigma$, $y_{HL}(\varepsilon) = y_{LL}(\varepsilon)$, $\forall\varepsilon$.

C.3   Proof of Lemma 3

Rewrite:

$$
\begin{aligned}
\zeta(z_1, z_2, x) &= [f(z_1) - f(z_1 + x)] - [f(z_2) - f(z_2 + x)] \\
&= -[f(z_1 + x) - f(z_1)] - \{-[f(z_2 + x) - f(z_2)]\}.
\end{aligned}
$$

We know that $f'(\cdot) < 0$ because $S''(\cdot) < 0$. Also, $S''(y_1) \gtrless S''(y_2)$, $\forall y_1 > y_2$ if $S''' \gtrless 0$. It implies that $f'(\cdot)$ is higher (lower) for the inverse of y_2 than y_1, i.e., for a_2 than a_1, where $S'(y_1) = a_1$, $S'(y_2) = a_2$. Then, $y_1 > y_2$ implies $a_2 > a_1$ so that $f''(\cdot) \gtrless 0$ and $-f''(\cdot) \lessgtr 0$, hence $-f'(a_i) \gtrless -f'(a_j)$ if $a_i < a_j$. For z_1, z_2, x such that $z_1 < z_2$ and $x > 0$:

$$
\begin{aligned}
-\int_0^x f'(v + z_1)\,dv &\gtrless -\int_0^x f'(v + z_2)\,dv \\
&\Leftrightarrow -\int_{z_1}^{z_1+x} f'(v)\,dv \gtrless -\int_{z_2}^{z_2+x} f'(v)\,dv \\
&\Leftrightarrow -[f(z_1 + x) - f(z_1)] \gtrless -[f(z_2 + x) - f(z_2)] \Leftrightarrow \zeta(z_1, z_2, x) \gtrless 0.
\end{aligned}
$$

C.4   Proof of Proposition 4 and Corollaries 3 - 5

C.4.1   Case $\Delta\theta > \Delta\sigma$

Consider Γ' and recall the rents (28) - (32). We see that $\Pi_{LL,2} \geq \Pi_{LL,1}$ if and only if:

$$\Delta\theta(q_{HH} - q_{HL}) \geq \Delta\sigma(r_{HH} - r_{HL}). \tag{40}$$

$\Pi_{LH,1} \geq \Pi_{LH,4}$ if and only if $q_{HH} \geq q_{HL}$; using (6), this is rewritten:

$$q_{HH} \geq q_{HL} \Leftrightarrow \zeta(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) \geq f(\overline{a}_{HL} + \underline{a}_{HL} - \underline{a}_{HH}) - f(\overline{a}_{HH}). \tag{41}$$

$\Pi_{LH,2} \geq \Pi_{LH,4}$ if and only if $r_{LL} \geq r_{HL}$, rewritten as

$$r_{LL} \geq r_{HL} \Leftrightarrow \zeta(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}) \geq f(\overline{a}_{LL} + \underline{a}_{LL} - \underline{a}_{HL}) - f(\overline{a}_{HL}). \tag{42}$$

$\Pi_{LH,3} \geq \Pi_{LH,1}$ if and only if $r_{LL} \geq r_{HL}$; $\Pi_{LH,3} \geq \Pi_{LH,2}$ if and only if (40) holds. Given the quantity ranking in Lemma 1, we have: $\underline{a}_{LL} \leq \underline{a}_{HH} \leq \underline{a}_{HL}$ and $\overline{a}_{LL} \leq \overline{a}_{HL} \leq \overline{a}_{HH}$. Then (40)

is equivalent to:

$$\left(\Delta\theta - \Delta\sigma\right)\left(y_{HH}\left(-1\right) - y_{HL}\left(-1\right)\right) \geq \left(\Delta\theta + \Delta\sigma\right)\left(y_{HL}\left(1\right) - y_{HH}\left(1\right)\right).$$

Computing:

$$
\begin{aligned}
&\left(\Delta\theta - \Delta\sigma\right)\left(y_{HH}\left(-1\right) - y_{HL}\left(-1\right)\right) - \left(\Delta\theta + \Delta\sigma\right)\left(y_{HL}\left(1\right) - y_{HH}\left(1\right)\right) \\
=\ &\left(\Delta\theta - \Delta\sigma\right)\left[\left(y_{HH}\left(-1\right) - y_{HL}\left(-1\right)\right) - \left(y_{HL}\left(1\right) - y_{HH}\left(1\right)\right)\right] - 2\Delta\sigma\left(y_{HL}\left(1\right) - y_{HH}\left(1\right)\right) \\
=\ &\left(\Delta\theta - \Delta\sigma\right)\left\{\zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) + \left[f\left(\overline{a}_{HH}\right) - f\left(\overline{a}_{HL} + \underline{a}_{HL} - \underline{a}_{HH}\right)\right]\right\} \\
&- 2\Delta\sigma\left(f\left(\overline{a}_{HL}\right) - f\left(\overline{a}_{HH}\right)\right),
\end{aligned}
$$

we see that (40) is further rewritten as:

$$\zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) \geq \zeta_1, \tag{43}$$

where ζ_1 is defined in (15).

Then, as $\zeta\left(\cdot,\cdot,\cdot\right)$ increases, $\Pi_{LL,2}$ becomes bigger relative to $\Pi_{LL,1}$; $\Pi_{LH,1}$ and $\Pi_{LH,2}$ become bigger relative to $\Pi_{LH,4}$; $\Pi_{LH,3}$ becomes bigger relative to both $\Pi_{LH,1}$ and $\Pi_{LH,2}$. This evolution highlights that IC_{LL}^{HH} becomes tighter relative to IC_{LL}^{HL}, and IC_{LH}^{LL} becomes tighter relative to both IC_{LH}^{HL} and IC_{LH}^{HH}. Hence, relevant incentive constraints are increasingly more related to the possibility of adjacent types being misrepresented rather than nonadjacent types (Proposition 4), and, equivalently, increasingly more related to the agent misrepresenting σ_j rather than θ_i (Corollary 3). When (41), (42) and (43) hold local incentive constraints imply global incentive constraints (Corollary 4). According to type ordering (2), there is no downward incentive constraint involving understatement of θ_H (Corollary 5).

C.4.2 Case $\Delta\theta \leq \Delta\sigma$

Given the quantity ranking in Lemma 1, we have $\underline{a}_{HH} \leq \underline{a}_{HL}$ and $\overline{a}_{HL} \leq \overline{a}_{HH}$. Recall that the rents are as follows: $\Pi_{LL} = \Pi_{HL}$ for types LL and HL; $\Pi_{HH} = \Pi_{HL} + \Delta\sigma r_{HL}$ for type HH; either $\Pi_{LH,1}$ or $\Pi_{LH,4}$ for type LH. As above, $\Pi_{LH,1} \geq \Pi_{LH,4}$ if and only if (41) is satisfied, which proves again Proposition 4 and Corollaries 3-5 for this case.

D A characterization of the optimal contract under fully sequential screening

D.1 Proof of Proposition 5

D.1.1 Check the solution to Γ'

Lemma 4 $\beta > 0$ if and only if $\zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) \leq \zeta_1$, where $\zeta_1 > 0$.

Proof. The equivalence $\beta > 0 \Leftrightarrow \zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) \leq \zeta_1$ is demonstrated in the proof of Proposition 1. Knowing that $f'\left(\cdot\right) < 0$ and $\overline{a}_{HH} > \overline{a}_{HL}$, we have $f\left(\overline{a}_{HL}\right) - f\left(\overline{a}_{HH}\right) > 0$. Further using $\underline{a}_{HL} > \underline{a}_{HH}$ and $\underline{a}_{HL} - \underline{a}_{HH} > \overline{a}_{HH} - \overline{a}_{HL}$, we find:

$$f\left(\overline{a}_{HL}\right) - f\left(\overline{a}_{HH}\right) = f\left(\overline{a}_{HL}\right) - f\left(\overline{a}_{HL} + \overline{a}_{HH} - \overline{a}_{HL}\right) > f\left(\overline{a}_{HL}\right) - f\left(\overline{a}_{HL} + \underline{a}_{HL} - \underline{a}_{HH}\right),$$

so that $\zeta_1 > 0$. ∎

Lemma 5 At the solution to Γ' :
  (i) If $\zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) \leq \varepsilon_H$, then $\gamma_3 = 1$, $\gamma_1 = 0$ and $\gamma_2 = 0$.
  (ii) If $\varepsilon_H < \zeta\left(\underline{a}_{HH},\overline{a}_{HL},\underline{a}_{HL} - \underline{a}_{HH}\right) \leq \zeta_1$, then $\beta > 0$, $\gamma_3 = 0$, $\gamma_2 > 0$ and $\gamma_1 \geq 0$.

(iii) *If* $\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) > \zeta_1$, *then* $\beta > 0$, $\gamma_3 = 0$, $\gamma_2 > 0$ *and* $\gamma_1 \geq 0$; *if also* $\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{HH} - \underline{a}_{LL}\right) \geq \varepsilon_L$, *where* $\varepsilon_L > 0$, *then.*$\gamma_2 = 1$.

Proof. *Proof of* (i).

Throughout the proof we let: $(\underline{a}_{HL} - \underline{a}_{HH}) - (\overline{a}_{HH} - \overline{a}_{HL}) = \beta\gamma_2 + \gamma_3 - \frac{(1-\beta)\mu}{1-\mu} \equiv \chi$.

Use the definition of ε_H and recall that, applying (6), it is $q_{HH} \geq q_{HL}$ if and only if:

$$\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) \geq \varepsilon_H.$$

Since $\varepsilon_H < \zeta_1$, $\beta = 1$ for $\zeta < \varepsilon_H$.

Furthermore, $\gamma_3 > 0$ if and only if $q_{HL} \geq q_{HH}$ and $r_{LL} \geq r_{HL}$, which is equivalent to:

$$\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{HL} - \underline{a}_{LL}\right) < \varepsilon_{LL}$$

for some $\varepsilon_{LL} = f\left(\overline{a}_{LL} + \underline{a}_{HL} - \underline{a}_{LL}\right) - f\left(\overline{a}_{HL}\right)$. Resting on:

$$(\underline{a}_{HL} - \underline{a}_{LL}) - (\overline{a}_{HL} - \overline{a}_{LL}) = 2\frac{(1-\mu)(1-\gamma_2)}{(1-\nu)\mu}\Delta\sigma, \tag{44}$$

we see that $\varepsilon_{LL} \leq 0$, where $\varepsilon_{LL} = 0$ when $\gamma_2 = 1$.

Overall, if $\zeta < \min\{\varepsilon_{LL}, \varepsilon_H\}$, then $q_{HL} > q_{HH}$ together with $r_{HL} > r_{LL}$ so that $\gamma_3 = 1$. We hereafter show that neither inequality can be satisfied without satisfying the other as well.

(I) Suppose $\exists\varepsilon$ such that $q_{HH} > q_{HL}$ and $r_{LL} < r_{HL}$, involving that $\varepsilon_H < \varepsilon_{LL}$ and $\varepsilon \in (\varepsilon_H, \varepsilon_{LL})$. Then:

$$\Delta\theta(q_{HH} - q_{HL}) > \Delta\sigma(r_{LL} - r_{HL}). \tag{45}$$

Recall that $\varepsilon_{LL} \leq 0 < \zeta_1$, where $\varepsilon_{LL} > \varepsilon$ under our hypothesis, and that $\beta > 0$ for $\zeta < \zeta_1$. From (7d) and (9) and from (45) it follows that $\gamma_1 = 1$, hence $\gamma_2 = 0$. Then, the wedge:

$$(\underline{a}_{HL} - \underline{a}_{HH}) - (\overline{a}_{HH} - \overline{a}_{HL}) = \beta\gamma_2 + \gamma_3 - \frac{(1-\beta)\mu}{1-\mu} \equiv \chi$$

reduces to $-\frac{(1-\beta)\mu}{1-\mu} < 0$, implying $\varepsilon_H > 0$. As $\varepsilon < \varepsilon_{LL}$ and $\varepsilon_{LL} \leq 0 < \varepsilon_H$, the hypothesis that $q_{HH} > q_{HL}$ leads to a contradiction. Hence, $\nexists\varepsilon$ such that $q_{HH} > q_{HL}$ and $r_{LL} < r_{HL}$. It is: $\varepsilon_H \geq \varepsilon_{LL}$.

(II) Suppose $\exists\varepsilon < \zeta_1$ such that $q_{HH} < q_{HL}$ and $r_{LL} > r_{HL}$, involving that $\varepsilon_H > \varepsilon_{LL}$ and $\varepsilon \in (\varepsilon_{LL}, \varepsilon_H)$. Then:

$$\Delta\sigma(r_{LL} - r_{HL}) > \Delta\theta(q_{HH} - q_{HL}). \tag{46}$$

Knowing that $\beta = 1$ for $\varepsilon < \varepsilon_H$, from the expressions of the rents and from (46) it follows that $\gamma_2 = 1$ so that $\varepsilon_{LL} = 0$. Moreover, $\chi = 1$ involving that $\varepsilon_H < 0$, hence the interval $(\varepsilon_{LL}, \varepsilon_H)$ does not exist. Therefore, $\varepsilon_H \leq \varepsilon_{LL}$.

From (I) and (II), it follows that $\varepsilon_{LL} = \varepsilon_H$, implying that $r_{LL} \geq r_{HL}$ if and only if $q_{HH} \geq q_{HL}$. It further follows that, if $\varepsilon < \varepsilon_H$, then $\gamma_3 = 1$. As $\beta = 1$, it is $\chi = 1$ involving that $\varepsilon_H < 0$.

Proof of (ii)

Take $\varepsilon_H < \zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) \leq \zeta_1$ so that $\gamma_3 = 0$ and $\beta > 0$. Recall from *Proof of* (i) that $\gamma_1 < 1$ because the condition $q_{HH} > q_{HL}$ is otherwise contradicted. Then, $\gamma_2 > 0$. Supposing that $\gamma_2 = 1$, $\chi = \beta - \frac{(1-\beta)\mu}{1-\mu}$. $\exists\beta > 0$ such that $\chi > 0$, hence the condition $q_{HH} > q_{HL}$ is not contradicted. Therefore, $\gamma_2 > 0$ and $\gamma_1 \geq 0$, where $\gamma_1 = 0$ if $\gamma_2 = 1$.

Proof of (iii)

Take now $\zeta \geq \zeta_1$, implying that $\beta = 0$, which means:

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

40

Suppose that $\gamma_1 = 1$. As above, we have $\chi < 0$, which contradicts the result that $q_{HH} > q_{HL}$ $\forall \zeta > \varepsilon_H$. Hence, $\gamma_1 < 1$ and so $\gamma_2 > 0$.

Suppose that $\gamma_2 = 1$, which means $r_{LL} > r_{HH}$. Using $\bar{a}_{HH} > \bar{a}_{LL}$ and $\underline{a}_{HH} < \underline{a}_{LL}, r_{LL} > r_{HH}$ if and only if:

$$\zeta \left(\underline{a}_{LL}, \bar{a}_{LL}, \underline{a}_{HH} - \underline{a}_{LL} \right) \geq \varepsilon_L.$$

Resting on:

$$\underline{a}_{HH} - \underline{a}_{LL} = \Delta\theta - \Delta\sigma + \frac{\nu}{1-\nu} \left[\frac{\Delta\theta}{1-\mu} - \left(\gamma_2 + \frac{\mu}{1-\mu} \right) \Delta\sigma \right] - \gamma_2 \frac{1-\mu}{\mu} \Delta\sigma \quad (47)$$

$$\bar{a}_{HH} - \bar{a}_{LL} = \Delta\theta + \Delta\sigma + \frac{\nu}{1-\nu} \left[\frac{\Delta\theta}{1-\mu} + \left(\gamma_2 + \frac{\mu}{1-\mu} \right) \Delta\sigma \right] + \gamma_2 \frac{1-\mu}{\mu} \Delta\sigma, \quad (48)$$

we see that $\underline{a}_{HH} - \underline{a}_{LL} < \bar{a}_{HH} - \bar{a}_{LL} \Leftrightarrow \bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL} < \bar{a}_{HH}$. Hence, $\varepsilon_L > 0$, where $\varepsilon_L = f \left(\bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL} \right) - f \left(\bar{a}_{HH} \right)$ by definition. $\blacksquare$

Lemma 4 and Lemma 5 prove point (i) in the proposition.

D.1.2   Check the solution to Γ''

With $y_{HH}(\varepsilon) = y_{HL}(\varepsilon)$, the rents in (7a) - (7d) are rewritten as:

$$\Pi_{HL} = 0; \; \Pi_{HH} = \Delta\sigma r_{HH}; \; \Pi_{LL} = \Delta\theta q_{HH} \quad (49)$$
$$\Pi_{LH} = \gamma_1 \left(\Delta\theta q_{HH} + \Delta\sigma r_{HH} \right) + \gamma_2 \left(\Delta\theta q_{HH} + \Delta\sigma r_{LL} \right).$$

The agent's expected rent amounts to: $\nu\Delta\theta q_{HH} + (1-\mu)\{[1-(1-\gamma_1)\nu]r_{HH} + \nu\gamma_2 r_{LL}\}\Delta\sigma$. Using this expression, the objective function in Γ'' is formulated as:

$$\sum_{ij\in\Upsilon} E_{ij} \left[\frac{1}{2} \left(S(\underline{y}_{ij}) + S(\bar{y}_{ij}) \right) - \left(\theta_i q_{ij} - \sigma_j r_{ij} \right) \right]$$
$$-\nu\Delta\theta q_{HH} - (1-\mu)\{[1-(1-\gamma_1)\nu]r_{HH} + \nu\gamma_2 r_{LL}\}\Delta\sigma.$$

The optimal quantities are characterized by the first-order conditions (10) and (14), together with $y_{HH}(\varepsilon) = y_{HL}(\varepsilon) \; \forall\varepsilon$.

As $y_{HH}(\varepsilon) = y_{HL}(\varepsilon) \; \forall\varepsilon$, $\gamma_3 = 0$ and $\beta = 0$. From (49), $\gamma_1 \geq 0$ is equivalent to $r_{HH} \geq r_{LL}$, where $\gamma_1 = 1$ if $r_{HH} > r_{LL}$. From the quantity solution, we get:

$$\underline{a}_{HH} - \underline{a}_{LL} = \frac{1-(1-\nu)\mu}{(1-\nu)(1-\mu)}\Delta\theta + \frac{\nu(1-\gamma_2)}{1-\nu}\Delta\sigma$$

$$\bar{a}_{HH} - \bar{a}_{LL} = \frac{1-(1-\nu)\mu}{(1-\nu)(1-\mu)}\Delta\theta - \frac{\nu(1-\gamma_2)}{1-\nu}\Delta\sigma.$$

Both $\underline{a}_{HH} > \underline{a}_{LL}$ and $\bar{a}_{HH} > \bar{a}_{LL}$. Then $r_{HH} \geq r_{LL}$ if and only if $\zeta\left(\underline{a}_{LL}, \bar{a}_{HH}, \underline{a}_{HH} - \underline{a}_{LL}\right) \leq \varepsilon_L$. We see that $\underline{a}_{HH} - \underline{a}_{LL} \geq \bar{a}_{HH} - \bar{a}_{LL} \Leftrightarrow \bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL} \geq \bar{a}_{HH}$, with strict inequality if and only if $\gamma_2 < 1$; hence $f\left(\bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL}\right) - f\left(\bar{a}_{HH}\right) \leq 0$, with strict inequality if and only if $\gamma_2 < 1$.

so that $\underline{a}_{HH} > \underline{a}_{LL}$ and $\bar{a}_{HH} > \bar{a}_{LL}$. Then, $r_{HH} \geq r_{LL}$ if and only if $\zeta\left(\underline{a}_{LL}, \bar{a}_{HH}, \underline{a}_{HH} - \underline{a}_{LL}\right) \leq \varepsilon_L$. We see that $\underline{a}_{HH} - \underline{a}_{LL} \geq \bar{a}_{HH} - \bar{a}_{LL} \Leftrightarrow \bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL} \geq \bar{a}_{HH}$, where the inequality is strict if and only if $\gamma_2 < 1$. Hence, $f\left(\bar{a}_{LL} + \underline{a}_{HH} - \underline{a}_{LL}\right) - f\left(\bar{a}_{HH}\right) \leq 0$ where the inequality is strict if and only if $\gamma_2 < 1$.

Therefore, $\gamma_1 = 1 \Leftrightarrow r_{HH} > r_{LL} \Leftrightarrow \zeta\left(\underline{a}_{LL}, \bar{a}_{HH}, \underline{a}_{HH} - \underline{a}_{LL}\right) < \varepsilon_L$, where $\varepsilon_L < 0$ when evaluated at $\gamma_1 = 1$ and $\gamma_2 = \gamma_3 = \beta = 0$; $\gamma_2 = 1 \Leftrightarrow r_{HH} < r_{LL} \Leftrightarrow \zeta\left(\underline{a}_{LL}, \bar{a}_{HH}, \underline{a}_{HH} - \underline{a}_{LL}\right) > \varepsilon_L$, where $\varepsilon_L > 0$ when evaluated at $\gamma_2 = 1$ and $\gamma_1 = \gamma_3 = \beta = 0$. This proves point (ii) in Proposition 5.

D.2 Proof of Corollary 6

In Proposition 5, one identifies the following results.

(a) At least one between IC_{LL}^{HL} and IC_{LL}^{HH} is binding; the former is slack and the second is binding when $\beta = 0$. This implies that $\zeta\left(\underline{a}_{HH}, \overline{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}\right) \geq \zeta_1$ if (13) holds. IC_{LL}^{HL} and IC_{LL}^{HH} are both binding otherwise.

(b) If $\Delta\theta < \delta\Delta\sigma$, then at least one between IC_{LH}^{HL} and IC_{LH}^{HH} is binding if and only if $\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) \leq \varepsilon_L \ (\gamma_2 \neq 1)$, whereas IC_{LH}^{LL} is binding when $\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) \geq \varepsilon_H \ (\gamma_2 > 0)$.

If $\Delta\theta \geq \delta\Delta\sigma$, then IC_{LH}^{HL} and IC_{LH}^{HH} are both binding, whereas IC_{LH}^{LL} is binding when $\zeta\left(\underline{a}_{LL}, \overline{a}_{LL}, \underline{a}_{LL} - \underline{a}_{HL}\right) > \varepsilon_L \ (\gamma_1 < 1$ so that $\gamma_2 > 0)$.

D.3 Proof of Corollary 7

The proof is immediate because the conditions in the corollary are those where $\beta = 1$ and $\gamma_2 = 1$ so that IC_{LL}^{HH} and IC_{LH}^{LL} are the only binding constraints for types LL and LH.

D.4 Proof of Proposition 6

D.4.1 Reduced problem Γ'

From the proof of Lemma 2, at the solution to Γ', $y_{HH}(-1) > y_{HL}(-1)$ and $y_{HH}(1) < y_{HL}(1)$.

Using (10) and $S'(y_{LH}(\varepsilon)) = \theta_L + \varepsilon\sigma_H$ we compute:

$$S'(y_{LL}(\varepsilon)) - S'(y_{LH}(\varepsilon)) = -\left(1 + \gamma_2 \frac{1-\mu}{\mu}\right)\varepsilon\Delta\sigma,$$

so that $\underline{a}_{LL} > \underline{a}_{LH}$ and $\overline{a}_{LL} < \overline{a}_{LH}$.

Using (10) and (11) we compute:

$$
\begin{aligned}
S'(y_{HH}(\varepsilon)) - S'(y_{LL}(\varepsilon)) &= \Delta\theta + \frac{\nu}{1-\nu}\left[\gamma_1 + (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\right]\Delta\theta \\
&+ \varepsilon\Delta\sigma - \gamma_2\frac{1-\mu}{\mu}\varepsilon\Delta\sigma + \frac{\nu}{1-\nu}(1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\varepsilon\Delta\sigma.
\end{aligned}
$$

If $\varepsilon > 0$, then this difference is positive so that $\overline{a}_{HH} > \overline{a}_{LL}$. If $\varepsilon < 0$, then the sign depends on β, γ_1 and γ_2. It is $\underline{a}_{HH} \geq \underline{a}_{LL}$ if and only if:

$$\frac{\Delta\theta}{\Delta\sigma} \geq \frac{1 + \frac{\nu}{1-\nu}(1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right) + \frac{1-\mu}{\mu}\gamma_2}{1 + \frac{\nu}{1-\nu}(1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right) + \frac{\nu}{1-\nu}\gamma_1}.$$

If $\gamma_2 = 0$, then this is clearly satisfied. Recall from Proposition 5 that $\gamma_2 > 0$ is equivalent to $\beta < 1$, involving that $r_{LL} \geq r_{HH} \Leftrightarrow y_{LL}(-1) - y_{HH}(-1) \geq y_{LL}(1) - y_{HH}(1)$. However, provided $y_{LL}(1) > y_{HH}(1)$ (as implied by $\overline{a}_{HH} > \overline{a}_{LL}$), this cannot hold true if $y_{LL}(-1) < y_{HH}(-1)$. Hence, $\underline{a}_{HH} \geq \underline{a}_{LL}$.

Using (11) and $S'(y_{LH}(\varepsilon)) = \theta_L + \varepsilon\sigma_H$ we compute:

$$
\begin{aligned}
&S'(y_{HH}(\varepsilon)) - S'(y_{LH}(\varepsilon)) \qquad\qquad\qquad\qquad\qquad\qquad\qquad (50) \\
&= \Delta\theta + \frac{\nu}{1-\nu}\left\{\left[\gamma_1 + (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\right]\Delta\theta + (1-\beta)\left(\gamma_2 + \frac{\mu}{1-\mu}\right)\varepsilon\Delta\sigma\right\}.
\end{aligned}
$$

This difference is positive $\forall \varepsilon$ so that $\bar{a}_{HH} > \bar{a}_{LH}$ and $\underline{a}_{HH} > \underline{a}_{LH}$.

We found:

$$\underline{a}_{HL} > \underline{a}_{HH} \geq \underline{a}_{LL} > \underline{a}_{LH} \quad \text{and} \quad \bar{a}_{LL} < \bar{a}_{LH} < \bar{a}_{HH}; \; \bar{a}_{HL} < \bar{a}_{HH},$$

from which we derive the quantity ranking in Lemma 1.

D.4.2 Reduced problem Γ''

Using (10) and (14) we compute:

$$
S'(y_{HH}(\varepsilon)) - S'(y_{LL}(\varepsilon)) \;=\; \frac{(1-\nu)(1-\mu)+\nu}{(1-\nu)(1-\mu)}\Delta\theta
$$
$$
+ \left(\gamma_2 \frac{\nu\mu + (1-\nu)(1-\mu)}{(1-\nu)\mu} - \frac{\nu}{1-\nu}\right)\varepsilon\Delta\sigma
$$

If $\varepsilon = 1$, then this is strictly positive so that $\bar{a}_{HH} > \bar{a}_{LL}$. If $\underline{a}_{HH} \leq \underline{a}_{LL}$, it involves that $r_{HH} > r_{LL}$, hence $\gamma_2 = 0$. However, when $\gamma_2 = 0$ the difference above is positive, contradicting the hypothesis that $\underline{a}_{HH} \leq \underline{a}_{LL}$. Hence, $\underline{a}_{HH} > \underline{a}_{LL}$.

Using (14) and $S'(y_{LH}(\varepsilon)) = \theta_L + \varepsilon\sigma_H$ we compute:

$$
S'(y_{HH}(\varepsilon)) - S'(y_{LH}(\varepsilon)) = \frac{1}{1-\nu}\left(\frac{(1-\nu)(1-\mu)+\nu}{1-\mu}\Delta\theta - (1-\gamma_2\nu)\varepsilon\Delta\sigma\right)
$$

If $\varepsilon < 0$, then this difference is positive. If $\varepsilon = 1$, then it is positive if:

$$
\frac{1}{1-\mu} \geq 1 - \gamma_2,
$$

which is true. Hence, $\underline{a}_{HH} > \underline{a}_{LH}$ and $\bar{a}_{HH} > \bar{a}_{LH}$.

Recalling that quantities are as in Γ' for types LL and LH, and that $y_{HH}(\varepsilon) = y_{HL}(\varepsilon) \, \forall\varepsilon$, the ranking conditions in Lemma 1 are satisfied.

D.4.3 The necessary conditions (4) hold

First take Γ'. From (12) and $S'(y_{LH}(\varepsilon)) = \theta_L + \varepsilon\sigma_H$ we compute:

$$
S'(y_{HL}(\varepsilon)) - S'(y_{LH}(\varepsilon)) \;=\; \Delta\theta + \varepsilon\Delta\sigma + \frac{\nu}{1-\nu}\left[\beta + (\gamma_2\beta + \gamma_3)\frac{1-\mu}{\mu}\right]\Delta\theta
$$
$$
- \frac{\nu}{1-\nu}\left[1 - \beta + (1-\nu\gamma_2\beta)\frac{1-\mu}{\nu\mu}\right]\varepsilon\Delta\sigma,
$$

which is positive if $\varepsilon < 0$, and either positive or negative otherwise. Hence, $\underline{a}_{HL} - \underline{a}_{LH} > 0$ and $\bar{a}_{HL} - \bar{a}_{LH} \lesseqgtr 0$. If $\bar{a}_{HL} - \bar{a}_{LH} > 0$, then $q_{LH} > q_{HL} \, \forall\zeta(\cdot,\cdot,\cdot)$; using (5), $r_{LH} \geq r_{HL}$ if and only if (22) is satisfied.

As $r_{LH} \geq r_{LL} \, \forall\zeta(\cdot,\cdot,\cdot)$ and $r_{HL} \geq r_{LL}$ if and only if $\zeta(\underline{a}_{HH}, \bar{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) \leq \varepsilon_H$, we deduce that $\varepsilon_{LH} < \varepsilon_H$. Moreover, because $\beta = \gamma_3 = 1$ when $\zeta(\underline{a}_{HH}, \bar{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) \leq \varepsilon_H$, ε_{LH} is evaluated at these values. If $\bar{a}_{HL} - \bar{a}_{LH} < 0$, then $r_{LH} > r_{HL} \, \forall\zeta(\cdot,\cdot,\cdot)$; using (6), $q_{LH} \geq q_{HL}$ if and only if (23) is satisfied. Since $q_{LH} \geq q_{HH} \, \forall\zeta(\cdot,\cdot,\cdot)$ and $q_{HL} \geq q_{HH}$ if and only if $\zeta(\underline{a}_{HH}, \bar{a}_{HL}, \underline{a}_{HL} - \underline{a}_{HH}) \leq \varepsilon_H$, we deduce that also ε'_{LH} is such that $\varepsilon'_{LH} < \varepsilon_H$. As above, ε'_{LH} is evaluated at $\beta = \gamma_3 = 1$.

Next take Γ''. We previously found that $\underline{a}_{HH} > \underline{a}_{LH}$ and $\bar{a}_{HH} > \bar{a}_{LH}$. Knowing that $\underline{a}_{HH} = \underline{a}_{HL}$ and $\bar{a}_{HH} = \bar{a}_{HL}$, we deduce that $q_{LH} > q_{HL} \, \forall\zeta(\cdot,\cdot,\cdot)$. As above, $r_{LH} \geq r_{HL}$ if and only if

(22) is satisfied.

E   More than four types

Index types by $h \in \{1, 2, 3\}$ and consider the case in which type $h - 1$ is more efficient than type h, $\forall h \in \{2, 3\}$. The incentive constraints whereby type 1 is unwilling to report 3 and 2 are respectively:

$$\Pi_1 \geq \Pi_3 + (\theta_3 - \theta_1) q_3 + (\sigma_1 - \sigma_3) r_3 \tag{51}$$
$$\text{and}$$
$$\Pi_1 \geq \Pi_2 + (\theta_2 - \theta_1) q_2 + (\sigma_1 - \sigma_2) r_2. \tag{52}$$

The incentive constraint whereby type 2 is unwilling to report 3 is:

$$\Pi_2 \geq \Pi_3 + (\theta_3 - \theta_2) q_3 + (\sigma_2 - \sigma_3) r_3.$$

Using this constraint in (52), we further see that:

$$\Pi_2 + (\theta_2 - \theta_1) q_2 + (\sigma_1 - \sigma_2) r_2$$
$$\geq \Pi_3 + (\theta_3 - \theta_2) q_3 + (\sigma_2 - \sigma_3) r_3 + (\theta_2 - \theta_1) q_2 + (\sigma_1 - \sigma_2) r_2$$
$$= \Pi_3 + (\theta_3 - \theta_1) q_3 + (\sigma_1 - \sigma_3) r_3 + (\theta_2 - \theta_1) (q_2 - q_3) + (\sigma_1 - \sigma_2) (r_2 - r_3).$$

(52) is tighter than (51) if:

$$(\theta_2 - \theta_1) (q_2 - q_3) \geq (\sigma_1 - \sigma_2) (r_3 - r_2)$$

or, equivalently:

$$[\theta_2 - \theta_1 + \sigma_1 - \sigma_2] (y_2 (-1) - y_3 (-1)) \geq [\theta_2 - \theta_1 - (\sigma_1 - \sigma_2)] (y_3 (1) - y_2 (1)). \tag{53}$$

We distinguish between the following three cases:

- When $\theta_2 - \theta_1 \geq -(\sigma_2 - \sigma_1) \geq 0$, (53) is satisfied if $y_2(1) \geq y_3(1)$ and $y_2(-1) \geq y_3(-1)$, which is true when it is also $\theta_3 - \theta_2 \geq -(\sigma_3 - \sigma_2) \geq 0$ and the quantities respect the first-best ranking.

- When $\theta_2 - \theta_1 \geq \sigma_2 - \sigma_1 \geq 0$, (53) is satisfied if $y_2(1) \geq y_1(1)$ and $y_2(-1) \geq y_1(-1)$, which is true when it is also $\theta_3 - \theta_2 \geq 0 \geq \sigma_2 - \sigma_3$ and quantities respect the first-best ranking.

- When $0 \leq \theta_2 - \theta_1 \leq \sigma_2 - \sigma_1$, (53) is satisfied if $y_2(-1) \leq y_3(-1)$ and $y_3(1) \leq y_2(1)$, which is true when it is also $0 \leq \theta_3 - \theta_2 \leq \sigma_3 - \sigma_2$ and quantities respect the first-best ranking.

Next consider the case in which type h is more efficient than type $h - 1$, $\forall h \in \{2, 3\}$. The incentive constraints whereby type 3 is unwilling to report 2 and 1 are respectively:

$$\Pi_3 \geq \Pi_2 + (\theta_2 - \theta_3) q_2 + (\sigma_3 - \sigma_2) r_2 \tag{54}$$
$$\text{and}$$
$$\Pi_3 \geq \Pi_1 + (\theta_1 - \theta_3) q_1 + (\sigma_3 - \sigma_1) r_1. \tag{55}$$

The incentive constraint whereby type 2 is unwilling to report 1 is:

$$\Pi_2 \geq \Pi_1 + (\theta_1 - \theta_2) q_1 + (\sigma_2 - \sigma_1) r_1.$$

Using this constraint in (55), we further see that:

$$
\begin{aligned}
&\Pi_2 + (\theta_2 - \theta_3)\, q_2 + (\sigma_3 - \sigma_2)\, r_2 \\
\geq\ & \Pi_1 + (\theta_1 - \theta_2)\, q_1 + (\sigma_2 - \sigma_1)\, r_1 + (\theta_2 - \theta_3)\, q_2 + (\sigma_3 - \sigma_2)\, r_2 \\
=\ & \Pi_1 + (\theta_1 - \theta_3)\, q_1 + (\sigma_3 - \sigma_1)\, r_1 + (\theta_3 - \theta_2)\, (q_1 - q_2) + (\sigma_3 - \sigma_2)\, (r_2 - r_3) \\
>\ & \Pi_3 + (\theta_3 - \theta_1)\, q_1 + (\sigma_3 - \sigma_1)\, r_1,
\end{aligned}
$$

hence (54) is tighter than (55).