Contract design with countervailing incentives, correlated information and limited liability

Daniel Danau* Annalisa Vinella†

Abstract

In a principal-agent relationship with limited liability on the agent's side, the incentive scheme that implements the first-best outcome for the widest class of technologies is structured as follows. The agent is assigned a reward for one sole realization of an *ex post* observable signal correlated with the marginal cost (his type), and the largest feasible loss for all the other realizations. When the agent's technology includes a fixed cost inversely related with the type, leading to countervailing incentives, first best is not at hand, even with large liability, unless the fixed cost is weakly convex in type. With tight liability and countervailing incentives, conditioning the agent's reward on one sole signal realization allows the principal to still implement first best for some intermediate types, as long as the fixed cost is convex. When it is concave, the second-best output profile is a fully separating one. The same occurs with a linear fixed cost, unless the agent is to break even *ex post*. In the latter case, pooling is induced for some intermediate types as a means to secure the agent's participation.

Keywords: Countervailing incentives; Limited liability; Informative signals; Pooling

J.E.L. Classification Numbers: D82

*Bournemouth University, Business School, Executive Business Centre, 89 Holdenhurst Road, Bournemouth, Dorset, BH8 8EB, UK. E-mail: d.danau@bournemouth.ac.uk

†University of Bari, Department of Economics and Mathematics, Via C. Rosalba 53, 70124 Bari, Italy. E-mail: a.vinella@dse.uniba.it

Introduction

Countervailing incentives arise in an agency relationship when the agent has the temptation either to overstate or to understate his private information (the type) in the report to the principal, depending upon its specific realization. In regulatory contexts, for instance, a firm may display countervailing incentives *vis-à-vis* the regulator when its production technology is such that the fixed cost depends negatively on the privately known marginal cost (Lewis and Sappington [9]), which is a plausible case as, in regulated sectors, low marginal costs are likely associated with high overhead costs (Maggi and Rodriguez-Clare [11]). Moreover, a universal service provider possibly exhibits countervailing incentives when the regulator sets the financial compensation for the universal service obligation by taking the profit foregone on the unregulated market as a measure of the provider's outside opportunity, which is thus type-dependent (Poudou *et alii* [12]).[1]

So far, the literature on countervailing incentives in agency problems has studied contract design in environments where no information correlated with the agent's type, is at the principal's hand. However, in regulatory settings, it is typically the case that some piece of correlated information becomes publicly available after the contract between the firm/agent and the regulator/principal is drawn up. An agency that regulates various firms simultaneously, tying them all in a unique grand-contract, learns more about one of the firms at the time that she observes the behaviour of the others (say, a report or a contractual choice), provided they share information stochastically.[2] A regulator that imposes a universal service obligation on an incumbent providing network access to competitors on the unregulated market, is likely to extract information about the regulated firm after observing the unregulated market performance. Thus, a more realistic and comprehensive approach should account for the possibility of correlated information being present in situations where countervailing incentives appear.

In this article, we characterize the optimal contract between a principal and an agent displaying countervailing incentives, in situations where the agent's type is indeed correlated with some external signal, which becomes publicly observable after the contract has been signed and the agent has reported his type to the principal (or made a choice within the menu of contracts offered by the latter).[3] In so doing, we focus on situations in which the agent

[1] Although the focus of our study is on regulatory settings, examples of countervailing incentives are found in various other contexts as well. For instance, they arise in vertical relationships in which retailers need to specialize some assets before contracting with the upstream suppliers (Acconcia *et alii* [1]); in agency conflicts on investment levels between uninformed shareholders and informed managers (Degryse and de Jong [5]); in linear and nonlinear pricing (Jullien [8]); in landowner-farmer relationships with up-front capital endowments (Lewis and Sappington [10]).

[2] In markets where direct competition among firms is either unfeasible or weak, regulators can harness competitive forces to provide discipline by basing each firm's compensation on its performance (or report) relative to the performance (or reports) of other firms that operate either in similar markets or in the same market (see Armstrong and Sappington [2] for an overview). Yardstick schemes are especially useful when data sources are limited (Vogelsang [15] - [16]).

[3] Those mentioned in the text are situations in which correlated information appears naturally after contracting, which explains why we concentrate on *ex post* signals. In other situations, correlated information

is protected by limited liability. This allows us to make the study truly positive, provided regulated firms are generally prevented from incurring financial difficulties in order to avoid activity interruptions. By bringing correlated information and limited liability into the analysis, we thus capture two important aspects of various real-world agency relationships characterized by the presence of countervailing incentives. These aspects being obviously not innocuous in terms of the strategies and achievements that are at the principal's hand, our study sheds light on the way that contractual performance is thereby affected.[4]

As standard in the literature, also in our setting correlated information enhances contracting (recall the pioneer article of Crémer and McLean [4]), whereas limited liability contains such benefits and, particularly, the scope for retaining rents (Demougin and Garvie [6], Gary-Bobo and Spiegel [7]). Nonetheless, the specific characterization of the optimal contract in the problem that we tackle, cannot be directly grasped from previous research, for the reasons illustrated hereafter.

On the one hand, in adverse selection problems with systematic (rather than countervailing) incentives to misrepresent type, the principal benefits most from correlated information when she offers a compensation scheme under which each type is assigned a reward for one signal realization and a loss for all other realizations. Indeed, as shown by Gary-Bobo and Spiegel [7], limited liability constraints, in that case, are most easily satisfied. This scheme, to be denominated the *M*inimum-*F*easible-*L*oss scheme (MFL) throughout the paper, is due in its original version to Riordan and Sappington [13], who evidence that its performance depends upon the shape of the agent's cost function. On the other hand, also in the presence of countervailing incentives, the principal's ability to screen types depends crucially upon the shape of the agent's cost function (Maggi and Rodriguez-Clare [11]).[5] Therefore, in frameworks where correlated information (under limited liability) is combined with countervailing incentives, the optimal contract will definitely depend upon the shape of the cost function. Nevertheless, the exact way this dependence will be realized is not apparent, unless a targeted analysis is developed. In particular, in settings with countervailing incentives but no correlated information, screening is especially costly to the principal when the cost function is strictly concave in type, in which case pooling is induced. Without a specific analysis, it is not clear for which level of liability and under which cost conditions correlated

could become available prior to contracting, either naturally or because the perspective of new correlated information appearing at a later stage induces the principal to delay contracting deliberately. However, regulators may not be in a position to postpone contracting with regulated firms that provide services of general interest.

[4] Vinella [14] considers a principal dealing with two limitedly liable agents who may both display countervailing incentives and whose types are correlated. While she focuses on the simple two-type case and compares Bayesian with dominant-strategy implementation, we work with a continuum of types in a Bayesian environment. As it will become apparent at a later stage, the results that we obtain could not be derived with a binary information structure.

[5] To be precise, Maggi and Rodriguez-Clare [11] show this result with regards to the agent's utility function in a context with type-dependent reservation utility (a context of this kind is also considered in Acconcia *et alii* [1], Brainard and Martimort [3], Jullien [8], Lewis and Sappington [10]). However, analogous result entails if countervailing incentives appear because the fixed cost depends upon the privately known marginal cost (as in Lewis and Sappington [9]).

information will be so helpful that pooling is completely removed in the optimal contract.

We address these issues, focusing on environments where countervailing incentives rest on the agent's technology including a fixed cost that declines with the privately known marginal cost (rather than on his reservation utility being type-dependent), as in Lewis and Sappington [9]. Moreover, we make the reasonable assumption that the conditional likelihood of the reward signal increases as the type becomes less efficient, but at a decreasing rate, in the same vein as Riordan and Sappington [13] and Gary-Bobo and Spiegel [7].

Our first (preliminary) result is that, in the framework that we consider, the MFL entails the first-best outcome (namely, efficiency and full surplus extraction) if and only if the (fixed) cost function is non-concave in type. Comparing this finding with the one that Riordan and Sappington [13] derive in a framework with systematic incentives to lie, it emerges that, under the MFL, the presence of countervailing incentives tightens the condition on costs for first-best implementation. Riordan and Sappington [13] show, indeed, that first best is at hand if the agent's cost function is less concave in type than is the conditional likelihood function of the relevant signal. The difference between our result and theirs is explained as follows. When fixed and (privately known) marginal cost are inversely related, the worst type from the principal's perspective (*i.e.*, the one that produces at highest *total* costs) corresponds to an "intermediate" (rather than an extreme) marginal cost value. The marginal gain from mis-reporting decreases as the report approaches this (least efficient) type, either from below or from above, and vanishes when this type is announced. Therefore, under the MFL, this type is assigned a payoff equal to zero whatever the signal. That is, correlated information plays no role in the payoff profile designed for the least efficient type. Because of this, the principal might be unable to remove incentives of non-neighboring types to mimic this particular type, without conceding an information rent. This occurs, in fact, when the cost function is concave, in which case first best is beyond reach under the MFL.

Our second result is that an incentive scheme that enforces first best more often (*i.e.*, also with a non-concave cost), can be designed if and only if countervailing incentives do *not* arise. This means that countervailing incentives exacerbating the requirements for first-best implementation is not a limit specific to the MFL. To see how the alternative scheme is structured and why it only performs better absent countervailing incentives, one should consider that, under the MFL, the compensation targeted to any given type is not sufficiently uncertain to remove the other types' temptation to announce this specific type. This follows because all agent's types are inflicted the minimum feasible loss. To lessen this difficulty, the principal should expose all types to higher uncertainty. We show that each type faces highest uncertainty when it is still assigned one reward and equal losses, but the latter are fixed at the Maximum (incentive-)Compatible level that the agent can bear under limited liability. This scheme, to be denominated MCL throughout the paper, is the one that enlarges most the class of technologies for which first best is at hand. The reason why this outcome does not extend to frameworks with countervailing incentives is that, while, under the MCL, all other types do face higher uncertainty, the least efficient type is still optimally assigned the all-zero-payoffs contract. Therefore, with regards to this particular type, the MCL displays exactly the same limit as the MFL.

Moving to situations in which limited liability prevents implementation of the efficient outcome under both the MFL and the MCL, we first show that there exist some types for which first best is still effected, in fact. This outcome, which rests on the possibility of exploiting informative signals by inflicting (bounded) losses, entails if and only if the cost function is weakly convex and is peculiar to environments with countervailing incentives. Specifically, first best survives for a continuum range of intermediate types neighboring the least efficient one. This occurs because, when the fixed cost function is weakly convex, the gain from cheating is low enough so that, by assigning small losses and rewards to intermediate types, it is possible to remove both their own incentives to mimic other types and the others' incentives to mimic these types. Their limited liability constraints are thus irrelevant. However, as one moves away from intermediate cost values, mis-reporting becomes sufficiently more attractive to prevent first-best implementation for all remaining types. For types immediately below and above the intermediate ones, the quantity is distorted just enough to retain all surplus and, at the same time, to elicit information and satisfy the limits on liability. On the other hand, an information rent is conceded to very low and very high types. As usual, this rent is contained by distorting the quantity till the ensuing loss exactly compensates the surplus-extraction gain (the familiar efficiency/rent-extraction trade-off).

The last relevant situation is that in which the cost function is concave in type. In this situation, the incentive problem is exacerbated to the point that both the MFL and the MCL fail to effect first best, as already mentioned. Presumably, an alternative scheme would be required, less parsimonious in terms of signals to be used. Nonetheless, in a second-best scenario with tight limited liability, a one-reward scheme still appears to be necessary because, by spreading losses over as many signal realizations as possible, the principal can elicit information at a lower cost. We show that the optimal second-best scheme under which, for all signal realizations but one, the agent bears the largest admissible loss, displays a critical difference with respect to the contract that would be optimal in the absence of correlated signals. From Lewis and Sappington [9] and Maggi and Rodriguez-Clare [11], we know that information would not be revealed, in that case, unless the principal were to pool quantities for a range of intermediate cost values. In our setting, the possibility of conditioning the agent's compensation on a piece of correlated information completely removes the need to introduce an inflexible rule compelling different types to produce all the same quantity. Importantly, this result holds however deep the agent's pocket is. That is, the principal designs a fully separating output profile whatever the agent's liability.

Overall, our study identifies one sole situation in which pooling arises in correlated-information frameworks. This is the specific situation in which the fixed cost is linear in type and the agent cannot be exposed to any deficit *ex post*. Noticeably, the reason for which pooling is induced in this case, is not the same as in environments without informative signals. In the latter, it is dictated by the impossibility to make the contract globally incentive compatible with a fully separating output profile. In correlated-information settings, it is a way to secure participation without leaving any rent to the agent.

The reminder of the article is organized as follows. In section 2, we present the model. Section 3 focuses on first-best implementation with large liability. In section 4, we move to situations with tight limited liability and characterize the optimal contract for different shapes of the cost function. As a conclusive step, in section 5, we compare our findings, particularly the pooling result, with those emerging in frameworks without correlated information. Mathematical details are relegated to the Appendix.

1 The model

A risk-neutral principal P contracts with a risk-neutral agent for the production of q units of some good. Production costs are given by

$$C(q, c) = cq + K(c), \tag{1}$$

where c is the marginal cost and $K(c)$ the fixed cost that depends upon c. The function $K(\cdot)$ is twice continuously differentiable. Similarly to Lewis and Sappington [9], we take it to be decreasing in marginal cost ($K'(c) < 0$). However, while these authors focus on concavity, we only assume that $K''(c)$ maintains the same sign for all values of c. This facilitates the exposition but does not affect results qualitatively.

At the contracting stage, the agent is privately informed about c (the type). It is commonly known that c is drawn from the continuous support $[\underline{c}, \overline{c}]$, $\underline{c} > 0$, with density function $f(c)$ and cumulative distribution function $F(c)$, both continuously differentiable. The marginal cost is correlated with a random signal s that is purely informational as in Riordan and Sappington [13] and in Demougin and Garvie [6]. The signal is drawn from the discrete support $N \equiv \{1, .., n\}$ and publicly observed after the contract has been signed and the agent has chosen how much to produce. We denote $p(c, s) = \Pr ob(s \,|c)$ the probability of observing the signal s conditional on type being c. We take $p(c, s)$ to be twice continuously differentiable everywhere with respect to c.

1.1 The programme of the principal

As usual, the Revelation Principle applies and attention can be restricted to direct revelation mechanisms in which the agent reports his true type. P can condition the compensation to the agent on the realization of the signal, which is observed at the end of the production period. Thus, a mechanism designed for an agent of type c and some signal s is an allocation

$\{q(c), t(c,s)\}$, with $q(c)$ the quantity to be produced and $t(c,s)$ the transfer to be paid.[6] Under truthful reporting, the agent's *ex post* and *interim* profits are given by

$$\pi(c,s) = t(c,s) - [cq(c) + K(c)] \tag{2}$$
$$E_s[\pi(c,s)] \equiv \sum_{s=1}^{n} \{t(c,s) - [cq(c) + K(c)]\} p(c,s), \tag{3}$$

respectively. Truthful reporting in a Bayesian setting is induced by satisfying the following incentive constraints

$$E_s[\pi(c,s)] \geq \sum_{s=1}^{n} \{t(r,s) - [cq(r) + K(c)]\} p(c,s), \ \forall c \in [\underline{c}, \overline{c}]. \tag{IC}$$

Besides, P needs to satisfy the participation constraints

$$E_s[\pi(c,s)] \geq 0, \ \forall c \in [\underline{c}, \overline{c}], \tag{PC}$$

and the limited liability constraints

$$\pi(c,s) \geq -L, \ \forall c \in [\underline{c}, \overline{c}], \ \forall s \in N, \tag{LL}$$

for some given $L \geq 0$.

Let $S(q(c))$ the gross utility that P obtains when $q(c)$ units of the good are provided, with $S(0) = 0$, $S' > 0$, $S'' < 0$, $S'(0) = +\infty$ and $S'(+\infty) = 0$. The objective of P is to achieve the highest attainable level of utility. The latter is taken to be a weighed sum of gross utility net of transfer, namely $V(q(c)) = S(q(c)) - t(c,s)$, and the agent's profit. Formally, the programme of P is written:

$$\underset{\{q(c);\pi(c,s)\}}{Max} \widetilde{W} \equiv \int_{\underline{c}}^{\overline{c}} \sum_{s=1}^{n} [V(q(c)) + \alpha\pi(c,s)] p(c,s) f(c) dc \quad \text{subject to (IC), (PC) and (LL),} \tag{Γ}$$

with $\alpha \in [0,1]$.[7]

[6] In Gary-Bobo and Spiegel [7], not only the payment but also the quantity is conditioned on the signal realization because the latter is a shock that affect the agent's cost/type. However, the circumstance that the signal is not purely informational in their model has no qualitative implications on the comparison between our solution and theirs.

[7] As standard in the literature, in a regulation context, the transfer to the agent (the regulated firm) can be thought of as made out of the public budget if the product is a public good. In the case of a private good, $S(\cdot)$ can be interpreted as the gross consumer surplus for the product (*i.e.*, the integral of the inverse demand function) and the transfer to the agent as including both the usage fees and the fixed fees paid by consumers (or, as an alternative to the latter, a subsidy made out of the public budget). In a procurement context, a natural choice would be to set $\alpha = 0$.

2 First-best implementation

The first-best outcome (FB hereafter) entails whenever P can design transfers $t^{fb}(c,s)$ such that, for all cost levels, she induces information revelation without incurring any agency cost $\left(E_s[\pi^{fb}(c,s)] = 0\right)$ and without distorting production away from the socially efficient level $\left(S'(q^{fb}(c)) = c\right)$. In this section, we explore in which ways and under which conditions this is feasible.

2.1 Preliminary analysis: the Minimum-Feasible-Loss scheme

To make (LL) most likely satisfied, a natural strategy for P is to offer the mechanism that minimizes the loss to be assigned to the agent for all possible types. We begin by showing how such a mechanism is to be designed in environments characterized by the presence of countervailing incentives to misrepresent information. In so doing, we extend the analysis that Gary-Bobo and Spiegel [7] develop for the case of systematic incentives to over-report, to settings where different types may want to lie in different directions.

Under FB implementation, (IC) is conveniently replaced by the pair of conditions (as derived in Appendix A.1)

$$q^{fb}(c) + K'(c) + \sum_{s=1}^{n} \pi^{fb}(c,s) \frac{dp(c,s)}{dc} = 0 \qquad \text{(LIC)}$$

$$\sum_{s=1}^{n} \left\{ t^{fb}(r,s) - cq^{fb}(r) - K(c) \right\} p(c,s) \leq 0, \qquad \text{(GIC)}$$

where $t^{fb}(r,s)$ is the FB transfer that P makes to the agent when the latter reports r and s is observed. (LIC) warrants that the agent has no incentive to report $r \neq c$ in a neighborhood of his true type c (*local* incentive compatibility). (GIC) ensures that the agent has no interest in reporting any $r \neq c$ within the entire feasible set (*global* incentive compatibility).

Consider now the reduced programme in which (LL) and (GIC) are neglected. It is defined as follows:

$$\underset{\{q(c);\pi(c,s)\}}{Max} \ \widetilde{W} \qquad (\Gamma')$$
$$s.t. \ \text{(LIC) and (PC)}.$$

Let any solution to (Γ') be the vector of profits $\Pi^{fb}(c) \equiv \left\{\pi^{fb}(c,1), ..., \pi^{fb}(c,n)\right\}$ together with the quantity $q^{fb}(c)$ for each given c. As there are more combinations of profits $\pi^{fb}(c,s)$ that solve (Γ') for each given c, define Ω the set of all vectors $\Pi^{fb}(c)$. Further define Φ the set that contains the lowest element of each vector $\Pi^{fb}(c)$. Lastly, let $\pi^*(c)$ be the largest element of Φ. This is the minimum feasible loss under which FB is implemented for any given c. Once $\pi^*(c)$ is identified, it is also possible to identify the specific set of profits, among all those in Ω, to which it belongs. Let $\Pi^*(c)$ denote such a set. This is the set of

FB profits under which (LL) is least likely to be binding in the original programme (Γ). From now on, we refer to the incentive scheme that implements FB with profits in $\Pi^*(c)$ as to the "Minimum-Feasible-Loss" scheme (MFL hereafter).

Lemma 1 *(Gary-Bobo and Spiegel [7]) Under the MFL, for all $c \in [\underline{c}, \overline{c}]$, the agent is rewarded whenever s takes some given value $\widetilde{s}(c)$ and bears the smallest feasible loss whenever $s \neq \widetilde{s}(c)$, the loss being equal in size for all $s \neq \widetilde{s}(c)$. The optimal payoffs are given by:*

$$\pi^{fb}(c,s)\big|_{s=\widetilde{s}(c)} = \left[q^{fb}(c) + K'(c)\right] \frac{1 - p(c, \widetilde{s}(c))}{dp(c, \widetilde{s}(c))/dc} \equiv \overline{\pi}^{fb}(c, \widetilde{s}(c)) \quad (4)$$

$$\pi^{fb}(c,s)\big|_{s\neq\widetilde{s}(c)} = \left[q^{fb}(c) + K'(c)\right] \frac{-p(c, \widetilde{s}(c))}{dp(c, \widetilde{s}(c))/dc} \equiv \underline{\pi}^{fb}(c, \widetilde{s}(c)). \quad (5)$$

By this lemma, under the MFL, the set $\Pi^*(c)$ reduces to only two values for each type c. Gary-Bobo and Spiegel [7] obtain this result for the case of $K'(\cdot) = 0$ *i.e.*, with systematic incentives to exaggerate cost. As the authors explain, spreading punishments over as many realizations of s as possible (*i.e.*, all feasible realizations but one) allows P to minimize the highest possible loss for each type of agent. This requires that the largest reward-loss wedge that can be realized over all possible realizations of s be minimized. Being independent of the characteristics of the cost function, this result is also valid in the presence of countervailing incentives (and, of course, in the presence of systematic incentives to under-report). Noticeably, the difference $\left[\overline{\pi}^{fb}(c, \widetilde{s}(c)) - \underline{\pi}^{fb}(c, \widetilde{s}(c))\right]$ is the lowest feasible wedge when FB is implemented under (LIC) and (PC), whether the technology is such that systematic or countervailing incentives to misrepresent information arise.

It is interesting to see how $\widetilde{s}(c)$ should be selected, particularly in a framework where different types may be tempted to lie in different directions. As $K'(\cdot) < 0$, the choice of $\widetilde{s}(c)$ depends upon the sign of the sum $q^{fb}(c) + K'(c)$, which may not be the same for all $c \in [\underline{c}, \overline{c}]$. When this sum is positive, the situation is similar to that Gary-Bobo and Spiegel [7] consider and the same result obtains. For (5) to be a loss ($\underline{\pi}^{fb}(c, \widetilde{s}(c)) < 0$), $\widetilde{s}(c)$ must be such that the probability of its realization raises with c (*i.e.*, $dp(c, \widetilde{s}(c))/dc > 0$). Moreover, for (5) to be the smallest feasible loss, $r(c)$ must be such that the ratio $\frac{p(c,\widetilde{s}(c))}{dp(c,\widetilde{s}(c))/dc}$ is minimized. This is tantamount to requiring that the ratio $\frac{dp(c,\widetilde{s}(c))/dc}{p(c,\widetilde{s}(c))}$ be maximized. One can interpret this result (and hence that of Gary-Bobo and Spiegel [7]) by observing that $\frac{dp(c,\widetilde{s}(c))/dc}{p(c,\widetilde{s}(c))}$ is the rate of increase of the conditional likelihood that signal $\widetilde{s}(c)$ be drawn as c increases. This means that, for any given c, $\widetilde{s}(c)$ must be the most likely signal to be drawn by *higher* types. Intuitively, because any type c that has an incentive to over-report is more likely to incur a deficit than higher types are, a smaller deficit suffices to remove the incentive to mimic of that type. Similar reasoning applies, *mutatis mutandis*, when $q^{fb}(c) + K'(c) < 0$. In that case, $\widetilde{s}(c)$ must be such that its conditional likelihood decreases with c (*i.e.*, $dp(c, \widetilde{s}(c))/dc < 0$), and the ratio $\frac{dp(c,\widetilde{s}(c))/dc}{p(c,\widetilde{s}(c))}$ is minimized. That is, for any given c, the agent is to be rewarded when the signal that is most likely to be drawn by *lower* possible types does materialize. By doing so, a smaller deficit can be imposed to remove incentives to under-report.

Remark 1 $\overline{\pi}^{fb}(\widehat{c}, s) = \underline{\pi}^{fb}(\widehat{c}, s) = 0$ *for all* $s \in N$. *There is no need to select* $\widetilde{s}(\widehat{c})$.

To identify the signal $\widetilde{s}(c)$ for all feasible $c \neq \widehat{c}$, it is useful to make the following assumptions.

Assumption 1 $K''(c) < -\frac{dq^{fb}(c)}{dc}$, $\forall c \in [\underline{c}, \overline{c}]$.

Assumption 2 *The conditional likelihood function is such that:*

$$\begin{aligned} \frac{dp(c,n)}{dc} &> 0 \quad and \quad \frac{d^2p(c,n)}{dc^2} < 0, \ \forall c \in [\underline{c}, \overline{c}] \qquad (6) \\ \frac{dp(c,1)}{dc} &< 0 \quad and \quad \frac{d^2p(c,1)}{dc^2} < 0, \ \forall c \in [\underline{c}, \overline{c}]. \qquad (7) \end{aligned}$$

Assumption 3 *The conditional likelihood function satisfies the following properties:*

$$\begin{aligned} \frac{d}{dc}\left(\frac{p(c,s)}{p(c,n)}\right) &\leq 0, \ \forall c \in [\underline{c}, \overline{c}], \ \forall s \in N \qquad (8) \\ \frac{d}{dc}\left(\frac{p(c,s)}{p(c,1)}\right) &\geq 0, \ \forall c \in [\underline{c}, \overline{c}], \ \forall s \in N. \qquad (9) \end{aligned}$$

Assumption 1 is equivalent to saying that the sum $q^{fb}(c) + K'(c)$ decreases with c. If there exists $\widehat{c} \in (\underline{c}, \overline{c})$ so that $q^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$, then the sum is positive for types below $\widehat{c}$ and negative for types above $\widehat{c}$. Otherwise, it is either positive or negative for all types. Condition (6) (resp. (7)) in Assumption 2 tells that the probability of drawing $s = n$ (resp. $s = 1$) increases (resp. decreases) with type c at a decreasing rate. Assumption 2 is thus a requirement on the behaviour of $p(c, s)$ at two values of s, which we take to be n and 1. Condition (8) (resp. (9)) in Assumption 3 requires that the rate of increase (resp. decrease) of the conditional likelihood that $s = n$ (resp. $s = 1$) be drawn is higher (resp. lower) than that of any other signal.

Lemma 2 *Suppose that there exists* $\widehat{c} \in [\underline{c}, \overline{c}]$ *so that* $q^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$. *Then, under Assumption* $1-3$, *in the MFL,* $\widetilde{s}(c) = n$ *when* $c \in [\underline{c}, \widehat{c})$ *and* $\widetilde{s}(c) = 1$ *when* $c \in (\widehat{c}, \overline{c}]$.

Unlike in environments with systematic incentives to lie, in which P can implement FB using information about a unique signal, *two* signals are necessary when the agent displays countervailing incentives.[8] It is however noteworthy that the particular choice of n and 1 as

[8] Condition (6) is analogous to the assumption that Riordan and Sappington [13] impose to ensure that P only needs to use a unique reward-signal in a framework with systematic incentives to lie. Moreover, condition (8) is analogous to the assumption that Gary-Bobo and Spiegel [7] introduce to identify the highest feasible signal as that for which the agent's loss is minimized. The necessity to refer to two signals when countervailing incentives arise, explains why we impose conditions (7) and (9), in addition to (6) and (8).

the signals that trigger a reward is without loss of generality in our model. The properties of the likelihood function in Assumption 2 and 3, which ensure that n and 1 are the optimal reward signals indeed, could refer to any other pair of cost values.

In the sequel of the analysis, we maintain the hypothesis that $\widehat{c}$ does exist, unless differently specified.

Proposition 1 *Suppose that*

$$L \geq \left[q^{fb}(c) + K'(c)\right] \frac{p(c, \widetilde{s}(c))}{dp(c, \widetilde{s}(c))/dc}, \ \forall c \in [\underline{c}, \overline{c}], \tag{10}$$

with $\widetilde{s}(c) = n$ *for* $c < \widehat{c}$ *and* $\widetilde{s}(c) = 1$ *for* $c < \widehat{c}$*. Then, under Assumption* $1-3$*, the first-best outcome is implemented with ex post payoffs* (4) *and* (5) *if and only if*

$$K''(c) \geq 0, \ \forall c \in [\underline{c}, \overline{c}]. \tag{11}$$

When either $q^{fb}(c) + K'(c) > 0$ *or* $q^{fb}(c) + K'(c) < 0$ *for all* $c \in [\underline{c}, \overline{c}]$*, condition* (11) *is replaced by*

$$K\text{"}(c) \geq \left[q^{fb}(c) + K'(c)\right] \frac{d^2p(c, \widetilde{s}(c))/dc^2}{dp(c, \widetilde{s}(c))/dc}, \ \forall c \in [\underline{c}, \overline{c}], \tag{12}$$

with $\widetilde{s}(c) = n$ *in the former case, and* $\widetilde{s}(c) = 1$ *in the latter.*

The conditions reported in Proposition 1 are explained as follows. Condition (10) rests on the circumstance that the agent cannot bear unbounded losses. The solution to (Γ) that is picked by the MFL does not implement FB unless (10) is satisfied. This condition is similar to that in Proposition 2 of Gary-Bobo and Spiegel [7]. The peculiarity here is that it specifies differently according to whether $c < \widehat{c}$ or $c > \widehat{c}$. Condition (11) shows that weak convexity of the fixed cost function is a necessary and sufficient condition for the MFL payoff profile to be globally incentive compatible in (Γ). This cost restriction is precisely due to the presence of countervailing incentives. Condition (12) shows how it would be lessened if the agent were to display systematic incentives to over/understate type whatever the true cost. Observe that (12) replicates, *mutatis mutandis*, a previous finding that we owe to Riordan and Sappington [13]. They show that, when the signal space is smaller than the type space, together with the conditions on the likelihood function of the relevant signal (the counterpart of (6) in our model, as we said), enforcement of FB requires that the agent's cost function be less concave in type than the conditional likelihood function of the relevant signal.

To see why a tighter restriction (as expressed by condition (11)) arises in the presence of countervailing incentives, notice that the transfer that an agent of type c receives when he reports r and $\widetilde{s}(c)$ is observed, is given by

$$t(r, s) = rq^{fb}(r) + K(r) + \pi(r, s).$$

This transfer is composed of two elements. The first element, namely $rq^{fb}(r) + K(r)$, is a fixed payment equal to the total cost that the agent would bear if he were of type r. The second element, namely $\pi(r, s)$, is an uncertain payment, the value of which depends upon the signal realization. Because this realization is unknown to the agent, he faces a lottery with expected value

$$\begin{aligned}\sum_{s=1}^{n} \pi(r,s)\, p(c,s) &= -\left[\overline{\pi}^{fb}(r, \widetilde{s}(c)) - \underline{\pi}^{fb}(r, \widetilde{s}(c))\right]\left[p(r, \widetilde{s}(r)) - p(c, \widetilde{s}(r))\right] \\ &= -\frac{q^{fb}(r) + K'(r)}{dp(r, \widetilde{s}(r))/dr}\left[p(r, \widetilde{s}(r)) - p(c, \widetilde{s}(r))\right].\end{aligned}$$

The introduction of this lottery is meant to offset the benefit that the agent might obtain with a convenient report, as a difference between the fixed payment and his true cost. For this to occur, the lottery should yield sufficiently high expected costs to mimicking types. This requires that the wedge between the reward and the loss designed for type r, as expressed by the ratio $\frac{q^{fb}(r)+K'(r)}{dp(r,\widetilde{s}(r))/dr}$, be large enough. Indeed, this allows P to exploit the correlation between types, as represented by the difference $[p(r, \widetilde{s}(r)) - p(c, \widetilde{s}(r))]$, to extract surplus. Recall however that, under the MFL, the wedge $\overline{\pi}^{fb}(r, \widetilde{s}(c)) - \underline{\pi}^{fb}(r, \widetilde{s}(c))$ is set at the minimum feasible level for each r and, in particular, it equals zero for $r = \widehat{c}$, type $\widehat{c}$ being the sole that always obtains a payoff equal to zero. As a result, whenever $\widehat{c}$ is reported, the lottery disappears. Under this circumstance, type $c \neq \widehat{c}$ is discouraged from reporting $\widehat{c}$ if and only if (11) is satisfied. It is thus clear that the presence of the all-zero-payoffs solution, which glues at $c = \widehat{c}$ the one-reward-and-equal-losses contracts accruing to smaller and bigger types, restricts the class of technologies for which FB can be effected.

2.2 The Maximum-Compatible-Loss scheme

The restrictions on the shape of the fixed cost function, that are expressed by conditions (11) and (12), appear because, as we pointed out, the lottery that the agent faces under the MFL is limitedly effective, at least for some of the reports that he can make. To lessen this difficulty, and thus relax (11) and (12), P should be able to design a scheme under which the agent faces a lottery that is overall more effective but still complies with (LL). We hereafter explore this possibility, which can only arise as long as (10) is satisfied.

We begin by introducing the following assumption.

Assumption 4 *There exist $i, j \in N$ such that*

$$\frac{dp(c,i)}{dc} > 0 \quad \textit{and} \quad \frac{p(c,n)}{dp(c,n)/dc} < \frac{p(c,i)}{dp(c,i)/dc} \leq \frac{L}{q^{fb}(c) + K'(c)}, \tag{13}$$

for all $c \in [\underline{c}, \overline{c}]$ *for which* $q^{fb}(c) + K'(c) > 0$, *and*

$$\frac{dp(c,j)}{dc} < 0 \quad and \quad \frac{L}{q^{fb}(c) + K'(c)} \leq \frac{p(c,j)}{dp(c,j)/dc} < \frac{p(c,1)}{dp(c,1)/dc}, \tag{14}$$

for all $c \in [\underline{c}, \overline{c}]$ *for which* $q^{fb}(c) + K'(c) < 0$.

This assumption warrants that there exist (at least) two signals, namely i and j, that P can target as reward signals to design a scheme under which losses exceed the minimum feasible amount (as identified in (5)) but not the maximum admissible level (L). To reflect this peculiarity of the scheme, we hereafter refer to it as to the "Maximum-Compatible-Loss" scheme (MCL hereafter).

Lemma 3 *Suppose that condition* (10) *holds. Then, under Assumption* $1-4$, *the ex post payoffs in the MCL are given by* (4) *and* (5), *with* $\widetilde{s}(c) = i$ *when* c *is such that* $q^{fb}(c) + K'(c) > 0$ *and* $\widetilde{s}(c) = j$ *when* c *is such that* $q^{fb}(c) + K'(c) < 0$. *In particular,* $\overline{\pi}^{fb}(\widehat{c}, s) = \underline{\pi}^{fb}(\widehat{c}, s) = 0$ *for all* $s \in N$.

In the proof of the lemma (see Appendix A.4.1), we show that, for the expected value of the lottery to be minimized for all possible reports, P should offer a scheme under which all agent's types bear a loss for all signal realizations but one, which triggers a reward. The scheme should thus be structured as the MFL. Recall, indeed, that spreading losses over as many signal realizations as possible, allows P to minimize the loss that the agent can be required to incur. Besides, our proof evidences that spreading losses yields also a second benefit to P. That is, it allows P to maximize the uncertainty to be imposed on the agent. However, at this aim, losses should be raised as much as it is compatible with (LL), rather than being minimized as in the MFL. Contrarily to what one might expect, the highest attainable loss is not L. A scheme under which all types are assigned losses equal to L cannot be made globally incentive compatible (in addition to satisfying (LL), (LIC) and (PC)). In fact, the lottery that is thereby induced is so extreme that, for certain types, it ends up creating unusual incentives to mis-report. Hence, the contract targeted to some c may become attractive for some c' that would not otherwise be tempted to mimic c. This difficulty is circumvented when the MCL is offered, in which case the agent is faced with the same pair of payoffs as under the MFL, namely (4) and (5), with the novelty that the reward signals n and 1 are replaced by i and j, respectively.

Of course, this does not ensure that, when the MCL is adopted, incentives to lie globally are removed whatever the technology used by the agent. Nonetheless, by making the expected value of the lottery as low as possible and thus exacerbating the agent's uncertainty, the MCL potentially enlarges the class of technologies for which FB is enforced, as compared to the MFL. The following proposition clarifies how the performance of the MCL varies, depending upon the specific nature of the agent's incentives to misrepresent information.

Proposition 2 *Suppose that condition* (10) *holds and that Assumption* $1-4$ *are satisfied. When either* $q^{fb}(c)+K'(c)>0$ *or* $q^{fb}(c)+K'(c)<0$ *for all* $c\in[\underline{c},\overline{c}]$, *the MCL enlarges the class of technologies for which the first-best outcome is implemented, as compared to the MFL. It does not, instead, whenever there exists* $\widehat{c}\in[\underline{c},\overline{c}]$ *so that* $q^{fb}(\widehat{c})+K'(\widehat{c})=0$.

According to the proposition, in situations where the agent is tempted to over/understate type systematically, the MCL does make the lottery more effective, as compared to the MFL, for all possible reports. This explains why the MCL enables P to enforce FB also for more concave fixed cost functions, without violating the limit on the *ex post* deficit that the agent can bear. The precise extent to which the class of technologies is enlarged, depends upon the signals that comply with Assumption 4. Condition (12) is relaxed most when i (resp. j) is such that the ratio $\frac{p(c,i)}{dp(c,i)/dc}$ (resp., $\frac{p(c,j)}{dp(c,j)/dc}$) is closest to the upper (resp. lower) bound $\frac{L}{q^{fb}(c)+K'(c)}$. This result is interesting in that, while P may have a preference for the MFL because the latter makes the financial burden as light as possible for the agent,[9] it points to a different conclusion. That is, as long as all types are tempted to lie in the same direction, the loss that the MFL yields might prove excessively low for the efficient outcome to be achieved. To lessen this difficulty, principals should thus accept to adopt a scheme that inflicts more important (though still sustainable) deficits to agents.

However, in the presence of countervailing incentives, P is unable to relax (11) even by adopting the MCL. To see why this is the case, observe that, as specified in Lemma 3, the MCL is such that type $\widehat{c}$ still receives a payoff equal to zero for all signal realizations. Hence, despite that the lottery is made more effective for reports other than $\widehat{c}$, it still vanishes for $r=\widehat{c}$. This means that not only the presence of the all-zero-payoffs solution exacerbates the restriction on the class of technologies for which the MFL effects FB, as compared to situations in which type $\widehat{c}$ is absent. It also impedes that the class of technologies for which FB is at hand, be expanded when the MFL is replaced with the MCL.

It should by now be clear that a distinction between the MFL and the MCL can only be made as long as (10) holds so that, under our assumptions, it is possible to design losses that lie between the lowest feasible amount and the largest sustainable amount. Such a distinction no longer exists when (10) is violated and we move to a second-best (SB hereafter) setting.

3 The second-best contract with tight limited liability

In this section, we consider environments in which P cannot find a profile of transfers (hence, of profits) that effect FB because the agent's pocket is not sufficiently deep.

[9] This is likely the case, for instance, if the principal is a regulator dealing with a firm that provides services of general interest and/or warrants unprofitable market coverage. Gary-Bobo and Spiegel [7] stress that regulators tend to avoid financial difficulties for the regulated firms not only to prevent activity interruptions but also because inducing deficits can be embarrassing for themselves.

3.1 Weakly convex fixed cost

Suppose that $K''(\cdot) \geq 0$ and that (LL) is so tight that condition (10) fails to hold. To identify the SB scheme, it is first essential to notice that, in fact, (10) is not violated for all feasible values of c.

Lemma 4 *Suppose that* $K''(\cdot) \geq 0$ *and that condition* (10) *does not hold (at least for some* c*). Then, under Assumption* $1-3$, *for any* $L \geq 0$, *at the solution to* (Γ), *there exists a unique range of types* $[c_2, c_3] \subseteq [\underline{c}, \overline{c}]$, *such that* $\widehat{c} \in [c_2, c_3]$, *for which the first-best outcome is implemented.*

First of all, limited liability is not an issue as far as type $\widehat{c}$ is concerned. Indeed, for this type, (10) is surely satisfied as $q^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$. Furthermore, (10) holds for the types in a neighborhood of $\widehat{c}$ *i.e.*, for all values of c for which the absolute value of $q^{fb}(c) + K'(c)$ is sufficiently low. Condition (LL) is more and more likely to be binding as c gets farther from $\widehat{c}$.

Lemma 5 *Suppose that* $K''(\cdot) \geq 0$ *and that condition* (10) *does not hold (at least for some* c*). Then, under Assumption* $1-3$, *there exists at most one cost value* $c_1 \in (\underline{c}, c_2)$ *(resp.* $c_4 \in (c_3, \overline{c})$*) such that, at the solution to* (Γ), (PC) *is slack for all* $c \in [\underline{c}, c_1)$ *(resp.* $c \in (c_4, \overline{c}]$*) and binding for all* $c \in [c_1, c_2]$, *(resp.* $c \in [c_3, c_4]$*). The cost value* c_1 *(resp.* c_4*) exists if and only if* $\frac{dp(c,n)/dc}{p(c,n)}L$ *(resp.* $\frac{dp(c,1)/dc}{p(c,1)}L$*) is sufficiently large (resp. small). When* c_1 *(resp.* c_4*) does not exist,* (PC) *is binding for all* $c \in [\underline{c}, c_2)$ *(resp.* $c \in (c_3, \overline{c}]$*).*

In the framework here considered, at the solution to (Γ), not only P enforces FB for all types in $[c_2, c_3]$. She is also able to extract all surplus from some types below c_2 and some types above c_3.[10] For any given $L > 0$, the range of types below c_2 (resp. above c_3) from which surplus is fully retained spans to the whole set $[\underline{c}, c_2]$ (resp. $[c_3, \overline{c}]$) as long as the rate of increase (resp. decrease) of the conditional likelihood that signal n (resp. 1) be drawn as c raises, is sufficiently large (resp. small). In line with the insights from the FB analysis, this means that P is more likely to induce truthtelling at zero rent the more informative the two signals are about type. When $\frac{dp(c,n)/dc}{p(c,n)}L$ (resp. $\frac{dp(c,1)/dc}{p(c,1)}L$) is not large (resp. small) enough, surplus extraction becomes unfeasible for very low and very high types, whose incentives to mis-report are most intense. In that case, these types are assigned a positive *interim* payoff.

The possibility that efficient types obtain an information rent, depending upon the conditional likelihood of signal n, is reminiscent of the findings in Gary-Bobo and Spiegel [7]. Indeed, in their model, the agent's participation constraint holds strictly for all types but the least efficient one if the derivative of the conditional likelihood function at n is small enough.[11] Observe however that, in our setting, the possibility of (PC) being slack for all

[10] Although surplus is retained, FB is not implemented for these types because quantities are distorted away from the efficient level, as it will become clear shortly.

[11] See page 5 of the technical appendix in Gary-Bobo and Spiegel [7].

types but one is ruled out due to the fact that the fixed cost decreases with type. This facilitates surplus extraction, indeed, by weakening the incentives to cheat of types sufficiently close to $\widehat{c}$.

Now suppose that the conditions described above do hold so that the cost values c_1 and c_4 exist. We shall see how the SB output is characterized in this situation. Consider that the incentive to overstate (resp. understate) type that an agent with $c < \widehat{c}$ (resp. $c > \widehat{c}$) would display if he were to receive the sole fixed payment to produce the FB quantity, are more intense for c close to $\underline{c}$ (resp. $\overline{c}$). To remove the incentive to mimic by means of the lottery, while keeping output at the FB level, P would need to progressively increase the wedge between reward and losses as c moves away from $\widehat{c}$. Nevertheless, (LL) imposes a bound on how large losses can be set, for FB does not attain when $c \notin [c_2, c_3]$. Without quantity distortions, P could elicit information only by raising the reward sufficiently, which would yield an information rent to the agent. This would be too costly though. The optimal strategy is thus to reduce the rent by fixing output away from the efficient level. For types with weak incentives to cheat, namely those in $[c_1, c_2)$ and $(c_3, c_4]$, P distorts output till all surplus is extracted. This further clarifies why, over these cost ranges, participation constraints are saturated. For types with more intense incentives to mis-report, namely those in $[\underline{c}, c_1]$ and $[c_4, \overline{c}]$, P distorts output to contain the rent, but it would be too costly to remove the rent entirely.

The whole SB output profile and the thresholds of the relevant cost ranges will be characterized in a moment. Before proceeding, it is however useful to make the following standard assumption.

Assumption 5 *The conditional density and cumulative distribution function satisfy the following properties:*

$$\frac{d}{dc}\left(\frac{F(c\,|n)}{f(c\,|n)}\right) \geq 0, \ \forall c \in [\underline{c}, \overline{c}] \tag{15}$$

$$\frac{d}{dc}\left(\frac{1-F(c\,|1)}{f(c\,|1)}\right) \leq 0, \ \forall c \in [\underline{c}, \overline{c}]. \tag{16}$$

This assumption states the monotonicity of the conditional hazard rates $\frac{F(c|n)}{f(c|n)}$ and $\frac{1-F(c|1)}{f(c|1)}$ with respect to c. According to (15), once types between $\underline{c}$ and c have been drawn, it becomes more likely that a type higher than c be drawn, conditional on signal n being observed. According to (16), once types between c and $\overline{c}$ have been drawn, it is less likely that a type higher than c be drawn, conditional on signal 1 being observed.

In the following proposition, roman numbers are appended to denote SB quantities and payoffs over the five relevant cost ranges.

Proposition 3 *Suppose that* $K''(\cdot) \geq 0$ *and that condition* (10) *does not hold. Then, under*

Assumption $1-3$ *and* 5, *at the solution to* (Γ), *quantities are characterized as follows:*

$$S'\left(q^{I}(c)\right) = c+(1-\alpha)\frac{F(c\,|n)}{f(c\,|n)},\ \forall c\in[\underline{c},c_1] \tag{17}$$

$$q^{II}(c) = \frac{dp(c,n)/dc}{p(c,n)}L-K'(c),\ \forall c\in[c_1,c_2] \tag{18}$$

$$q^{III}(c) = q^{fb}(c),\ \forall c\in[c_2,c_3] \tag{19}$$

$$q^{IV}(c) = \frac{dp(c,1)/dc}{p(c,1)}L-K'(c),\ \forall c\in[c_3,c_4] \tag{20}$$

$$S'\left(q^{V}(c)\right) = c-(1-\alpha)\frac{1-F(c\,|1)}{f(c\,|1)},\ \forall c\in[c_4,\overline{c}], \tag{21}$$

with c_1, c_2, c_3 *and* c_4 *as defined in* $(27)-(30)$ *below. Moreover, interim profits (rents) are given by*

$$E_s\left[\pi^{I}_{c,s}\right]=p(c,n)\int_{c}^{c_1}\frac{q^{I}(x)+K'(x)}{p(x,n)}dx-\left[1-\frac{p(c,n)}{p(c_1,n)}\right]L,\ \forall c\in[\underline{c},c_1] \tag{22}$$

$$E_s\left[\pi^{k}_{c,s}\right]=0,\ \forall c\in[c_1,c_2],[c_2,c_3],[c_3,c_4],\ \forall k\in\{II,III,IV\} \tag{23}$$

$$E_s\left[\pi^{V}_{c,s}\right]=p(c,1)\int_{c_4}^{c}\frac{q^{V}(x)+K'(x)}{-p(x,1)}dx-\left[1-\frac{p(c,1)}{p(c_4,1)}\right]L,\ \forall c\in[c_4,\overline{c}]. \tag{24}$$

This solution satisfies (GIC) *under the conditions:*

$$\frac{dq^{I}(c)}{dc} \leq -\left[q^{I}(c)+K'(c)\right]\frac{dp(c,n)/dc}{p(c,n)},\ \forall c\in[\underline{c},c_1] \tag{25}$$

$$\frac{dq^{V}(c)}{dc} \leq -\left[q^{V}(c)+K'(c)\right]\frac{dp(c,1)/dc}{p(c,1)},\ \forall c\in[c_4,\overline{c}]. \tag{26}$$

Condition (19) confirms that output is still efficiently set as long as $c\in[c_2,c_3]$. According to (17) and (21), the same occurs at both the lowest and the highest marginal cost realization. Condition (17) further highlights that output is downward distorted for all types in $(\underline{c},c_1]$, which allows P to contain the rent in (22). Moreover, under the first part of Assumption 5, q^I decreases with c all over this set.[12] Condition (21) further evidences that output is upward distorted for all types in $[c_4,\overline{c})$, which helps P limit the rent in (24). Under the second part of Assumption 5, also q^V decreases with type for all $c\in[c_4,\overline{c})$. Lastly, (18) and (20) define how output is downward and upward distorted in the second and fourth region respectively, just enough to fully extract surplus in an incentive compatible way.

We now define the thresholds of the relevant cost ranges, which we have only mentioned in Proposition 3 but not yet characterized. The cost values c_1, c_2, c_3 and c_4 are defined as

[12] In Gary-Bobo and Spiegel [7], the SB quantity solution is characterized precisely as in (17) for all possible types because the agent displays a systematic incentive to overstate type.

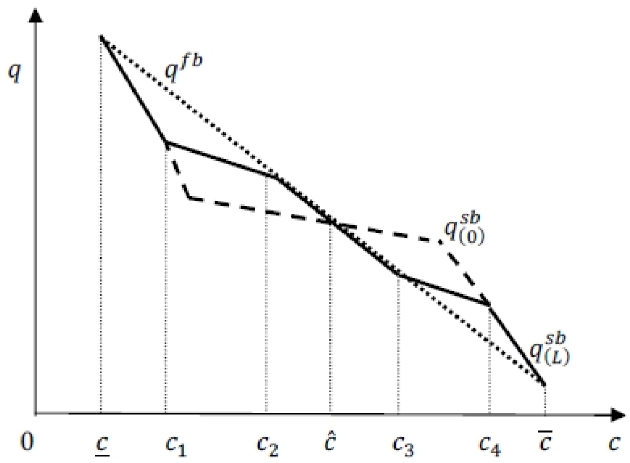

Figure 1: The FB output profile (q^{fb}; dotted line) and the output profile in the SB contract with $L > 0$ ($q^{sb}_{(L)}$; thick line) and with $L = 0$ ($q^{sb}_{(0)}$; dashed line) when $K''(\cdot) > 0$.

follows:

$$q^{I}(c_1) + K'(c_1) = \frac{dp(c_1, n)/dc_1}{p(c_1, n)} L \quad (27)$$

$$q^{II}(c_2) = q^{fb}(c_2) \quad (28)$$

$$q^{IV}(c_3) = q^{fb}(c_3) \quad (29)$$

$$q^{V}(c_4) + K'(c_4) = \frac{dp(c_4, 1)/dc_4}{p(c_4, 1)} L. \quad (30)$$

Interpreting (27) - (30) together with the results previously presented, it should be clear that c_1 is the cost value at which P retains all surplus from the agent by sufficiently deflating output q^I below the FB level, c_2 is the value at which P retains all surplus by keeping output q^{II} at the FB level and similarly for c_3 and c_4.

A graphical illustration of the full profile of quantities, as defined by (17) - (21), is provided in Figure 1 for the case of convex fixed cost. The graph evidences that the set of cost values around $\widehat{c}$ for which FB is still enforced under tight limited liability, enlarges as L raises. On the opposite, it would collapse onto the singleton $\{\widehat{c}\}$ in the extreme case in which $L = 0$.[13] The graph further shows that the SB quantity decreases with c all over the support *i.e.*, $\frac{dq^k(c)}{dc} \leq 0 \; \forall k \in \{I, II, III, IV, V\}, \forall c \in [\underline{c}, \overline{c}]$, with a rate of decrease that is specific to each cost interval. In particular, it is $\frac{dq^I(c)}{dc} < \frac{dq^{fb}(c)}{dc} < \frac{dq^{II}(c)}{dc}$ and $\frac{dq^V(c)}{dc} < \frac{dq^{fb}(c)}{dc} < \frac{dq^{IV}(c)}{dc}$.

Let us now discuss the conditions under which the SB solution described so far is globally incentive compatible. We have previously explained that, under Assumption 5, quantities q^I and q^V decrease with type. Conditions (25) and (26) further evidence that, for the contract presented in Proposition 3 to be globally incentive compatible, it suffices that those quantities decrease sufficiently fast over the respective cost ranges (see Figure 1 again).

[13]Actually, this is also the case when $K''(\cdot) = 0$, whatever the magnitude of L. See the explanation that follows Corollary 1 below.

According to (25) and (26), how fast q^I and q^V should decrease depends upon the rate of change of the conditional likelihood that is relevant in the concerned region. To illustrate why this is the case, let us focus on (25), keeping in mind that analogous reasoning applies to (26), *mutatis mutandis.* Take $c \in [\underline{c}, c_1)$. As the report r is raised above the true type c, under Assumption 2, the probability of reward increases. Because the loss that the agent might bear equals L, whatever the statement $r \in [\underline{c}, c_1]$, over-reporting yields a higher *interim* profit, as compared to truthtelling, unless the quantity is diminished sufficiently. The incentive to over-report is removed if $q^I(c)$ decreases as fast as (25) dictates.[14]

Corollary 1 *Suppose that condition* (10) *does not hold, whereas* (25) *and* (26) *are satisfied. When* $K"(\cdot) = 0$ *and* $L = 0$, *at the solution to* (Γ), $q^{sb}(c) = q^{fb}(\widehat{c})$ *for all* $c \in [c_1, c_4]$.

This corollary refers to the specific situation in which the fixed cost is linear in type and the agent can bear no deficit *ex post.* When $K''(\cdot) = 0$, the range of types for which FB is enforced collapses onto the singleton $\{\widehat{c}\}$. To see this, recall that c_2 and c_3 are defined by $q^{II}(c_2) = q^{fb}(c_2)$ and $q^{II}(c_3) = q^{fb}(c_3)$, respectively. When, additionally, $L = 0$, we further have $q^{II}(c) = -K'(c)$ and $q^{IV}(c) = -K'(c)$. Remembering also the definition of $\widehat{c}$, it is immediate to conclude that $c_2 \equiv \widehat{c} \equiv c_3$ when $L = 0$. Further observe that quantities $q^{II}(c)$ and $q^{IV}(c)$ are constant over types when so is $K'(c)$. Hence, all types within the set $[c_1, c_4]$, from which surplus is entirely extracted, are required to produce the same amount of output *i.e.*, the optimal contract entails pooling at $q^{fb}(\widehat{c})$ in a neighborhood of $\widehat{c}$. The output profile for this case is represented in Figure 2.

The outcome presented in Corollary 1 is reminiscent of that Maggi and Rodriguez-Clare [11] find in a setting without informative signals. They characterize the optimal contract in the presence of countervailing incentives for different possible shapes of the agent's reservation utility. They show that, when the reservation utility is linear in type, the contract entails pooling of quantities over some interval of types that earn zero rents. The case of $K"(\cdot) = 0$ in our model is the counterpart for the case of linear reservation utility in Maggi and Rodriguez-Clare [11]. Corollary 1 evidences that, when $K"(\cdot) = 0$, the optimal contract exhibits analogous features (namely, pooling and no rent in a neighborhood of $\widehat{c}$) in a correlated-information framework as soon as the agent cannot be punished *ex post.* This is explained by considering that having $L = 0$ in the presence of correlated information is tantamount to assuming that the agent has to break even *ex post* (rather than at *interim*), whereas *ex post* and *interim* participation are equivalent without correlated information. Nevertheless, the optimal contract does not simply replicate the one that would be offered in the absence of informative signals because, as usual, correlation enhances contracting.

[14] The sufficient condition for global incentive compatibility in Gary-Bobo and Spiegel [7] would be analogous to (25) if, in their model, the marginal cost were assumed to be constant in type, as it is in ours, rather than strictly increasing and convex. This can be seen by comparing the inequality at the end of page 5 in the technical appendix of Gary-Bobo and Spiegel [7] with (75) in the proof of (25) in our Appendix.

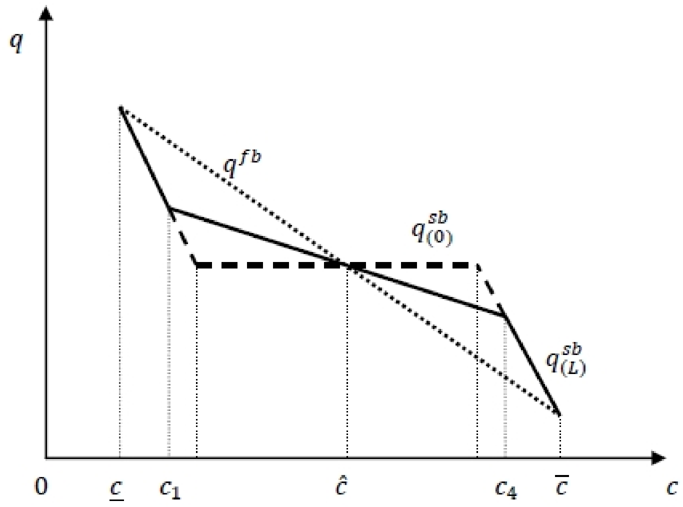

Figure 2: The FB output profile (q^{fb}; dotted line) and the output profile in the SB contract with $L > 0$ ($q^{sb}_{(L)}$; thick line) and with $L = 0$ ($q^{sb}_{(0)}$; dashed line) when $K''(\cdot) = 0$.

Corollary 2 *Suppose that condition* (10) *does not hold, whereas* (25) *and* (26) *are satisfied. Further suppose that* $K"(\cdot) = 0$ *and* $L = 0$. *Then, at the solution to* (Γ), *quantities are pooled for a smaller range of types, as compared to situations in which informative signals are not available. Moreover, non-pooled types are assigned* (i) *less distorted quantities and* (ii) *lower rents.*

3.2 Concave fixed cost

When $K''(c) < 0$, the incentive problem is exacerbated to the point that an incentive scheme under which all types obtain a reward for one sole signal realization and bear losses otherwise, no longer allows P to implement FB. As long as limited liability is not an issue, P would need to design an alternative mechanism, presumably less parsimonious in terms of signals to be used, to effect FB.

However, when (LL) is so stringent that FB is beyond reach, a one-reward scheme comes back to be useful. From the FB analysis, recall, indeed, that spreading losses over as many signal realizations as possible and raising them to the highest admissible level, allows P to maximize the uncertainty that the agent is exposed to, thereby inducing truthtelling at a lower cost. Thus, obviously, when $K''(c) < 0$, the SB scheme is still such that, for all $c \neq \widehat{c}$, losses are spread over all signal realizations but one and, with (LL) binding, raised to L.

We shall present the SB scheme in further details in a moment. Prior to that, we need to define:

$$h \equiv \arg\min_{\widetilde{s}(c)\in N} \frac{dp\left(c,\widetilde{s}\left(c\right)\right)/dc}{p\left(c,\widetilde{s}\left(c\right)\right)} \text{ and such that } \frac{dp\left(c,h\right)}{dc} > 0 \; \forall c \in [\underline{c},\overline{c}] \tag{31}$$

$$k \equiv \arg\max_{\widetilde{s}(c)\in N} \frac{dp\left(c,\widetilde{s}\left(c\right)\right)/dc}{p\left(c,\widetilde{s}\left(c\right)\right)} \text{ and such that } \frac{dp\left(c,k\right)}{dc} < 0 \; \forall c \in [\underline{c},\overline{c}] \,. \tag{32}$$

Assumption 2 and 3 ensure that the pair of signals h and k does exist, provided signals n and 1 are candidate to display the characteristics in (31) and (32), respectively. Moreover, Assumption 5 warrants that the conditions

$$\frac{d}{dc}\left(\frac{F(c|h)}{f(c|h)}\right) \geq 0, \ \forall c \in [\underline{c}, \overline{c}], \tag{33}$$

$$\frac{d}{dc}\left(\frac{1-F(c|k)}{f(c|k)}\right) \leq 0, \ \forall c \in [\underline{c}, \overline{c}], \tag{34}$$

are satisfied at least for $h = n$ and $k = 1$, respectively. We can now describe the optimal contract.

Proposition 4 *Suppose that $K''(\cdot) < 0$ and that condition (10) does not hold. Then, under Assumption $1-3$ and 5 and condition (33) and (34), the quantity solution to (Γ) is given by*

$$S'\left(q^{sb}(c)\right) = c + (1-\alpha)\frac{F(c|h)}{f(c|h)}, \ \forall c \in \left[\underline{c}, c^{-}\right) \tag{35}$$

$$q^{sb}(c) = q^{fb}(\widehat{c}) + \int_{c}^{\widehat{c}} \left[q^{sb}(x) + K'(x)\right] \frac{dp(x,h)/dx}{p(x,h)} dx, \ \forall c \in \left[c^{-}, \widehat{c}\right) \tag{36}$$

$$q^{sb}(\widehat{c}) = q^{fb}(\widehat{c}) \tag{37}$$

$$q^{sb}(c) = q^{fb}(\widehat{c}) - \int_{\widehat{c}}^{c} \left[q^{sb}(x) + K'(x)\right] \frac{dp(x,k)/dx}{p(x,k)} dx, \ \forall c \in \left(\widehat{c}, c^{+}\right] \tag{38}$$

$$S'\left(q^{sb}(c)\right) = c - (1-\alpha)\frac{1-F(c|k)}{f(c|k)}, \ \forall c \in \left(c^{+}, \overline{c}\right], \tag{39}$$

where c^{-} and c^{+} are the smallest and the largest cost value for which, respectively, the global incentive compatibility conditions here below are binding:

$$\frac{dq^{sb}(c)}{dc} \leq -\left[q^{sb}(c) + K'(c)\right]\frac{dp(c,h)/dc}{p(c,h)}, \ \forall c \in [\underline{c}, \widehat{c}) \tag{40}$$

$$\frac{dq^{sb}(c)}{dc} \leq -\left[q^{sb}(c) + K'(c)\right]\frac{dp(c,k)/dc}{p(c,k)}, \ \forall c \in (\widehat{c}, \overline{c}]. \tag{41}$$

Type $\widehat{c}$ obtains a payoff equal to zero for all $s \in N$. Types $c \in [\underline{c}, \widehat{c})$ obtain the rent in (22), where h and $\widehat{c}$ replace n and c_1, respectively. Types $c \in (\widehat{c}, \overline{c}]$ obtain the rent in (24), where k and $\widehat{c}$ replace 1 and c_4, respectively.

Conditions (35), (37) and (39) show that output is still efficiently set only for types $\underline{c}$, $\widehat{c}$ and $\overline{c}$. Conditions (35) and (39) further evidence that, for low and high types, the output solution is found to be like in (17) and (21), except that reward signals are now h and k. The solution in (35) no longer applies as soon as (40) becomes binding, at the cost value c^{-}. Thereafter, with the usual exception of $\widehat{c}$, P is unable to remove the temptation to cheat, unless she fixes the quantity according to (36) (*i.e.*, at the exact level that (40) dictates) for all $c \in [c^{-}, \widehat{c})$, and according to (38) (*i.e.*, at the level that exactly complies with (41)) for all $c \in (\widehat{c}, c^{+}]$. Beyond the cost value c^{+}, (41) is no longer binding and the solution in (39) applies. The difficulties that P faces in the attempt to elicit information when $K(\cdot)$ is concave, especially from types around $\widehat{c}$ that have incentives to report "closer" to the true cost value, explain why $\widehat{c}$ is now the sole type from which P retains surplus, despite that, unlike $\widehat{c}$, all other types are inflicted the maximum sustainable loss (L).

It is important to understand how P takes advantage of the informative signals to improve contractual performance in this environment. Observe that, *ceteris paribus*, condition (40) (resp. (41)) is relaxed as the ratio $\frac{dp(x,\cdot)/dx}{p(x,\cdot)}$ is decreased (resp. increased). Of course, under this circumstance, P prefers to reward the agent when she observes a signal at which this ratio is small (resp. large). More precisely, the best strategy for P is to pick the signal h (resp. k) such that the ratio $\frac{dp(x,\cdot)/dx}{p(x,\cdot)}$ is minimized (resp. maximized), which explains (31) (resp. (32)). The value of c^{-} (resp. c^{+}) is determined accordingly. Therefore, even in the event that $h = n$ and $k = 1$ so that the solution in (35) and (39) coincides with that in (17) and (21), respectively, the subsets of types $[\underline{c}, c^{-})$ and $(c^{+}, \overline{c}]$ do not (need to) coincide with the subsets $[\underline{c}, c_1)$ and $(c_4, \overline{c}]$, respectively.

By linking the agent's compensation to the signal realization, P is able to limit quantity distortions even for types whose temptation to lie is most hardly removed. In particular, this means that correlated information rules out the need to pool output in a neighborhood of $\widehat{c}$, that would arise otherwise. Indeed, in a framework where the fixed cost is concave but informative signals are not available, Lewis and Sappington [9] find that quantities are set equal to $q^{fb}(\widehat{c})$ for the whole bunch of types around $\widehat{c}$ for which the incentive compatibility condition is tightest. That solution appears as a limit case of the optimal contract presented in Proposition 4.

Corollary 3 *Suppose that $K''(\cdot) < 0$ and that condition* (10) *does not hold. Then, under Assumption* $1 - 3$ *and conditions* (33) *and* (34), *all types $c \in [c^{-}, c^{+}]$ produce the quantity $q^{sb}(c) = q^{fb}(\widehat{c})$ if and only if informative signals are not available.*

The output profiles with and without informative signals for the case of $K(\cdot)$ concave are graphically represented in Figure 3.

4 Comparison with no-correlated-information settings

Maggi and Rodriguez-Clare [11] show that, in environments where correlated signals are not available, the linear-fixed-cost case (or, equivalently, the linear-reservation-utility case), in which quantities are bunched for certain types that earn zero rents, can be viewed as a "knife-edge" situation. Pooling is removed as soon as the fixed cost becomes convex. All types but one obtain a rent as soon as the fixed cost becomes concave.

Our second-best analysis unveils that, in frameworks where the payment of the agent can be conditioned on the realization of a correlated signal, in the linear-fixed-cost case, the

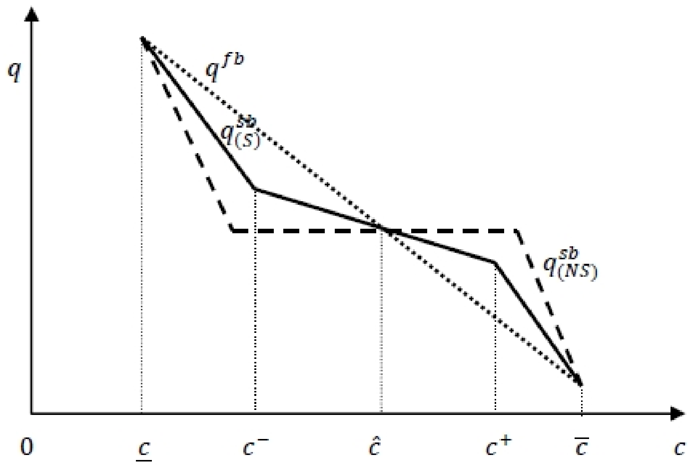

Figure 3: The FB output profile (q^{fb}; dotted line) and the output profile in the SB contract with informative signals ($q^{sb}_{(S)}$; thick line) and without informative signals ($q^{sb}_{(NS)}$; dashed line) when $K''(\cdot) < 0$.

principal avoids to give up rents to some types without the need to require that the latter produce all the same output, at least as long as the agent can be exposed to some deficit *ex post*, though his pocket is not very deep. The linear case still represents a "knife-edge" situation, but in a different way. As soon as the fixed cost becomes convex, the first-best outcome is at hand for some of the types from which surplus can be entirely extracted. As soon as the fixed cost becomes concave, all types but one obtain a rent, all of them being assigned different quantities.

Overall, while in uncorrelated-information settings pooling arises whenever the fixed cost is non-convex in type, our results show that this is no longer the case in correlated-information frameworks. In the latter, pooling is solely induced in the specific case in which the fixed cost is linear and the agent must break even whatever the signal realization. Yet, even in that case, pooling does not follow from the impossibility to find a fully separating quantity profile for which the contract results globally incentive compatible, as it occurs in the absence of informative signals. Rather, it serves the purpose of securing the agent's participation in the relationship with the principal at zero rent.

The situation that we refer to here above belongs to the class of situations in which not only the benefit associated with the agent's liability appears in the form of a rent reduction. It is also directly visible in both the output distortions and the cost values that identify the intervals over which the various solutions apply. Such situations only arise when the principal is able to use only the agent's compensation to solve the incentive compatibility problem. Then, quantity adjustments can be targeted to secure that the agent participates in the contract grasping no surplus. This task becomes easier as the deficit that the agent can bear raises, because correlation can be better exploited. The principal can thus afford to keep the agent in the contract by inducing a smaller quantity distortion for a wider range of types. For sufficiently high liability, the distortion even disappears for some cost values and the first-best outcome entails. Our study highlights that this additional benefit solely appears with (weakly) convex fixed cost (recall (18), (20) and (27) to (30)). It does not, on the opposite, when the fixed cost is concave, in which case the principal needs to use not only the compensation but also the quantity of the agent to handle with the incentive problem. It is thus explained why, unlike in Figure 1 and 2, the output profile with informative signals represented in Figure 3 is not differentiated according to whether $L > 0$ or $L = 0$.

Acknowledgements

We gratefully acknowledge useful comments from participants at the first NERI Meeting (Pescara), the XXI Siep Meeting (Pavia), the 2009 ASSET Meeting (Istanbul), the 2010 EEA Meeting (Glasgow) as well as seminar participants at the University of St Andrews, the Max-Planck Institute for Research on Collective Goods (Bonn) and the SIRE FORUM (Edinburgh). The usual disclaimer applies.

References

[1] Acconcia, A., R. Martina and S. Piccolo (2008), "Vertical Restraints under Asymmetric Information: On the Role of Participation Constraints," *Journal of Industrial Economics*, 56(2), 379-401

[2] Armstrong, M., and D.E.M. Sappington (2007), "Recent Developments in the Theory if Regulation," *Handbook of Industrial Organization*, Volume 3, Chapter 27, Ed. M. Armstrong and R. Porter, Elsevier

[3] Brainard, S.L., and D. Martimort (1996), "Strategic Trade Policy Design with Asymmetric Information and Public Contracts," *The Review of Economic Studies*, 63(1), 81-105

[4] Crémer, J., and R. McLean (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56(6), 1247-1258

[5] Degryse, H., and A. de Jong (2006), "Investment and Internal Finance: Asymmetric Information or Managerial Discretion", *International Journal of Industrial Organization*, 24, 125-147

[6] Demougin, D.M., and D.A. Garvie (1991), "Contractual design with correlated information under limited liability," *The RAND Journal of Economics*, 22(4), 477-489

[7] Gary-Bobo, R., and Y. Spiegel (2006), "Optimal State-Contingent Regulation under Limited Liability," *RAND Journal of Economics*, 37, 431-448

[8] Jullien, B. (2000), "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93, 1-47

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

[10] Lewis, T.R., and D.E.M. Sappington (1989b), "Inflexible Rules in Incentive Problems," *American Economic Review*, 79(1), 69-84

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

[12] Poudou, J.C., M. Roland and L. Thomas (2009), "Universal Service Obligations and Competition with Asymmetric Information," *The B.E. Journal of Theoretical Economics*, 9(1) (Topics), Article 35

[13] Riordan, M., and D.E.M. Sappington (1988), "Optimal Contracts with Public Ex Post Information," *Journal of Economic Theory*, 45, 189-199

[14] Vinella, A. (2010), "Bayesian-Nash *vs* dominant-strategy implementation with countervailing incentives: the two-type case," *Economics Bulletin*, 30(1), 636-649

[15] Vogelsang, I. (1999), "Optimal Price Regulation for Natural and Legal Monopolies," *Economía Mexicana. Nueva Época*, VIII(1), 5-43

[16] Vogelsang, I. (2002), "Incentive Regulation and Competition in Public Utility Markets: A 20-Year Perspective," *Journal of Regulatory Economics*, 22(1), 5-27

A First-best implementation

A.1 Local incentive constraint (LIC)

Let $\widetilde{\pi}(r,s) = t(r,s) - cq(r) - K(c)$ the *ex post* profit of the agent when he has type c and reports r while the state of nature is s. His *interim* profit is written

$$E_s[\widetilde{\pi}(r,s)] \equiv \sum_{s=1}^{n} \{t(r,s) - cq(r) - K(c)\} p(c,s). \tag{42}$$

From (42), the first order-condition of the programme of the agent, evaluated at $r = c$, is given by

$$\sum_{s=1}^{n} \left[\frac{dt(c,s)}{dc} - cq'(c)\right] p(c,s) = 0. \tag{43}$$

From (2) we can compute

$$\frac{dt(c,s)}{dc} = \frac{d\pi(c,s)}{dc} + cq'(c) + q(c) + K'(c). \tag{44}$$

Replacing (44) into (43), we have

$$\sum_{s=1}^{n} \left\{\frac{d\pi(c,s)}{dc} + [q(c) + K'(c)]\right\} p(c,s) = 0. \tag{45}$$

At FB $\pi(c,s) = \pi^{fb}(c,s)$ and (PC) is binding for all c, so that

$$\sum_{s=1}^{n} \pi^{fb}(c,s) p(c,s) = 0 \tag{46}$$

and

$$\sum_{s=1}^{n} \frac{d\pi^{fb}(c,s)}{dc} p(c,s) = -\sum_{s=1}^{n} \pi^{fb}(c,s) \frac{dp(c,s)}{dc}.$$

Using this in (45) together with $q(c) = q^{fb}(c)$, (LIC) is obtained.

A.2 Proof of Lemma 2

Suppose $\widetilde{s}(c) = \overline{s}$, $\forall c \in [\underline{c}, \overline{c}]$, with $\overline{s}$ some constant in N. Take also $\frac{dp(c,\overline{s})}{dc} > 0$. For $c < \widehat{c}$, the punishment is as from (5) *i.e.*, $\underline{\pi}^{fb}(c, \overline{s}) < 0$. Similarly, for $c > \widehat{c}$, the punishment is as from (4) *i.e.*, $\overline{\pi}^{fb}(c, \overline{s}) < 0$. Furthermore, because these profits belong to the MFL, $\overline{s}$ must maximize both $\underline{\pi}^{fb}(c, \overline{s})$ for $c < \widehat{c}$ and $\overline{\pi}^{fb}(c, \overline{s})$ for $c > \widehat{c}$ at once. The former requires that $\frac{d}{dc}\left(\frac{p(c,s)}{p(c,\overline{s})}\right) < 0$, $\forall s \neq \overline{s}$, the latter that $\frac{d}{dc}\left(\frac{1-p(c,s)}{p(c,\overline{s})}\right) > 0$, $\forall s \neq \overline{s}$. Suppose $\frac{d}{dc}\left(\frac{p(c,s)}{p(c,\overline{s})}\right) < 0$. Together with $\frac{dp(c,\overline{s})}{dc} > 0$, this involves that $\frac{d}{dc}\left(\frac{1-p(c,s)}{p(c,\overline{s})}\right) \leq 0$, contradicting the hypothesis that the MFL is obtained with $\widetilde{s}(c) = \overline{s}$. The proof proceeds similarly for $\frac{dp(c,\overline{s})}{dc} < 0$.

A.3 Proof of Proposition 1

Recall that payoffs (4) and (5) are such that (PC) and (LIC) are satisfied. Hence, a solution to (Γ) can be obtained with payoffs (4) and (5) if and only if they satisfy (LL) and (GIC).

A.3.1 Limited liability (LL)

From (4) and (5), with $\widetilde{s}(r) = n$ for $r \leq \widehat{c}$ and $\widetilde{s}(r) = 1$ for $r > \widehat{c}$, and from Assumptions 1 - 3, we deduce that $\underline{\pi}^{fb}(c, \widetilde{s}(r)) < 0 < \overline{\pi}^{fb}(c, \widetilde{s}(c))$ $\forall r \in [\underline{c}, \overline{c}]$. Hence, (LL) becomes $\underline{\pi}^{fb}(c, \widetilde{s}(r)) \geq -L$. Using (5), this is rewritten as (10).

A.3.2 Global incentive compatibility (GIC)

We rewrite $E_s[\widetilde{\pi}(r,s)]$ in (42) as follows:

$$
\begin{aligned}
E_s[\widetilde{\pi}(r,s)] &\equiv q^{fb}(r)(r-c) + K(r) - K(c) + \sum_{s=1}^{n} \pi(r,s)\, p(c,s) \qquad (47)\\
&= q^{fb}(r)(r-c) + K(r) - K(c) + \underline{\pi}^{fb}(c, \widetilde{s}(r)) \\
&\quad + \left[\overline{\pi}^{fb}(r, \widetilde{s}(c)) - \underline{\pi}^{fb}(r, \widetilde{s}(c))\right] p(c, \widetilde{s}(r))
\end{aligned}
$$

Replacing (4) and (5) into (47), the latter becomes

$$
\begin{aligned}
E_s[\widetilde{\pi}(r,s)] &= \int_c^r \left\{ \left[q^{fb}(r) + K'(r)\right] \left[1 - \frac{dp(x, \widetilde{s}(r))/dx}{dp(r, \widetilde{s}(r))/dr}\right] \right. \qquad (48)\\
&\quad \left. + K'(x) - K'(r) \right\} dx.
\end{aligned}
$$

(GIC) requires that $E_s[\widetilde{\pi}(r,s)] \leq 0$. Using (48), this is tantamount to having

$$
K'(r) - K'(c) \geq \left[q^{fb}(r) + K'(r)\right] \left[1 - \frac{dp(c, \widetilde{s}(r))/dc}{dp(r, \widetilde{s}(r))/dr}\right] \text{ if } r \geq c \qquad (49)
$$

$$
K'(r) - K'(c) \leq \left[q^{fb}(r) + K'(r)\right] \left[1 - \frac{dp(c, \widetilde{s}(r))/dc}{dp(r, \widetilde{s}(r))/dr}\right] \text{ if } r \leq c. \qquad (50)
$$

Recall that $\widetilde{s}(r) = n$ if $r < \widehat{c}$ and $\widetilde{s}(r) = 1$ if $r > \widehat{c}$. Moreover, K" has constant sign for all c. It follows that both (49) and (50) are satisfied if and only if (11) holds.

Condition (12) follows immediately from (49) and (50), with $q^{fb}(r) + K'(r) \neq 0$ $\forall r \in [\underline{c}, \overline{c}]$.

A.4 Proof of Lemma 3

We first show that, with (PC) binding, the expected value of the lottery $\sum_{s=1}^{n}\pi(r,s)\,p(c,s)$ is minimized when P assigns losses for all signal realizations but one, the loss being fixed at the highest feasible amount. We then identify the highest feasible loss $M \leq L$ for any type c and the corresponding reward at which (LIC), (PC) and (LL) are all satisfied. We shall refer to this loss as to the highest "compatible" loss.

A.4.1 The expected value of the lottery is minimized at the highest compatible loss

Denote M the highest compatible loss, such that (LL) is satisfied, namely $M \leq L$. We proceed as follows. For any given c, we first calculate the expected value of the lottery when the scheme includes one reward and losses all equal to M. We then calculate the expected value of the lottery with three distinct payoff levels, the smallest of which equal to $-M$. We lastly compare the expected value of the lottery in the two cases and show that it is higher in the latter case.

As a first step, assume that, when the agent has type c and reports r, he receives $\overline{\pi}(r,\widetilde{s}(r))>0$ if the state is some $s=\widetilde{s}(r)$ and $\underline{\pi}(r,\widetilde{s}(r))=-M$ in any state $s\neq\widetilde{s}(r)$. P seeks to minimize

$$\sum_{s=1}^{n}\pi(r,s)\,p(c,s)=\underline{\pi}(r,\widetilde{s}(r))+\left[\overline{\pi}(r,\widetilde{s}(r))-\underline{\pi}(r,\widetilde{s}(r))\right]p(c,\widetilde{s}(r)). \tag{51}$$

With (PC) binding for type r, we have

$$\overline{\pi}(r,\widetilde{s}(r))=-\frac{1-p(r,\widetilde{s}(r))}{p(r,\widetilde{s}(r))}\underline{\pi}(r,\widetilde{s}(r)).$$

Replacing this expression into (51) together with $\underline{\pi}(r,\widetilde{s}(r))=-M$, we get

$$\sum_{s=1}^{n}\pi(r,s)\,p(c,s)=-M\left[1-\frac{p(c,\widetilde{s}(r))}{p(r,\widetilde{s}(r))}\right]. \tag{52}$$

Assume next that P implements FB with three distinct payoff levels, namely $\underline{\pi}(c,\widetilde{s}(c))$, $\widehat{\pi}(c,\widehat{s}(c))$ and $\overline{\pi}(c,\widetilde{s}(c))$, such that $\underline{\pi}(c,\widetilde{s}(c))=-M$, $\underline{\pi}(c,\widetilde{s}(c))<\widehat{\pi}(c,\widehat{s}(c))<\overline{\pi}(c,\widetilde{s}(c))$ and $\widehat{s}(c)\in N\setminus\{\widetilde{s}(c)\}$. The expected value of the lottery becomes

$$\begin{aligned}\sum_{s=1}^{n}\pi(r,s)\,p(c,s) &= \left[\overline{\pi}(r,\widetilde{s}(r))-\underline{\pi}(r,\widetilde{s}(r))\right]p(c,\widetilde{s}(r)) \\ &\quad +\underline{\pi}(r,\widetilde{s}(r))+\left[\widehat{\pi}(c,\widehat{s}(c))-\underline{\pi}(r,\widetilde{s}(r))\right]p(c,\widehat{s}(r)),\end{aligned} \tag{53}$$

whereas the binding (PC) is now written

$$\overline{\pi}(r,\widetilde{s}(r))-\underline{\pi}(r,\widetilde{s}(r))=-\frac{\underline{\pi}(r,\widetilde{s}(r))+\left[\widehat{\pi}(r,\widehat{s}(r))-\underline{\pi}(r,\widetilde{s}(r))\right]p(r,\widehat{s}(r))}{p(r,\widetilde{s}(r))}.$$

Replacing this expression into (53) , together with $\underline{\pi}(r, \widetilde{s}(r)) = -M$, we obtain

$$\sum_{s=1}^{n} \pi(r,s)\, p(c,s) = -M \left[1 - \frac{p(c,\widetilde{s}(r))}{p(r,\widetilde{s}(r))}\right] + \left[\widehat{\pi}(r,\widehat{s}(r)) + M\right] \left[1 - \frac{p(r,\widehat{s}(r))\, p(c,\widetilde{s}(r))}{p(c,\widehat{s}(r))\, p(r,\widetilde{s}(r))}\right] p(c,\widehat{s}(r)). \quad (54)$$

Calculating the difference between (52) and (54), we obtain

$$\left[\widehat{\pi}(r,\widehat{s}(r)) + M\right] \left[-\frac{p(r,\widehat{s}(r))\, p(c,\widetilde{s}(r))}{p(c,\widehat{s}(r))\, p(r,\widetilde{s}(r))} + 1\right] p(c,\widehat{s}(r)).$$

From Proposition 2, $\widetilde{s}(r) = i$ if $r < \widehat{c}$ and $\widetilde{s}(r) = j$ if $r > \widehat{c}$. Under Assumption 4 and because $\widehat{\pi}(r,\widehat{s}(r)) + M > 0$, the above difference is positive. Hence, the expected value of the lottery is higher with any triplet $\{\underline{\pi}(c,\widetilde{s}(c)), \widehat{\pi}(c,\widehat{s}(c)), \overline{\pi}(c,\widetilde{s}(c))\}$, such that $\underline{\pi}(c,\widetilde{s}(c)) = -M$ and $\underline{\pi}(c,\widetilde{s}(c)) < \widehat{\pi}(c,\widehat{s}(c)) < \overline{\pi}(c,\widetilde{s}(c))$, than it is with the pair of profits $\{\underline{\pi}(c,\widetilde{s}(c)), \overline{\pi}(c,\widetilde{s}(c))\}$.

A.4.2 The value of M

The loss and the reward at which the expected value of the lottery is minimized are written as follows:

$$\begin{aligned} \underline{\pi}(c,\widetilde{s}(c), M) &= -M = \text{ punishment of type } c \text{ in states } s \neq \widetilde{s}(c) \\ \overline{\pi}(c,\widetilde{s}(c), M) &= \frac{1 - p(c,\widetilde{s}(c))}{p(c,\widetilde{s}(c))} M = \text{ reward that type } c \text{ obtains in state } \widetilde{s}(c). \end{aligned}$$

When type c reports r, his expected payoff is computed as

$$\begin{aligned} E_s[\widetilde{\pi}(r,s)] &= q^{fb}(r)(r-c) + K(r) - K(c) + \sum_{s=1}^{n} \underline{\pi}^*(r,\widetilde{s}(r), M)\, p(c,s) \\ &\quad + \left[\overline{\pi}^*(r,\widetilde{s}(r), M) - \underline{\pi}^*(r,\widetilde{s}(r), M)\right] p(c,\widetilde{s}(r)). \end{aligned}$$

Replacing the punishment and the reward reported above, this becomes

$$E_s[\widetilde{\pi}(r,s)] = \int_c^r \left[q^{fb}(r) + K'(r) - M \frac{dp(x,\widetilde{s}(r))/dx}{p(r,\widetilde{s}(r))} + K'(x) - K'(r)\right] dx.$$

(GIC) is satisfied if and only if $E_s[\widetilde{\pi}(r,s)] \leq 0$ *i.e.*,

$$\int_c^r \left[K'(r) - K'(x)\right] dx \geq \int_c^r \left[q^{fb}(r) + K'(r) - M \frac{dp(x,\widetilde{s}(r))/dx}{p(r,\widetilde{s}(r))}\right] dx$$

or, equivalently,

$$\int_c^r \int_x^r K"(y) dy dx \geq \int_c^r \left[q^{fb}(r) + K'(r) - M \frac{dp(x,\widetilde{s}(r))/dx}{p(r,\widetilde{s}(r))}\right] dx.$$

This is further rewritten

$$\int_c^r \int_x^r K"(y)dydx \geq \int_c^r \int_x^r [\frac{q^{fb}(r) + K'(r) - M\frac{dp(r,\widetilde{s}(r))/dr}{p(r,\widetilde{s}(r))}}{r-x} + M\frac{d^2p(y,\widetilde{s}(r))/dy^2}{p(r,\widetilde{s}(r))}]dydx. \tag{55}$$

As $(r-x)$ becomes infinitely small, the right-hand side (RHS hereafter) of this condition remains finite if and only if

$$M = \left[q^{fb}(r) + K'(r)\right] \frac{p(r,\widetilde{s}(r))}{dp(r,\widetilde{s}(r))/dr}.$$

Replacing this value of M into $\underline{\pi}(c,\widetilde{s}(c),M)$ and $\overline{\pi}(c,\widetilde{s}(c),M)$, we obtain (4) and (5). Moreover, replacing M into (55), we come back to (12), except that here $\widetilde{s}(c)$ is not yet defined. Lastly, under Assumption 4, (3) is relaxed most when $\widetilde{s}(c) = i$ for $c < \widehat{c}$ and $\widetilde{s}(c) = j$ for $c > \widehat{c}$, whereas it is independent of $\widetilde{s}(c)$ for $c = \widehat{c}$.

A.5 Proof of Proposition 2

From the proof of Lemma (3), condition (GIC) under the MCL is the same as (12) under the MFL, except that $\widetilde{s}(c)$ takes different values under the two schemes. Under Assumption 2, 3 and 4, switching from the MFL to the MCL, the (GIC) defined by (12) is relaxed for all $c \neq \widehat{c}$. For $c = \widehat{c}$, (12) is the same under the two schemes.

B The SB contract with tight limited liability

B.1 Proof of Lemma 4

From the definition of $\widehat{c}$ in Lemma 2, (10) holds for $c = \widehat{c}$, $\forall L \geq 0$. Take now $c < \widehat{c}$ and suppose that (10) is violated for c, in which case

$$q^{fb}(c) + K'(c) > L\frac{dp(c,n)/dc}{p(c,n)}. \tag{56}$$

(i) Suppose that

$$\frac{dq^{fb}(c)}{dc} + K"(c) < \frac{L}{p(c,n)}\left[\frac{d^2p(c,n)}{dc^2} - \frac{(dp(c,n)/dc)^2}{p(c,n)}\right]. \tag{57}$$

and recall that $\frac{dq^{fb}(c)}{dc} + K"(c) < 0$ (Assumption 1). It follows that, as c raises, the left-hand side (LHS hereafter) of (56) decreases faster, as compared to the RHS. Because (56) does not hold for $c = \widehat{c}$, there is at most one value $c_2 \in [\underline{c}, \widehat{c}]$ such that (56) does not hold for any $c \in [\underline{c}, c_2)$ and holds for all $c \in [c_2, \widehat{c}]$. This value exists if (56) holds for $c = \underline{c}$.

(ii) Next suppose that (57) is not satisfied, so that, as c raises, the LHS of (56) decreases less fast than the RHS. Hence, if (56) does not hold for $c = \underline{c}$, then it does not hold for any $c \in [\underline{c}, \widehat{c}]$, in which case there is no c in this interval for which (10) is violated. If (56) holds for $c = \underline{c}$, then it must hold for any $c \in [\underline{c}, \widehat{c}]$, involving that (10) is violated for all types within this interval. This contradicts the definition of $\widehat{c}$, under which (10) is satisfied for $c = \widehat{c}$. Therefore, (56) does not hold for $c = \underline{c}$, so that (10) is satisfied for all $c \in [\underline{c}, \widehat{c}]$.

Considering (i) and (ii) altogether, we deduce that there exists at most one subset $[\underline{c}, c_2) \subseteq [\underline{c}, \widehat{c}]$ over which (10) is violated, with $c_2 \in [\underline{c}, \widehat{c}]$. This value exists if and only if (10) is violated for $c = \underline{c}$.

A similar reasoning applies when $c > \widehat{c}$, meaning that there exists at most one subset $(c_3, \overline{c}] \subseteq [\widehat{c}, \overline{c}]$, with $c_3 \in [\widehat{c}, \overline{c}]$, for which (10) is violated.

B.2 Proof of Lemma 5

Take $c \in [\underline{c}, c_2)$. By Lemma 4, (LL) is binding, which means that all losses are equal to $-L$. Furthermore, the proof of Lemma 3 shows that, for all $c \in [\underline{c}, \widehat{c}]$ to which a rent accrues with this profile of profits, the SB quantity is given by $q^I(c)$ as defined by (17), whereas, in case no rent, accrues the SB quantity is given by $q^{II}(c)$ as defined by (18). A rent is given up to type $c \in [\underline{c}, \widehat{c}]$ if and only if

$$q^I(c) + K'(c) > L \frac{dp(c,n)/dc}{p(c,n)}. \tag{58}$$

Indeed, for the types for which (58) is violated, P is better off by choosing the quantity $q^{II}(c) \geq q^I(c)$ such that all surplus is extracted.

From Assumption 1 and because $\frac{dq^I(c)}{dc} < \frac{dq^{fb}(c)}{dc}$, it is $K"(c) < -\frac{dq^I(c)}{dc}$. Using this condition, we proceed identically as in the proof of Lemma 4 but replacing $q^{fb}(c)$ with $q^I(c)$, c_2 with c_1, $\widehat{c}$ with c_2 and (56) with (58). We find that there exists at most one cost value in $[\underline{c}, c_2)$, that we denote c_1, for which $q^I(c) + K'(c) = L\frac{dp(c,n)/dc}{p(c,n)}$. This value c_1 exists if and only if (58) is satisfied for $c = \underline{c}$ (and a rent is given up at least to type $\underline{c}$).

The procedure is similar for $c \in [\widehat{c}, \overline{c}]$.

B.3 Proof of Proposition 3

We use Lemma 4 and 5 to identify the intervals in which (LL) is binding and pin down the SB solution that satisfies (LIC), (PC) and (LL). We then find the conditions under which (GIC) is satisfied as well.

B.3.1 The solution

Define

$$\widetilde{W}(a,b) \equiv \int_a^b \sum_{s=1}^{n} \left[V(q(c)) + \alpha\pi(c,s)\right] p(c,s) f(c)\, dc, \tag{59}$$

so that the objective function in (Γ) is rewritten

$$\widetilde{W} = \left[\widetilde{W}(\underline{c}, c_1) + \widetilde{W}(c_1, c_2) + \widetilde{W}(c_2, c_3) + \widetilde{W}(c_3, c_4) + \widetilde{W}(c_4, \overline{c})\right].$$

As the maximization of the expected utility in each cost interval is independent of that in any other interval, we treat the various intervals separately. We have already established that, in the situation under scrutiny, FB attains $\forall c \in [c_2, c_3]$ (Lemma 4) and we shall not come back to this case.

The solution for $c \in [\underline{c}, c_1)$ This proof is close to that of Gary-Bobo and Spiegel [7]. We first calculate the *ex post* transfer, then the expected transfer for $c \in [\underline{c}, c_1)$, namely $E(t_1)$. We finally replace it into the expression of $\widetilde{W}(\underline{c}, c_1)$ and optimize with respect to quantity.

The *ex post* transfer when $c \in [\underline{c}, c_1)$ It is useful to define $t(c,s) \equiv g(c)$ the transfer the agent receives when $s = \widetilde{s}(c)$ and $t(c,s) \equiv h(c,s)$ the transfer he receives when $s \neq \widetilde{s}(c)$. For sake of simplicity, $h(c,s)$ is defined for any $s \in N$, although in reality $h(c, \widetilde{s}(c))$ does not exist (as the agent is not punished in state $\widetilde{s}(c)$). Replacing into (43) and rearranging, we get

$$g'(c) = \sum_{s=1}^{n} cq'(c)\frac{p(c,s)}{p(c,\widetilde{s}(c))} - \sum_{s=1}^{n}\frac{dh(c,s)}{dc}\frac{p(c,s)}{p(c,\widetilde{s}(c))} + \frac{dh(c,\widetilde{s}(c))}{dc}$$

Define $c_k \in \{c_1, c_4\}$ any type c for which $E_s[\pi(c_k,s)] = 0$. Integrating all terms above from c to c_k we obtain

$$\begin{aligned} g(c) &= g(c_k) - \int_c^{c_k}\sum_{s=1}^{n} xq'(x)\frac{p(x,s)}{p(x,\widetilde{s}(c))}dx \\ &\quad + \int_c^{c_k}\sum_{s=1}^{n}\frac{dh(x,s)}{dx}\frac{p(x,s)}{p(x,\widetilde{s}(c))}dx - h(c_k,\widetilde{s}(c)) + h(c,\widetilde{s}(c)). \end{aligned} \tag{60}$$

Integrating by parts the second and the third term in the RHS of (60), we rewrite it as

$$\begin{aligned} g(c) &= g(c_k) - h(c_k,\widetilde{s}(c)) + h(c,\widetilde{s}(c)) \\ &\quad - \sum_{s=1}^{n} c_k q(c_k)\frac{p(c_k,s)}{p(c_k,\widetilde{s}(c))} + \sum_{s=1}^{n} h(c_k,s)\frac{p(c_k,s)}{p(c_k,\widetilde{s}(c))} \\ &\quad + \sum_{s=1}^{n} cq(c)\frac{p(c,s)}{p(c,\widetilde{s}(c))} - \sum_{s=1}^{n} h(c,s)\frac{p(c,s)}{p(c,\widetilde{s}(c))} \\ &\quad + \int_c^{c_k}\sum_{s=1}^{n} q(x)\frac{d}{dx}\left[x\frac{p(c,s)}{p(x,\widetilde{s}(c))}\right]dx - \int_c^{c_k}\sum_{s=1}^{n} h(x,s)\frac{d}{dx}\left[\frac{p(x,s)}{p(x,\widetilde{s}(c))}\right]dx. \end{aligned} \tag{61}$$

Denote

$$\psi_s(c) \equiv \frac{p(c,s)}{p(c,\widetilde{s}(c))}. \tag{62}$$

Using it in (61) we obtain

$$\begin{aligned} g(c) &= \sum_{s=1}^{n}\left[h(c_k,s) - c_k q(c_k)\right]\psi_s(c_k) + g(c_k) \\ &\quad - h(c_k,\widetilde{s}(c)) + \sum_{s=1}^{n} q(c)\,c\psi_s(c) + \int_c^{c_k}\sum_{s=1}^{n} q(x)\frac{d}{dx}\left[x\psi_s(x)\right]dx \\ &\quad - \sum_{s=1}^{n} h(c,s)\psi_s(c) - \int_c^{c_k}\sum_{s=1}^{n} h(x,s)\psi_s'(x)\,dx + h(c,\widetilde{s}(c)). \end{aligned} \tag{63}$$

Using (2) we can group the expression

$$\begin{aligned} &\sum_{s=1}^{n}\left[h(c_k,s) - c_k q(c_k)\right]\psi_s(c_k) + r(c_k) - h(c_k,\widetilde{s}(c_k)) \\ = &\sum_{s=1}^{n}\left[\pi(c_k,s) + K(c_k)\right]\psi_s(c_k). \end{aligned}$$

Replacing into (63) returns

$$\begin{aligned} g(c) &= \sum_{s=1}^{n}\left[\pi\left(c_{k}, s\right)+K\left(c_{k}\right)\right] \psi_{s}\left(c_{k}\right)-\sum_{s=1}^{n} h(c, s) \psi_{s}(c)+\sum_{s=1}^{n} c q(c) \psi_{s}(c) \\ &+h(c, \widetilde{s}(c))+\int_{c}^{c_{k}} \sum_{s=1}^{n} q(x) \frac{d}{d x}\left[x \psi_{s}(x)\right] d x-\int_{c}^{c_{k}} \sum_{s=1}^{n} h(x, s) \psi_{s}^{\prime}(x) d x . \end{aligned} \tag{64}$$

Using (2) as well as $t(c,s)=h(c,s)$ for $s \neq \widetilde{s}(c)$ and letting $\underline{\pi}^{sb}(c,s)$ the loss, we have $\underline{\pi}^{sb}(c,s)=h(c,s)-[cq(c)+K(c)]$. We use this to rewrite the expression

$$\begin{aligned} & \int_{c}^{c_{k}} \sum_{s=1}^{n} q(x) \frac{d}{d x}\left[x \psi_{s}(x)\right] d x-\int_{c}^{c_{k}} \sum_{s=1}^{n} h(x, s) \psi_{s}^{\prime}(x) d x \\ = & \int_{c}^{c_{k}} \sum_{s=1}^{n}\left[q(x) \psi_{s}(x)+[x q(x)-h(x, s)] \psi_{s}^{\prime}(x)\right] d x \\ = & \int_{c}^{c_{k}} \sum_{s=1}^{n}\left[q(x) \psi_{s}(x)-\left[\underline{\pi}^{sb}(x, s)+K(x)\right] \psi_{s}^{\prime}(x)\right] d x . \end{aligned}$$

Replacing this into (64) yields the *ex post* transfer

$$\begin{aligned} g(c) = & \sum_{s=1}^{n}\left[\pi\left(c_{k}, s\right)+K\left(c_{k}\right)\right] \psi_{s}\left(c_{k}\right)+\sum_{s=1}^{n} c q(c) \psi_{s}(c) \\ & -\left[\sum_{s=1}^{n} h(c, s) \psi_{s}(c)-h(c, \widetilde{s}(c))\right] \\ & +\int_{c}^{c_{k}} \sum_{s=1}^{n} q(x) \psi_{s}(x)-\left[\underline{\pi}^{sb}(x, s)+K(x)\right] \psi_{s}^{\prime}(x) d x . \end{aligned} \tag{65}$$

The expected transfer for $c \in [\underline{c}, c_1)$ Using the notation $h(c,s)$ and $g(c)$ as defined above, the expected transfer $E(t_1)$ when $c < c_1$ is given by

$$E\left(t_{1}\right)=\int_{\underline{c}}^{c_{1}}\left[\sum_{s=1}^{n-1} h(c, s) p(c, s)+g(c) p(c, n)\right] f(c) d c$$

Substitute $\widetilde{s}(c)=n$ and $c_k=c_1$ into (65) and then substitute $g(c)$ from (65) into the above expression. This yields

$$\begin{aligned} E\left(t_{1}\right)= & \int_{\underline{c}}^{c_{1}}\left\{\sum_{s=1}^{n}\left(\left[\pi\left(c_{1}, s\right)+K\left(c_{1}\right)\right] \psi_{s}\left(c_{1}\right)+c q(c) \psi_{s}(c)\right)\right. \\ & \left.+\sum_{s=1}^{n}\left(q(c) \psi_{s}(c)-\left[\underline{\pi}^{sb}(c, s)+K(c)\right] \psi_{s}^{\prime}(c)\right)\right\} p(c, n) f(c) d c \end{aligned}$$

Define

$$\phi(c) \equiv \int_{\underline{c}}^{c} p(x, n) f(x) d x, \forall c \in\left[\underline{c}, c_{1}\right) \tag{66}$$

for any $c \in [\underline{c}, c_1)$. We calculate

$$\int_{\underline{c}}^{c_1} \left\{ \int_{c}^{c_1} \left[\sum_{s=1}^{n} \left(q(x)\, \psi_s(x) - \left[\underline{\pi}^{sb}(x,s) + K(x) \right] \psi_s'(x) \right) \right] dx \right\} p(c,n)\, f(c)\, dc$$

$$= \int_{\underline{c}}^{c_1} \left\{ \sum_{s=1}^{n} \left(q(c)\, \psi_s(c) - \left[\underline{\pi}^{sb}(c,s) + K(c) \right] \psi_s'(c) \right) \right\} \phi(c)\, dc$$

We thus find

$$\begin{aligned} E(t_1) &= \int_{\underline{c}}^{c_1} \left\{ \sum_{s=1}^{n} \left([\pi(c_1,s) + K(c_1)]\, \psi_s(c_1) + cq(c)\, \psi_s(c) \right) \right\} p(c,n)\, f(c)\, dc \\ &\quad + \int_{\underline{c}}^{c_1} \left\{ \sum_{s=1}^{n} \left(q(c)\, \psi_s(c) - \left[\underline{\pi}^{sb}(c,s) + K(c) \right] \psi_s'(c) \right) \right\} \phi(c)\, dc. \qquad (67) \end{aligned}$$

The optimal output for $c \in [\underline{c}, c_1)$ Substituting (67) into (2) and then (2) into the expression of $\widetilde{W}(\underline{c}, c_1)$ from (59), we rewrite it as follows:

$$\begin{aligned} \widetilde{W}(\underline{c}, c_1) &= \int_{\underline{c}}^{c_1} \sum_{s=1}^{n} \left[S(q(c)) - \alpha c q(c) - \alpha K(c) \right] p(c,s)\, f(c)\, dc \qquad (68) \\ &\quad - (1-\alpha) \int_{\underline{c}}^{c_1} \left\{ \sum_{s=1}^{n} \left([\pi(c_1,s) + K(c_1)]\, \psi_s(c_1) \right. \right. \\ &\quad \left. \left. + cq(c)\, \psi_s(c) \right) \right\} p(c,n)\, f(c)\, dc \\ &\quad - (1-\alpha) \int_{\underline{c}}^{c_1} \left\{ \sum_{s=1}^{n} \left(q(c)\, \psi_s(c) - \left[\underline{\pi}^{sb}(c,s) + K(c) \right] \psi_s'(c) \right) \right\} \phi(c)\, dc \end{aligned}$$

From the definition of c_1 (see Lemma 5), $E_s[\pi(c_1,s)] = 0$. Also, because $\widetilde{W}(\underline{c}, c_1)$ decreases with $\underline{\pi}^{sb}(c,s)$, it is optimal to set the latter at the lowest feasible value *i.e.*, $\underline{\pi}^{sb}(c,s) = -L$. Replacing into $\widetilde{W}(\underline{c}, c_1)$, the first-order condition with respect to q, $\forall c \in [\underline{c}, c_1)$, is given by

$$\begin{aligned} & \left[S'(q(c)) - \alpha c \right] p(c,s)\, f(c) \\ = \; & (1-\alpha) \left[c\psi_s(c)\, p(c,n)\, f(c) + \psi_s(c)\, \phi(c) \right]. \end{aligned}$$

Denoting $q^I(c)$ the quantity that satisfies the condition above together with (62) and (66), we can rewrite

$$\begin{aligned} S'\left(q^I(c)\right) &= \alpha c + (1-\alpha) \frac{\psi_s(c)}{p(c,s)\, f(c)} \left[cp(c,n)\, f(c) + \phi(c) \right] \\ &= c + (1-\alpha) \frac{\int_{\underline{c}}^{c} p(x,s)\, f(x)\, dx}{p(c,n)\, f(c)} \\ &= c + (1-\alpha) \frac{F(c\,|n)}{f(c\,|n)}, \end{aligned}$$

with $F(c\,|n) = \frac{\int_{\underline{c}}^{c} p(x,n) f(x) dx}{\int_{\underline{c}}^{\overline{c}} p(x,n) f(x) dx}$ and $f(c\,|n) = \frac{p(c,n) f(c)}{\int_{\underline{c}}^{\overline{c}} p(x,n) f(x) dx}$.

The solution for $c \in [c_1, c_2)$ From Lemma 5 one has $E_s[\pi(c,s)] = 0$ whenever $c \in [c_1, c_2)$. It means that the functional form of the *ex post* profit $\pi(c,s)$ is similar to that in (4) and (5), except that $q^{fb}(c)$ is replaced by $q^{II}(c)$, the value of which we need to determine. In particular, the punishment is given by

$$\underline{\pi}^{sb}(c,s) = -\left[q^{II}(c) + K'(c)\right] \frac{p(c,n)}{dp(c,n)/dc}.$$

Moreover, by Lemma 4, $\underline{\pi}^{sb}(c,s) = -L, \forall c \in [\underline{c}, c_2)$. Using this in the expression above, $q^{II}(c)$ is found to be as defined in (18).

The proof is identical for $c \in (c_3, c_4]$.

The solution for $c \in (c_4, \overline{c}]$ Define

$$\varphi(c) \equiv \int_c^{\overline{c}} p(x,1) f(x)\, dx, \forall c \in (c_4, \overline{c}].$$

Proceeding as for $c \in [\underline{c}, c_1)$, one finds the expected transfer $E(t_2)$ that corresponds to $c \in (c_4, \overline{c}]$ as follows

$$\begin{aligned} E(t_2) &= \int_{c_4}^{\overline{c}} \left\{ \sum_{s=1}^{n} \left([\pi(c_4,s) + K(c_4)]\, \psi_s(c_4) + cq(c)\, \psi_s(c) \right) \right\} p(c,1) f(c)\, dc \\ &+ \int_{c_4}^{\overline{c}} \left\{ \sum_{s=1}^{n} \left(\left[\underline{\pi}^{sb}(c,s) + K(c)\right] \psi_s'(c) - q(c)\, \psi_s(c) \right) \right\} \varphi(c)\, dc \end{aligned} \tag{69}$$

Substituting (69) into $\widetilde{W}(c_4, \overline{c})$, we can characterize the optimal output $q^V(c)$ as

$$S'\left(q^V(c)\right) = c - (1-\alpha) \frac{1 - F(c\,|1)}{f(c\,|1)},$$

with $[1 - F(c\,|1)] = \frac{\int_c^{\overline{c}} p(x,1)f(x)dx}{\int_{\underline{c}}^{\overline{c}} p(x,1)f(x)dx}$ and $f(c\,|1) = \frac{p(c,1)f(c)}{\int_{\underline{c}}^{\overline{c}} p(x,1)f(x)dx}$.

B.3.2 Verify that (GIC) is satisfied

The *interim* profit of the agent when he reports r is written similarly to (47) *i.e.*,

$$E_s[\widetilde{\pi}(r,s)] = q^{sb}(r)(r-c) + K(r) - K(c) + \sum_{s=1}^{n} \pi(r,s)\, p(c,s), \tag{70}$$

with $q^{sb} \in \left\{q^I, q^{II}, q^{III}, q^{IV}, q^V\right\}$ (see the proof above). The *ex post* profit $\pi(r,s)$ that appears in the expression in (70) is calculated in a different way according to the value the report r takes. We thus develop the analysis case by case.

Case $r \in [\underline{c}, c_1]$ We proceed as follows. We first calculate the *ex post* profit $\pi(r,s)$, $\forall r \in [\underline{c}, c_1]$. We replace into (70) so as to calculate $E_s[\widetilde{\pi}(r,s)]$. We finally state the global incentive condition $E_s[\widetilde{\pi}(r,s)] \leq E_s[\pi(c,s)]$ for any report $r \in [\underline{c}, c_1]$. Two sub-cases are considered, namely $c \in [\underline{c}, c_1]$ and $c \in [c_1, \overline{c}]$.

The *ex post* profit $\pi(r,s)$ Recall that $c_k \in \{c_1, c_4\}$ is by definition a type c for which $E_s(\pi(c_k,s)) = 0$. Using the definition of c_k and replacing $\sum_{s=1}^{n} \psi_s(c_k) = \frac{1}{p(c_k,\widetilde{s}(c_k))}$ (from (62)) into (65), we obtain

$$\begin{aligned} g(c) &= \frac{K(c_k)}{p(c_k,\widetilde{s}(c_k))} + cq^{sb}(c)\sum_{s=1}^{n}\psi_s(c) - \sum_{s=1}^{n} h(c,s)\psi_s(c) + h(c,\widetilde{s}(c)) \\ &\quad + \int_{c}^{c_k}\left\{\sum_{s=1}^{n}\left[q^{sb}(x)\psi_s(x) - \left[\underline{\pi}^{sb}(x,s) + K(x)\right]\psi_s'(x)\right]\right\}dx. \end{aligned}$$

We further calculate

$$\begin{aligned} &\sum_{s=1}^{n} cq^{sb}(c)\psi_s(c) - \sum_{s=1}^{n} h(c,s)\psi_s(c) + h(c,\widetilde{s}(c)) \\ =&\ \sum_{s=1}^{n}\left[cq^{sb}(c) - h(c,s)\right]\psi_s(c) + h(c,\widetilde{s}(c)) \\ =&\ \sum_{s=1}^{n}\left[-K(c) - \underline{\pi}^{sb}(c,s)\right]\psi_s(c) + \underline{\pi}^{sb}(c,\widetilde{s}(c)) + cq^{sb}(c) + K(c) \\ =&\ \sum_{s=1}^{n}(-K(c)+L)\psi_s(c) - L + cq^{sb}(c) + K(c) \\ =&\ [L-K(c)]\frac{1-p(c,\widetilde{s}(c))}{p(c,\widetilde{s}(c))} + cq^{sb}(c) \end{aligned}$$

and then substitute into the expression of $g(c)$ above. Rearranging yields

$$\begin{aligned} g(c) &= cq^{sb}(c) + \int_{c}^{c_k} q^{sb}(x)\left(\sum_{s=1}^{n}\psi_s(x)\right)dx + \frac{1-p(c,\widetilde{s}(c))}{p(c,\widetilde{s}(c))}[L-K(c)] \\ &\quad + L\int_{c}^{c_k}\left(\sum_{s=1}^{n}\psi_s'(x)\right)dx - \int_{c}^{c_k}\sum_{s=1}^{n}(K(x)\psi_s'(x))\,dx + \frac{K(c_k)}{p(c_k,\widetilde{s}(c))}. \end{aligned}$$

Integrate $\int_c^{c_k}\left(\sum_{s=1}^{n}\psi_s'(x)\right)dx$ by parts, where $\psi_s(x)$ is defined by (62). Further integrate $\int_c^{c_k}\sum_{s=1}^{n}(K(x)\psi_s'(x))\,dx$. Then, replacing into the above expression of $g(c)$, we find

$$\begin{aligned} g(c) &= cq^{sb}(c) + K(c) + \frac{1-p(c_k,\widetilde{s}(c))}{p(c_k,\widetilde{s}(c))}L \\ &\quad + \int_{c}^{c_k}\sum_{s=1}^{n}\left[q^{sb}(x) + K'(x)\right]\psi_s(x)\,dx. \end{aligned} \tag{71}$$

Using (71) in (2) for $t(r,\widetilde{s}(r)) = g(r)$ (knowing that $g(r)$ is the transfer that corresponds to type $\widetilde{s}(r)$), the reward of the agent when he reports $r \in [\underline{c}, c_1]$ and $\widetilde{s}(r) = n$ is written

$$\pi(r,n) = \int_{r}^{c_1}\frac{q^I(x) + K'(x)}{p(x,n)}dx + \frac{1-p(c_1,n)}{p(c_1,n)}L. \tag{72}$$

From the proof above, $\pi(r,s) = -L$ whenever $r \in [\underline{c}, c_1]$ and $s \neq n$.

The *interim* profit Using (72) and $\pi(r,s) = -L$ for $s \neq n$ in (70), $E_s[\widetilde{\pi}(r,s)]$ is rewritten

$$\begin{aligned} E_s[\widetilde{\pi}(r,s)] &= -\int_r^c \left[q^I(r) + K'(x)\right] dx + p(c,n) \int_r^{c_1} \frac{q^I(x) + K'(x)}{p(x,n)} dx \\ &\quad - \left[1 - \frac{p(c,n)}{p(c_1,n)}\right] L, \end{aligned} \tag{73}$$

whereas the *interim* profit associated with a truthful report $r = c$ is given by

$$E_s[\pi(c,s)] = p(c,n) \int_c^{c_1} \frac{q^I(x) + K'(x)}{p(x,n)} dx - \left[1 - \frac{p(c,n)}{p(c_1,n)}\right] L. \tag{74}$$

Sub-case $c \in [\underline{c}, c_1]$ Using (73) and (74), we have $E_s[\pi(c,s)] \geq E_s[\widetilde{\pi}(r,s)]$ if and only if

$$\int_c^r \left[q^I(x) + K'(x)\right] \left[1 - \frac{p(c,n)}{p(x,n)}\right] dx + \int_c^r \left[q^I(r) - q^I(x)\right] dx \leq 0, \tag{75}$$

from which (25) follows.

Sub-case $c \in [c_1, \overline{c}]$ Assume that $r = c_1$ and calculate from (73) the following

$$\begin{aligned} \frac{dE_s[\widetilde{\pi}(c,s)]}{dc} &= -\left[q^I(c_1) + K'(c)\right] + \frac{L}{p(c_1,n)} \frac{dp(c,n)}{dc} \\ &= -\left[q^I(c_1) + K'(c_1)\right] + \frac{L}{p(c_1,n)} \frac{dp(c_1,n)}{dc_1} \\ &\quad + K'(c_1) - K'(c) + \frac{L}{p(c_1,n)} \left[\frac{dp(c,n)}{dc} - \frac{dp(c_1,n)}{dc_1}\right] \\ &= K'(c_1) - K'(c) + \frac{L}{p(c_1,n)} \left[\frac{dp(c,n)}{dc} - \frac{dp(c_1,n)}{dc_1}\right]. \end{aligned}$$

With $K"(\cdot) \geq 0$, we have $\frac{dE_s[\widetilde{\pi}(c_1,s)]}{dc} \leq 0$. Moreover, $E_s[\widetilde{\pi}(c_1,s)] = E_s[\pi(c_1,s)]$. From Lemma 5, $E_s[\pi(c_1,s)] = 0$. Therefore, $E_s[\widetilde{\pi}(c_1,s)] \leq 0$ whenever $c \in [c_1, \overline{c}]$ and $r = c_1$.

Take now $r \leq c_1$ and calculate

$$\frac{dE_s[\widetilde{\pi}(r,s)]}{dr} = -\int_r^c \left\{\left[q^I(r) + K'(r)\right] \frac{dp(x,n)/dx}{p(r,n)} + \frac{dq^I(r)}{dr}\right\} dx.$$

We look for the condition under which $\frac{dE_s[\widetilde{\pi}(r,s)]}{dr} \geq 0$. Because $c \geq c_1$ and $r \leq c_1$, this inequality holds if and only if

$$\frac{dq^I(r)}{dr} \leq -\left[q^I(r) + K'(r)\right] \frac{dp(x,n)/dx}{p(r,n)}, \forall r \in [\underline{c}, c_1] \text{ and } x \geq c_1, \tag{76}$$

which is implied by (25) together with Assumption 2 and $x \geq r$. As $\frac{dE_s[\widetilde{\pi}(r,s)]}{dr} \geq 0 \; \forall r \in [\underline{c}, c_1]$ and $c \in [c_1, \overline{c}]$, whereas $E_s[\widetilde{\pi}(c_1,s)] \leq 0$ (as previously found), one has $E_s[\widetilde{\pi}(r,s)] \leq 0$, $\forall r \in [\underline{c}, c_1]$ and $c \in [c_1, \overline{c}]$.

Overall, (25) and $K" \geq 0$ ensure that the agent has no incentive to report $r \in [\underline{c}, c_1]$ such that $r \neq c$, whatever his real type.

Case $r \in [c_1, c_2]$ As $\pi(r,s) = -L$ for all $s \neq n$ and, from Lemma 4, $E_s[\pi(r,s)] = 0$ for all $r \in [c_1, c_2]$, we have $\pi(r,n) = \frac{1-p(r,n)}{p(r,n)}L$. Substituting these values of $\pi(r,s)$ into $E_s[\widetilde{\pi}(r,s)]$, together with $q^{II}(r) = \frac{dp(r,n)/dr}{p(r,n)}L - K'(r)$ (as obtained in the proof above), we find that, with regards to this interval, (70) specifies as

$$E_s\left[\widetilde{\pi}^{II}\right] = \int_c^r \left\{ \frac{L}{p(r,n)} \left[\frac{dp(r,n)}{dr} - \frac{dp(x,n)}{dx} \right] + K'(x) - K'(r) \right\} dx.$$

From the expression above, $K''(\cdot) \geq 0$ and $\widetilde{s}(c) = n$, we deduce that $E_s\left[\widetilde{\pi}^{II}\right] \leq 0$.

Case $r \in [c_2, c_3]$ The condition for global incentive compatibility is given by (12), just as for Proposition 1 and 2.

Case $r \in [c_3, c_4]$ Proceeding as we did for $r \in [c_1, c_2]$, we find that the payoff of the agent when he reports r is written

$$\begin{aligned} E\left[\widetilde{\pi}^{IV}\right] &= \int_c^r \left[L\frac{dp(r,1)/dr}{p(r,1)} - K'(r) + K'(x) - \frac{L}{p(r,1)}\frac{dp(x,1)}{dx} \right] dx \\ &= \int_c^r \left\{ \frac{L}{p(r,1)} \left[\frac{dp(r,1)}{dr} - \frac{dp(x,1)}{dx} \right] + K'(x) - K'(r) \right\} dx. \end{aligned}$$

$K"(\cdot) \geq 0$ and $\widetilde{s}(c) = 1$ imply that $E\left[\widetilde{\pi}^{IV}\right] \leq 0$.

Case $r \in [c_4, \overline{c}]$ We follow the same steps as we did to treat the very first case.

The *ex post* profit $\pi(r,s)$ Using (71) in (2) for $t(r, \widetilde{s}(r)) = g(r)$, the reward of the agent when he reports $r \in [c_4, \overline{c}]$ and $\widetilde{s}(r) = 1$ is written

$$\pi(r,1) = -\int_{c_4}^r \frac{q^V(x) + K'(x)}{p(x,1)} dx + \frac{1 - p(c_4,1)}{p(c_4,1)} L, \tag{77}$$

From the proof above, $\pi(r,s) = -L$ whenever $r \in [\underline{c}, c_4]$ and $s \neq 1$.

The *interim* profit The *interim* profit of the agent when he reports r is given by

$$\begin{aligned} E\left[\widetilde{\pi}^V\right] &= \int_c^r \left[q^V(r) + K'(x)\right] dx - \int_{c_4}^r \left[q^V(x) + K'(x)\right] \frac{p(c,1)}{p(x,1)} dx \\ &\quad -L\left[1 - \frac{p(c,1)}{p(c_4,1)}\right], \end{aligned} \tag{78}$$

whereas the *interim* profit in case of truthtelling is written

$$E_s[\pi(c,s)] = -\int_{c_4}^c \left[q^V(x) + K'(x)\right] \frac{p(c,1)}{p(x,1)} dx - L\left[1 - \frac{p(c,1)}{p(c_4,1)}\right]. \tag{79}$$

Sub-case $c \in [c_4, \overline{c}]$ Using (78) and (79), it is $E_s[\pi(c,s)] \geq E_s[\widetilde{\pi}(r,s)]$ if and only if

$$\int_c^r \left[q^V(x) + K'(x)\right]\left[1 - \frac{p(c,1)}{p(x,1)}\right] dx + \int_c^r \left[q^V(r_i) - q^V(x)\right] dx \leq 0,$$

which is implied by (26).

Sub-case $c \notin [c_4, \overline{c}]$ Take first $r = c_4$ and calculate

$$\begin{aligned}
\frac{dE\left[\widetilde{\pi}^V\right]}{dc} &= -\left[q^V(c_4) + K'(c)\right] + \frac{dp(c,1)/dc}{p(c_4,1)}L \\
&= -\left[q^V(c_4) + K'(c_4)\right] + \frac{dp(c_4,1)/dc_4}{p(c_4,1)}L \\
&\quad + K'(c_4) - K'(c) + \frac{dp(c,1)/dc - dp(c_4,1)/dc_4}{p(c_4,1)}L \\
&= K'(c_4) - K'(c) + \frac{dp(c,1)/dc - dp(c_4,1)/dc_4}{p(c_4,1)}L.
\end{aligned}$$

As $K"(\cdot) \geq 0$, $\frac{dE\left[\widetilde{\pi}^V\right]}{dc} \geq 0$. Moreover, $E\left[\widetilde{\pi}^V\right] = 0$ if $c = r = c_4$. This shows that any type $c \notin [c_4, \overline{c}]$ that reports $r = c_4$ obtains $E\left[\widetilde{\pi}^V\right] \leq 0$. Furthermore,

$$\begin{aligned}
\frac{dE\left[\widetilde{\pi}^V\right]}{dr} &= \left[q^V(r) + K'(r)\right]\left[1 - \frac{p(c,1)}{p(r,1)}\right] + \int_c^r \frac{dq^V(r)}{dr} dx \\
&= \int_c^r \left\{\frac{dq^V(r)}{dr} + \left[q^V(r) + K'(r)\right]\frac{dp(x,1)/dx}{p(r,1)}\right\} dx.
\end{aligned}$$

$E\left[\widetilde{\pi}^V\right] \leq 0$ for any report $r \in [c_4, \overline{c}]$ if $\frac{dE\left[\widetilde{\pi}^V\right]}{dr} \leq 0$, which is implied by

$$\frac{dq^V(r)}{dr} \leq -\left[q^V(r) + K'(r)\right]\frac{dp(x,1)/dx}{p(r,1)}.$$

In turn, this is implied by (26) together with Assumption 2 and $x \leq r$.

Overall, $K"(\cdot) \geq 0$ and (26) imply that the agent has no incentive to report $r \in [c_4, \overline{c}]$ such that $r \neq c$.

B.4 Proof of Corollary 1

See the text below the corollary.

B.5 Proof of Corollary 2

B.5.1 Proof of (i)

Recall (17) in Proposition 3. Let $q^{I}_{(NS)}(c)$ denote the optimal quantity that would arise in the absence of informative signals. It would be characterized as follows:

$$S'(q^{I}_{(NS)}(c)) = c + (1-\alpha)\frac{F(c)}{f(c)}.$$

For any given $c \in (\underline{c}, c_1)$, it is $q^{I}(c) > q^{I}_{(NS)}(c)$ if and only if $S'\left(q^{I}(c)\right) < S'(q^{I}_{(NS)}(c))$, which requires that

$$\frac{F(c\,|n)}{f(c\,|n)} < \frac{F(c)}{f(c)}. \tag{80}$$

Recall that, for all feasible c,

$$F(c\,|n) = \frac{\int_{\underline{c}}^{c} p(x,n)\, f(x)\, dx}{\int_{\underline{c}}^{\overline{c}} p(x,n)\, f(x)\, dx} \quad \text{and} \quad f(c\,|n) = \frac{p(c,n)\, f(c)}{\int_{\underline{c}}^{\overline{c}} p(x,n)\, f(x)\, dx}.$$

Hence, we can write

$$\frac{F(c\,|n)}{f(c\,|n)} = \frac{\frac{\int_{\underline{c}}^{c} p(x,n) f(x) dx}{\int_{\underline{c}}^{\overline{c}} p(x,n) f(x) dx}}{\frac{p(c,n) f(c)}{\int_{\underline{c}}^{\overline{c}} p(x,n) f(x) dx}} = \frac{\int_{\underline{c}}^{c} p(x,n)\, f(x)\, dx}{p(c,n)\, f(c)}.$$

Therefore, (80) holds if and only if

$$\frac{\int_{\underline{c}}^{c} p(x,n)\, f(x)\, dx}{p(c,n)\, f(c)} < \frac{F(c)}{f(c)}$$

or, equivalently,

$$\int_{\underline{c}}^{c} p(x,n)\, f(x)\, dx < p(c,n)\, F(c) = \int_{\underline{c}}^{c} p(c,n)\, f(x)\, dx.$$

Under Assumption 2, $p(x,n) < p(c,n)$ for all x. Hence, we have

$$\int_{\underline{c}}^{c} p(x,n)\, f(x)\, dx < \int_{\underline{c}}^{c} p(c,n)\, f(x)\, dx, \ \forall c.$$

This means that, for all $c \in [\underline{c}, c_1)$, $q^{I}(c) > q^{I}_{(NS)}(c)$.

The proof that $q^{V}(c) < q^{V}_{(NS)}(c)$ for all $c \in (c_4, \overline{c})$, with $q^{V}_{(NS)}(c)$ the optimal quantity that would arise in the absence of informative signals, is analogous and thus omitted.

B.5.2 Proof of (ii)

It is evident from (22) and (24).

B.5.3 The presence of informative signals reduces the amount of pooling

From (i) recall that $q^I_{(NS)}(c) < q^I(c)$ for all $c \in (\underline{c}, c_1)$ and that $q^V(c) < q^V_{(NS)}(c)$ for all $c \in (c_4, \overline{c})$. Moreover, from the explanation that follows Corollary 1, recall that pooling starts at the cost value c_1 that is defined by $q^I(c_1) = q^{II}(c_1)$, with $q^{II}(c) = -K'(c)$. With no informative signal, pooling starts at some cost value $c_1^{(NS)}$ that is defined by $q^I_{(NS)}(c_1^{(NS)}) = q^{II}(c_1^{(NS)})$. Both $q^I(c)$ and $q^I_{(NS)}(c)$ decrease with c and $q^I_{(NS)}(c) < q^I(c)$ for all $c \in (\underline{c}, c_1)$. It follows that $c_1^{(NS)} < c_1$. The proof that pooling ends at some cost value above c_4 in the absence of informative signals, is analogous and thus omitted.

B.6 Proof of Proposition 4

We proceed as follows. We begin by showing that, whenever $K''(\cdot) < 0$, there exists no $c \neq \widehat{c}$ for which (PC) is binding at the SB solution. We then rewrite (Γ) for the situation in which (PC) is not binding for all $c \neq \widehat{c}$ and prove the proposition.

Suppose that (PC) is binding over some non-empty interval $\left[c^L, c^H\right]$, with either $c^L \neq \widehat{c}$ or $c^H \neq \widehat{c}$ or both. Assume also that FB is not implementable over this interval at the solution to (Γ). From the proof of Proposition 3, the SB quantity would be $q^{II}(c)$ for all $c \in \left[c^L, c^H\right]$. That same proof shows that the quantity $q^{II}(c)$ and the transfers that leave no rent to the agent are unfeasible when $K"(\cdot) < 0$. This contradicts the assumption that (PC) is binding and, at the same time, FB is not at hand for types in $\left[c^L, c^H\right]$. Now suppose that, at the solution to (Γ), (PC) is binding and FB is implemented for all $c \in \left[c^L, c^H\right]$. From Lemma 5, it follows that there exist other cost values around $\left[c^L, c^H\right]$ for which (PC) is binding and the SB quantity is $q^{II}(c)$. Again, from the proof of Proposition 3, this is unfeasible when $K"(\cdot) < 0$. In turn, the hypothesis that FB is enforced for all $c \in \left[c^L, c^H\right]$ is contradicted as well. Overall, there exists no subset $\left[c^L, c^H\right]$, with either $c^L \neq \widehat{c}$ or $c^H \neq \widehat{c}$ or both, in which (PC) is binding. It follows that (PC) is slack for all $c \neq \widehat{c}$. Hence, the interval $[c_2, c_4]$ defined by Lemma 5 reduces to the singleton $\{\widehat{c}\}$.

From the proof of Proposition 3, the scheme is globally incentive compatible whenever

$$\frac{dq^{sb}(c)}{dc} \leq -\left[q^{sb}(c) + K'(c)\right] \frac{dp(c, \widetilde{s}(c))/dc}{p(c, \widetilde{s}(c))}, \ \forall c \in [\underline{c}, \overline{c}], \tag{81}$$

$q^{sb}(c)$ being the SB quantity for type c. We can thus rewrite (Γ) as

$$\max_{q(c)} \widetilde{W} \equiv \left[\widetilde{W}(\underline{c}, \widehat{c}) + \widetilde{W}(\widehat{c}, \overline{c})\right] \quad s.t.\ (81),$$

where $\widetilde{W}(\underline{c}, \widehat{c})$ and $\widetilde{W}(\widehat{c}, \overline{c})$, as defined in the proof of Proposition 3, are such that, beside (81), all other relevant constraints are satisfied. In particular, $\widetilde{W}(\underline{c}, \widehat{c})$ is defined by (59), with $\widehat{c}$ replacing c_1 and $\underline{\pi}^{sb}(c, s) = -L$. As long as (81) is slack, the optimal quantities are given by $q^I(c)$ and $q^V(c)$, as characterized by (17) and (21), respectively, in Proposition 3. In what follows, we check under which conditions $q^I(c)$ and $q^V(c)$ do not satisfy (81). To begin with, we check whether a pooling interval exists. At this aim, we rewrite (81) as

$$q^{sb}(c) \geq q^{sb}(\widehat{c}) + \int_c^{\widehat{c}} \left[q^{sb}(x) + K'(x)\right] \frac{dp(x,h)/dx}{p(x,h)} dx, \ \forall c \in [\underline{c}, \widehat{c}) \tag{82}$$

$$q^{sb}(c) \leq q^{sb}(\widehat{c}) - \int_{\widehat{c}}^{c} \left[q^{sb}(x) + K'(x)\right] \frac{dp(x,j)/dx}{p(x,j)} dx, \ \forall c \in (\widehat{c}, \overline{c}]. \tag{83}$$

Suppose that $q^{sb}(x) = q^{sb}(\widehat{c})$ for all $x \in [c, \widehat{c}]$. In particular, $q^{sb}(c) = q^{sb}(\widehat{c})$. Then, with $K" < 0$ and $q^{sb}(\widehat{c}) + K'(\widehat{c}) = 0$, it is $q^{sb}(x) + K'(x) = q^{sb}(\widehat{c}) + K'(x) > 0$ for any $x < \widehat{c}$. This shows that (82) is not satisfied unless $q^{sb}(c) > q^{sb}(\widehat{c})$. This contradicts the hypothesis that the two quantities are equal. The proof is similar for $c > \widehat{c}$ and, thus, here omitted. We conclude that pooling of quantities does not arise at optimum.

We next check under which conditions $q^I(c)$ and $q^V(c)$ satisfy (81), knowing that pooling is not induced. We have $q^{sb}(c) > q^{sb}(\widehat{c})$ when $c < \widehat{c}$ and $q^{sb}(c) < q^{sb}(\widehat{c})$ when $c > \widehat{c}$. For $c = \widehat{c}$, (82) and (83) are both trivially satisfied. Moreover, from the proof of Proposition 3, $q^{sb}(\widehat{c}) = q^{fb}(\widehat{c})$. As c approaches $\widehat{c}$ from the left, the solution $q^I(c) < q^{fb}(\widehat{c})$ does not satisfy the necessary condition $q^{sb}(c) > q^{fb}(\widehat{c})$, under which (81) is satisfied. Hence, (81) is binding for all $x \in [c^-, \widehat{c})$, for some $c^- < \widehat{c}$. Analogously, (81) is binding for all $x \in (\widehat{c}, c^+]$, for some $c^+ > \widehat{c}$.

B.7 Proof of Corollary 3

When $\frac{dp(c,\widetilde{s}(c))/dc}{p(c,\widetilde{s}(c))}$ $\forall c \in [\underline{c}, \overline{c}]$, (81) reduces to $\frac{dq^{sb}(c)}{dc} \leq 0$ and we are back to the setup of Lewis and Sappington [9].