XXI CONFERENZA

PUBLIC CHOICE E POLITICAL ECONOMY
I fondamenti positivi della teoria di finanza pubblica

Pavia, Università, 24-25 settembre 2009

MULTI-AGENT CONTRACTING WITH COUNTERVAILING INCENTIVES AND LIMITED LIABILITY

DANIEL DANAU, ANNALISA VINELLA

società italiana di economia pubblica

dipartimento di economia pubblica e territoriale – università di Pavia

Multi-agent contracting with countervailing incentives and limited liability*

Daniel Danau[†] Annalisa Vinella[‡]

This version: September 6, 2009

Abstract

We consider a principal who deals with two privately informed agents protected by limited liability. Their technologies are such that the fixed costs decline with the marginal costs (the types), which are correlated. Because of this, they display countervailing incentives to misrepresent type. We show that, with high liability, the first-best outcome can be effected for any type if (1) the fixed cost is non-concave in type, under the contract that yields the smallest feasible loss to agents; (2) the fixed cost is not very concave in type, under the contract that yields the maximum sustainable loss to agents. We further show that, with low liability, the first-best outcome is still implementable for a non-degenerate range of types if the fixed cost is less concave in type than some given threshold, which tightens as the liability reduces, and that the optimal contract entails pooling otherwise.

Keywords: Countervailing incentives; Limited liability; Correlation; Pooling

J.E.L. Classification Numbers: D82

---

*We are grateful to participants at the first NERI Meeting (Pescara) for providing useful comments on a earlier version.

[†]University of St Andrews, School of Economics and Finance, Castlecliffe, The Scores, Fife, UK, KY16 9AL (Scotland). E-mail: daniel.danau@gmail.com

[‡]University of Bari, Department of Economics and Mathematics, Via C. Rosalba 53, 70124 Bari (Italy). *Corresponding author.* E-mail: a.vinella@dse.uniba.it

1 Introduction

In some agency problems, the agent displays countervailing incentives, *i.e.* the temptation both to overstate and to understate his private information (the type) in the report to the principal, depending upon its specific realization. The existing literature about countervailing incentives only focuses on cases in which the principal deals with a single agent. However, there exist real-world situations in which this assumption appears to be restrictive.

To fix ideas, consider the situations in which electricity is produced by local monopolists on franchised territories within some given geographical area, under the regulation of a unique central authority entrusted for the whole area. The technologies that are used for power generation typically involve an inverse relation between marginal and fixed costs. As Lewis and Sappington [6] point out, this is a source of countervailing incentives. This means that, as long as marginal costs are privately known in those contexts, regulator and producers are actually in a principal/multi-agent relationship in which agents have countervailing incentives to misrepresent type. Insisting on regulated utilities, further examples can be found in electricity distribution and water and sewage services. In the latter, higher maintenance costs (that are variable) typically allow for lower depreciation costs of capital (that are fixed) and firms go subject to centralized regulation, in general[1].

When agents have correlated types but behave non-cooperatively, the distortions that come from informational asymmetries can be either removed or (at least) alleviated by properly designing a unique grand-contract for all of them. With a centralized incentive scheme that conditions one agent's compensation on the others' reports, the principal can take advantage of the information externalities generated by the reports and extract surplus (Crémer and McLean [2], Riordan and Sappington [10], McAfee and Reny [9]). In the examples previously made, it is typically the case that, beside inducing countervailing incentives in the report made to the regulator, costs are correlated across firms, provided that they all operate in the same sector. In fact, the presence of information correlation is one of the reasons why regulation is entrusted to a unique authority.

Considering all this, in the present paper we study centralized contracting between a principal and multiple agents whose private information is correlated and

[1]Electricity distribution is a network activity that constitutes a natural monopoly. Both in Europe and in America, the activity is typically performed by local monopolists under geographical franchises that make the service mandatory to all existing and new users in the concerned areas. The situation is similar in the water sector. The distribution of supplies to individual properties and the subsequent removal of sewage are classic network monopolies. Bulk supply provision, water treatment and sewage treatment normally enjoy spatial monopoly because of the high costs involved in transporting bulky water products.

who have countervailing incentives. At this aim, we model a situation in which, for each agent, the production technology includes a fixed cost that declines with the privately known marginal cost (the type), which takes value in a continuum support. This representation of the technology is similar to that Lewis and Sappington [6] use to introduce countervailing incentives in their single-agent model. From this perspective, we extend the analysis of Lewis and Sappington [6] to a multi-agent environment with correlated information[2]. Moreover, for the analysis to be truly positive, we focus on the realistic case in which agents are limitedly liable, *i.e.* they can only sustain bounded financial losses *ex post*.

The extent to which the principal can benefit from correlation depends on how deep agents' pockets are. Could agents sustain unbounded deficits *ex post*, the principal would be able to enforce the efficient quantity and retain all surplus *ex ante* (the first-best outcome). Otherwise, quantity distortions and information rents may appear (a second-best outcome). The efficiency of the contractual outcome also depends on the curvature of the agent's cost function with respect to type, an aspect that becomes relevant under limited liability. The more concave the cost function, the larger the marginal gain from misreporting, so that the agent has a stronger incentive to cheat. Thus, under limited liability, it may be more costly for the principal to induce information release. Studying the relationship between the level of agents' liability and their technological features, which are the source of countervailing incentives, and assessing the way this affects the design of the optimal contract are core dimensions to our analysis.

We begin by exploring how first best can be implemented under limited liability. As a first step of the analysis, we characterize the incentive scheme that most likely enforces first best by yielding the lowest feasible deficit to agents (*i.e.*, by being the least likely to violate the limits on liability) and provide a sufficient condition on agents' cost functions for this outcome to arise.

Our first result is that the design of the incentive scheme that minimizes the agents' loss is specifically affected by the presence of countervailing incentives. Let us clarify this point. Under this scheme, whether countervailing incentives arise

[2] The modelling device we use for the technology is flexible enough to sketch circumstances other than technological features *stricto sensu*. Indeed, as Lewis and Sappington [6] evidence, it further captures the possibility that countervailing incentives appear when valuable managerial skills are associated with higher profits foregone in the activities that could be performed alternatively. Countervailing incentives also appear when the agent's reservation utility is decreasing with his productivity in the activity performed for the principal (compare, for instance, Lewis and Sappington [7], Maggi and Rodriguez-Clare [8], Brainard and Martimort [1]; Jullien [5]). As an application, Lewis and Sappington [7] describe the relationship between a landowner who endows her farmer with some capital grant at the outset of their agreement, which limits the farmer's incentive to exaggerate production costs. Jullien [5] provides further applications concerning linear and nonlinear pricing.

or not, each type of the first agent is rewarded if some critical type of the second agent is realized and incurs an equal deficit whenever it is not, the critical type being chosen such that the deficit is minimized. As long as agents have a systematic incentive to over-report, the principal picks the same critical type for all types of the first agent, namely the one that is most likely to be drawn by *higher* first agent's types. Hence, whether an agent is rewarded or bears a loss solely depends on the type that is realized for the second agent. This is the result Gary-Bobo and Spiegel [4] obtain in a correlated information setting with systematic incentives to overstate type[3]. When agents display countervailing incentives, this result remains valid for low-marginal-cost agents, who have an incentive to over-report, but it does not for high-marginal-cost agents, who have an incentive to under-report. For the latter, minimizing the deficit requires that the reward be assigned when the second agent's type that is most likely to be drawn by *lower* first agent's types is realized. Therefore, whether an agent is rewarded or bears a loss depends not only on the type that is realized for the second agent but also on the realization of his own type. From now on, we denominate "Maxmin" the incentive scheme that minimizes the loss, using the terminology of Gary-Bobo and Spiegel [4].

Our second result is that the Maxmin contract entails full efficiency if, provided the conditional likelihood function of one agent's type is concave in the other agent's type, the cost function (or, more precisely, the fixed cost) is non-concave in type for either agent. Importantly, the presence of countervailing incentives tightens the condition on costs, which would be less stringent otherwise. This emerges by comparing our result with the literature. In a model similar to that of Gary-Bobo and Spiegel [4] but without limited liability concerns, Riordan and Sappington [10] assess that one simple way to implement first best is to use the same payoff scheme as in Gary-Bobo and Spiegel [4], except that the relevant signal (corresponding to the second agent's type in our model) is not necessarily chosen to minimize the loss. They show that such a scheme does effect first best if the agent's cost function is less concave in type than so is the conditional likelihood function of the relevant signal[4]. Our result differs from that of Riordan and Sappington [10] for the following reason. In a setting where countervailing incentives arise, there exists some intermediary type that displays no incentive to cheat because its incentives

[3] Gary-Bobo and Spiegel [4] consider a situation in which the principal deals with a single limitedly liable agent, whose cost is privately known and correlated with the distribution of publicly observable shocks. In terms of information structure, this is tantamount to having the principal face two agents with correlated private information, as in our model.

[4] This sufficient condition is surely satisfied in the model of Gary-Bobo and Spiegel [4]. Indeed, the latter assume that the conditional likelihood function of the shock is concave in type, whereas the agent's cost function is strictly convex. By doing so, they remove any concern about how cost features can affect the contractual outcome.

to over and under-report perfectly compensate. From the principal's perspective, this is the least efficient type in that it produces at highest *total* costs. Under the Maxmin contract, this type is assigned a payoff equal to zero whatever the other agent's type. Information correlation plays no role in this payoff profile. Because of this, the principal would be unable to extract surplus from other types that were to mimic this particular type. It turns out that, with a concave cost function, the Maxmin scheme leaves all other types with an incentive to actually mimic this particular type.

The next contribution of our research rests on the observation that, under the Maxmin contract, first best can only arise when the penalty it yields is smaller than the maximum sustainable deficit for all types (or, at the limit, it equals that deficit for some types), but this penalty has no relation with the agents' actual liability. Because of this, the principal takes less advantage of type correlation than she would if any larger sustainable deficit were assigned to agents. This aspect is especially relevant in environments with countervailing incentives. Indeed, in the latter, intermediary types are assigned particularly low losses and this contractual offer may become attractive for other types. To circumvent this difficulty, we propose an alternative scheme under which the minimum feasible loss is replaced with the maximum sustainable deficit for all types (including the one that produces at highest total costs). We show that this is the best possible contractual option for the principal as it allows to exploit type correlation at maximum, given the agents' liability. The benefit is that first best is effected under milder condition. While under the Maxmin contract first best arises if the cost function is non-concave in type, a condition that solely reflects the presence of countervailing incentives, under the alternative scheme it arises if the cost function is "not excessively" concave, a condition that is relaxed the more liable agents are and the more type correlation is exploited, given the agents' liability.

When limited liability is sufficiently tight that the minimum feasible loss exceeds the maximum deficit agents can sustain, albeit the relaxed condition described here above is satisfied, first best cannot be (fully) implemented. Our subsequent contribution is to characterize the optimal second-best contract for this case. Our findings show that, once again, the contractual features heavily depend both on the nature of agents' incentives and on the cost characteristics that determine their intensity. Two relevant situations can arise, depending on the curvature of the cost function with respect to type.

If the cost function is less concave in type than some relevant threshold (the first possible situation), then there exist some types for which first best is still effected. This outcome follows from the possibility to exploit type correlation by inflicting (bounded) losses in the presence of countervailing incentives. It does not appear in correlated information environments with systematic incentives to over-report, as represented by Gary-Bobo and Spiegel [4], in which tight limited liability prevents first-best implementation for *any* type. In our model, first best survives for a continuum range of intermediate types neighboring the "least efficient". This is explained by considering that, since such types display weak incentives to cheat, as they are turned between the desire to over-report and that to under-report, the principal does not need to assign large losses (and rewards) to induce them to truthtell. Therefore, as far as intermediate types are concerned, the limits on liability remain irrelevant in the contractual design. The first-best outcome is beyond reach for all remaining types, instead. Yet, this does not mean that all remaining types receive a contract that is similar to the second-best contract in Gary-Bobo and Spiegel [4]. Indeed, in our setting, the second-best contract reflects the circumstance that, moving away from the intermediate types, lower and higher types exhibit increasingly stronger incentives to over and under-report respectively. This involves that, for the types immediately below and above the intermediate ones, the quantity is distorted just enough to retain all surplus and, at the same time, to solicit information release 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 second-best contract in our framework compares with that in Gary-Bobo and Spiegel [4] with sole regards to this last case. More precisely, the similarity concerns the contract designed for low types with intense incentives to over-report. This is so because the countervailing effect is weak for such types, so that the principal faces a (nearly) standard adverse selection problem.

If the cost function is more concave in type than the relevant threshold aforementioned (the second possible situation), then also this cost characteristics, and not only the limits on liability, has an impact on contractual performance. The incentive problem is exacerbated to the point that information is not released unless the principal induces pooling in the contract, *i.e.* an inflexible rule for some given bunch of types. Under this rule, the quantity that is efficient for the type that has no incentive to cheat (the "least efficient" type) is assigned to all types in its neighborhood. Since that type is also the sole from which all surplus is retained *ex ante*, it is the sole for which the first-best outcome is still enforced. Comparing with Lewis and Sappington [6] - [7], it emerges that this incentive scheme is similar in structure to the contract that is optimal in single-agent settings with countervailing incentives when the fixed cost is concave in type.

Further comparing the whole bulk of our results with those obtained in single-agent contexts, we are able to shed light on how the presence of correlation and liability affects the "knife-edge" situation between pooling and separating equilibria when moving from single to multi-agent relationships with countervailing incentives. This is the last contribution of our research. Maggi and Rodriguez-Clare [8] show that, in single-agent settings, the knife-edge situation is represented by the case of linear fixed cost (linear reservation utility, in their model), in which the optimal contract entails pooling and no rent for a range of intermediate types. Pooling is removed as soon as the fixed cost becomes convex. It persists with all types but one getting a rent as soon as the fixed cost becomes concave. According to our results, in multi-agent environments with correlated information, the linear case would still be the knife-edge situation if agents were unable to sustain any deficit *ex post*. With (limitedly) liable agents, it is rather given by the concavity threshold we mentioned to distinguish the two situations that can be realized with tight limited liability. Importantly, we find that this threshold relaxes as the maximum loss agents can bear raises. We thus conclude that the possibility to take advantage of type correlation by inflicting penalties to agents removes pooling in a class of situations in which it would otherwise arise, *i.e.* in contexts where technologies are such that the fixed cost is concave but not too concave in type, and that this class enlarges as agents' pockets become deeper.

The reminder of the paper is organized as follows. In section 2, we present the model. Section 3 focuses on implementation of the first-best outcome. In section 4, we characterize the optimal contract for the case of tight limited liability. Section 5 concludes. Mathematical details are relegated to an Appendix.

2 The model

A risk-neutral principal P contracts with two risk-neutral agents A_1 and A_2. Each agent $i \in \{1, 2\}$ has a distinct task, which is to produce a good in some quantity q_i. As in Lewis and Sappington [6], production costs are given by

$$C_i(q_i; c_i) = c_i q_i + K(c_i). \tag{1}$$

A_i produces at marginal cost c_i and bears the fixed cost $K(c_i)$. The technology is such that the fixed cost depends negatively on c_i, *i.e.* $K'(c_i) < 0$.

At the contracting stage, agent A_i is privately informed about his own type c_i, but he does not know the type c_j of agent A_j, $j \in \{1, 2\}$ and $j \neq i$. Neither type is known to P. The distribution of each c_i, is commonly known to be taken over the support $[\underline{c}, \overline{c}]$. It is also commonly known that types are correlated. The joint density function is $f(c_1, c_2)$ and the cumulative distribution function $F(c_1, c_2) = \int_{\underline{c}}^{c_1} \int_{\underline{c}}^{c_2} f(c_1, c_2)\, dc_1 dc_2$. The marginal density of each c_i is $f_i(c_i) = \int_{\underline{c}}^{\overline{c}} f(c_i, c_j)\, dc_j$ and the conditional distribution function is written $f(c_j | c_i) = f(c_i, c_j) / f_i(c_i)$.

Since the analysis is perfectly symmetric for the two agents, from now on we refer to A_i as the generic agent and, when needed, to A_j, $j \neq i \in \{1, 2\}$, as the other agent.

2.1 The principal's programme

P offers a grand-contract to agents. The Revelation Principle applies. Attention can thus be restricted to direct revelation mechanisms, in which each agent truthfully reports his type. To write the programme formally, we denote $q_i(r_i, r_j)$ and $t_i(r_i, r_j)$ the quantity A_i is to produce and the transfer he has to receive when agents of type c_i and c_j report r_i and r_j respectively. The *ex post* and the *interim* profit of A_i when either agent reports truthfully, are respectively given by

$$\pi_i(c_i, c_j) = t_i(c_i, c_j) - [c_i q_i(c_i, c_j) + K(c_i)] \tag{2a}$$

$$E_{c_j}[\pi_i(c_i, c_j)] \equiv \int_{\underline{c}}^{\overline{c}} \{t_i(c_i, c_j) - c_i q_i(c_i, c_j) - K(c_i)\} f(c_j | c_i)\, dc_j. \tag{2b}$$

Truthful reporting in a Bayesian setting is induced by satisfying, for each agent A_i and each type c_i, the following *interim* incentive constraint

$$E_{c_j}[\pi_i(c_i, c_j)] \geq \int_{\underline{c}}^{\overline{c}} \{t_i(r_i, c_j) - c_i q_i(r_i, c_j) - K(c_i)\} f(c_j | c_i)\, dc_j, \quad \forall r_i, c_i \in [\underline{c}, \overline{c}]. \tag{IC}$$

Besides, P needs to satisfy the *interim* participation constraint

$$E_{c_j}[\pi_i(c_i, c_j)] \geq 0, \ \forall c_i \in [\underline{c}, \overline{c}], \tag{PC}$$

and the *ex post* limited liability constraint

$$\pi_i(c_i, c_j) \geq -L, \ \forall c_i, c_j \in [\underline{c}, \overline{c}], \tag{LL}$$

for some given $L \geq 0$.

Let $S(q_i(c_i, c_j))$ the gross surplus that is generated, under truthful reporting, when A_i provides $q_i(c_i, c_j)$ units of the good, with $S(0) = 0$, $S' > 0$, $S'' < 0$, $S'(0) = +\infty$ and $S'(+\infty) = 0$. P's objective is to achieve the highest attainable level of welfare. The latter is taken to be a weighed sum of P's net surplus, namely

$\sum_{i \neq j} V(q_i(c_i, c_j)) = \sum_{i \neq j} [S(q_i(c_i, c_j)) - t_i(c_i, c_j)]$, and the agents' profits. Formally, P's programme is written as:

$$\underset{\{q_i(c_i,c_j);\pi_i(c_i,c_j)\}}{Max} \widetilde{W}(c_i, c_j) \equiv \sum_{i \neq j} \int_{\underline{c}}^{\overline{c}} \int_{\underline{c}}^{\overline{c}} [V(q_i(c_i, c_j)) + \alpha \pi_i(c_i, c_j)] f(c_i, c_j) dc_j dc_i$$

$$\text{subject to} \qquad (\Gamma)$$

$$\text{(IC), (PC) and (LL),}$$

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

3 First-best implementation

At the first-best outcome (FB hereafter), for all types, the quantity is such that marginal benefit and marginal cost are equal, *i.e.* $S'(q_i^{fb}(c_i)) = c_i$, and surplus is entirely retained *ex ante*, *i.e.* $E_{c_j}\left[\pi_i^{fb}(c_i, c_j)\right] = 0$. We devote this section to explore in which ways and under which conditions P can enforce this outcome.

To make (LL) most likely satisfied, a natural strategy for P is to offer the mechanism that minimizes the largest loss that each type c_i might be assigned (compare Gary-Bobo and Spiegel [4]). We first show how such a mechanism is to be designed in the presence of countervailing incentives and highlight that it is rather requiring in terms of global incentive-compatibility. To circumvent this difficulty, we subsequently characterize an alternative incentive scheme that implements FB under milder conditions.

3.1 The Maxmin scheme with countervailing incentives

Before moving to the analysis, we observe that, under FB implementation, (IC) is conveniently replaced by the pair of conditions

$$q_i^{fb}(c_i) + K'(c_i) - \int_{\underline{c}}^{\overline{c}} \frac{d\pi_i^{fb}(c_i, c_j)}{dc_i} f(c_j | c_i) dc_j = 0 \qquad \text{(LIC)}$$

$$\int_{\underline{c}}^{\overline{c}} \left\{ t_i^{fb}(r_i, c_j) - c_i q_i^{fb}(r_i) - K(c_i) \right\} f(c_j | c_i) dc_j \leq 0, \qquad \text{(GIC)}$$

where $t_i^{fb}(r_i, c_j)$ is the FB transfer P makes to A_i when he reports r_i and A_j reports c_j. (LIC) requires that A_i has no incentive to report $r_i \neq c_i$ in a neighborhood of his true type c_i (*local* incentive compatibility). (GIC) ensures that A_i has no interest in reporting any $r_i \neq c_i$ within the feasible set (*global* incentive compatibility).

Now let $\left\{\pi_i^l(c_i, \underline{c}), ..., \pi_i^l(c_i, \overline{c})\right\}$, with $l \in \mathbb{N}_+$, any feasible profile of *ex post* profits under which FB is implemented. Proceeding similarly to Gary-Bobo and Spiegel [4], within this class, we shall identify the specific profile under which the largest loss that either agent can be required to sustain is minimized. These are the profits under which (LL) is least likely to be binding. For this purpose, we temporarily neglect (GIC) and (LL) and, for each c_i, we solve the following problem:

$$Max \left\{\min\left\{\pi_i^1(c_i,\underline{c}),...,\pi_i^1(c_i,\overline{c})\right\};\ \min\left\{\pi_i^2(c_i,\underline{c}),...,\pi_i^2(c_i,\overline{c})\right\},...\right\}$$
$$\text{subject to} \qquad \text{(Maxmin)}$$
$$\text{(LIC) and (PC).}$$

The solution to (Maxmin) is found in two steps. First, the lowest profit (*i.e.* the highest loss) that A_i might incur over all possible c_j is identified within any profile of first-best profits. Second, within the whole set of lowest profits, the largest one (*i.e.* the minimum loss) is selected. Once this solution is found, it is possible to determine the whole profile of profits to be used for FB implementation. Following Gary-Bobo and Spiegel [4], we name this profile the "Maxmin" scheme. We describe it in the lemmas stated hereafter.

Lemma 1 *Under the Maxmin scheme, for any $c_i \in [\underline{c},\overline{c}]$, A_i is rewarded whenever A_j is of some type $c_{jr}(c_i) \in [\underline{c},\overline{c}]$ and bears the smallest feasible loss whenever $c_j \neq c_{jr}(c_i)$, the loss being equal in size for all $c_j \neq c_{jr}(c_i)$.*

By this lemma, under the Maxmin scheme, for each type c_i, the profile of payoffs $\left\{\pi_i^l(c_i,\underline{c}),\ ...\ ,\pi_i^l(c_i,\overline{c})\right\}$ reduces to only two values. This result is analogous to that of Gary-Bobo and Spiegel [4]. As the latter explain, spreading punishments over as many realizations of c_j as possible (*i.e.*, all feasible realizations but one) allows to minimize the highest possible loss for each type of agent. This requires to minimize the largest reward-loss wedge that can be realized over all possible realizations of c_j.

Based on Lemma 1, for all $c_i \in [\underline{c},\overline{c}]$, the Maxmin pair of profits is found to be

$$\pi_i^l(c_i,c_j)\Big|_{c_j=c_{jr}(c_i)} = \left[q_i^{fb}(c_i)+K'(c_i)\right]\frac{1-f(c_{jr}(c_i)|c_i)}{df(c_{jr}(c_i)|c_i)/dc_i} \equiv \pi_{i,r}^{fb}(c_i) \qquad (3)$$

$$\pi_i^l(c_i,c_j)\Big|_{c_j\neq c_{jr}(c_i)} = \left[q_i^{fb}(c_i)+K'(c_i)\right]\frac{-f(c_{jr}(c_i)|c_i)}{df(c_{jr}(c_i)|c_i)/dc_i} \equiv \pi_{i,p}^{fb}(c_i), \qquad (4)$$

the difference $\left[\pi_{i,r}^{fb}(c_i)-\pi_{i,p}^{fb}(c_i)\right]$ being the lowest feasible wedge when FB is implemented under (LIC) and (PC).

To see how $c_{jr}(c_i)$ should be selected, first notice that this choice depends on the sign of the sum $q_i^{fb}(c_i)+K'(c_i)$, which may not be the same for all $c_i \in [\underline{c},\overline{c}]$. Suppose this sum is positive. Then, for (4) to be a loss ($\pi_{i,p}^{fb}(c_i)<0$), $c_{jr}(c_i)$ must be such that its conditional likelihood raises with c_i, *i.e.* $df\left(c_{jr}\left(c_i\right)|c_i\right)/dc_i > 0$. Moreover, for (4) to be the smallest feasible loss, $c_{jr}\left(c_i\right)$ must be such that the ratio $\frac{f(c_{jr}(c_i)|c_i)}{df(c_{jr}(c_i)|c_i)/dc_i}$ is minimized. This is tantamount to requiring that the ratio $\frac{df(c_{jr}(c_i)|c_i)/dc_i}{f(c_{jr}(c_i)|c_i)}$ be maximized. Observe that the latter ratio is the conditional hazard rate that type $c_j = c_{jr}\left(c_i\right)$ be drawn as c_i increases. This means that, for any given c_i, $c_{jr}\left(c_i\right)$ is to be the value of c_j that *higher* types are most likely to draw. Intuitively, under *interim* break-even, making A_i most likely to be rewarded, given his own type c_i, allows P to contain the deficit that A_i is to sustain for all $c_j \neq c_{jr}\left(c_i\right)$. Similar reasoning applies, *mutatis mutandis*, when $q_i^{fb}\left(c_i\right) + K'\left(c_i\right) < 0$. In that case, $c_{jr}\left(c_i\right)$ must be such that its conditional likelihood decreases with c_i, *i.e.* $df\left(c_{jr}\left(c_i\right)|c_i\right)/dc_i < 0$, and the conditional hazard rate $\frac{df(c_{jr}(c_i)|c_i)/dc_i}{f(c_{jr}(c_i)|c_i)}$ is minimized. That is, for any given c_i, A_i is to be rewarded when the second agent's type that is most likely to be drawn by *lower* possible own types does materialize. The sole situation in which the value of $c_{jr}\left(c_i\right)$ is irrelevant arises when $q_i^{fb}\left(c_i\right) + K'\left(c_i\right) = 0$, in which case both (3) and (4) reduce to zero.

To identify the values of $c_{jr}\left(c_i\right)$ in (3) and (4), it is useful to make the following assumptions.

Assumption 1 $K''\left(c_i\right) < -\frac{dq_i^{fb}(c_i)}{dc_i}, \forall c_i \in \left[\underline{c}, \overline{c}\right].$

Assumption 2 $\frac{df(\overline{c}|c_i)}{dc_i} > 0$, $\frac{d^2 f(\overline{c}|c_i)}{dc_i^2} < 0$, $\frac{df(\underline{c}|c_i)}{dc_i} < 0$, $\frac{d^2 f(\underline{c}|c_i)}{dc_i^2} < 0$, $\forall c_i \in \left[\underline{c}, \overline{c}\right].$

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

$$\frac{d}{dc_i}\left(\frac{f\left(c_j\,|c_i\right)}{f\left(\overline{c}\,|c_i\right)}\right) < 0, \ \forall c_i, c_j \in \left[\underline{c}, \overline{c}\right), \tag{5}$$

$$\frac{d}{dc_i}\left(\frac{f\left(c_j\,|c_i\right)}{f\left(\underline{c}\,|c_i\right)}\right) > 0, \ \forall c_i, c_j \in \left(\underline{c}, \overline{c}\right]. \tag{6}$$

Assumption 1 is equivalent to saying that the sum $q_i^{fb}\left(c_i\right) + K'\left(c_i\right)$ is positive at least at $c_i = \underline{c}$ and decreases with c_i for all its feasible values. Thus, if it decreases fast enough, then there exists some $\widehat{c} \in \left[\underline{c}, \overline{c}\right]$ at which it equals zero. Moreover, it takes negative values $\forall c_i \in \left(\widehat{c}, \overline{c}\right]$. Assumption 2 tells that the probability of drawing type $c_j = \overline{c}$ (resp. $c_j = \underline{c}$) for A_j increases (resp. decreases) with the type of A_i at a decreasing rate. This is thus a requirement on the behaviour of f at two values of c_j, which we take to be $\underline{c}$ and $\overline{c}$. Assumption 3 imposes a pair of monotone likelihood ratio properties of the sort that is standard in the literature. Condition (5) involves that $\frac{df(c_j|c_i)/dc_i}{f(c_j|c_i)} < \frac{df(\overline{c}|c_i)/dc_i}{f(\overline{c}|c_i)}$, *i.e.* the conditional hazard rate that type $c_j = \overline{c}$ be drawn is higher than that of any $c_j \neq \overline{c}$ for all c_i. On the other hand, condition

(6) involves that $\frac{df(\underline{c}|c_i)/dc_i}{f(\underline{c}|c_i)} < \frac{df(c_j|c_i)/dc_i}{f(c_j|c_i)}$, *i.e.* the conditional hazard rate that type $c_j = \underline{c}$ be drawn is lower than that of any $c_j \neq \underline{c}$ for all c_i.

Our previous assumptions allow us to determine the values of $c_{jr}(c_i)$ in (3) and (4), as stated in the lemma hereafter.

Lemma 2 *Suppose there exists $\widehat{c} \in [\underline{c}, \overline{c}]$ such that $q_i^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$. Then, under Assumption 1 - 3, the type $c_{jr}(c_i)$ for which A_i is rewarded takes only two values, namely $c_{jr}(c_i) = \overline{c}$ $\forall c_i \in [\underline{c}, \widehat{c})$ and $c_{jr}(c_i) = \underline{c}$ $\forall c_i \in (\widehat{c}, \overline{c}]$.*

In what follows, we maintain the hypothesis that $\widehat{c}$ actually exists, unless differently specified. Because of Assumption 1, this means that we have $c_{jr}(c_i) = \overline{c}$ for all c_i for which $q_i^{fb}(c_i) + K'(c_i) > 0$ and $c_{jr}(c_i) = \underline{c}$ for all c_i for which $q_i^{fb}(c_i) + K'(c_i) < 0$. Observe that the specific choice of the reward types c_j is without loss of generality in the model. The properties of the likelihood function in Assumption 2 and 3 could refer to any other pair of cost values. Yet, taking the assumption to be satisfied at the extremes of the cost support conveniently warrants that A_i is assigned a loss over a continuum of types c_j, namely $[\underline{c}, \overline{c})$ when $c_i < \widehat{c}$ and $(\underline{c}, \overline{c}]$ when $c_i > \widehat{c}$.

FB implementation in contexts with a continuum of types and signals, like ours, has been studied by McAfee and Reny [9]. They show that (almost) full surplus extraction entails if there exist two signals at which $f(\cdot|c_i)$ displays monotonicity and concavity as it is the case in Assumption 2 (Remark 6 in their paper). This is not a sufficient condition in our model though, because of the need to minimize agents' losses under limited liability.

Furthermore, Riordan and Sappington [10] explore a single-agent framework in which the agent's types are correlated with some observable signals, the space of which is smaller than that of the possible types. With regards to this framework, they show that, under the circumstances described above, P can implement FB using information about a unique signal $c_{jr}(c_i) = c_{jr}$, $\forall c_i$, provided the conditional likelihood function has analogous features to those described in Assumption 2 and 3 at that sole signal. This is what occurs in Gary-Bobo and Spiegel [4], who actually consider a similar setting. By contrast, we focus on situations in which, for either agent, types and signals (which are here the second agent's types) take values on a continuous support. Despite this, the need to satisfy (LL) involves that P refers to fewer signals than available also in our case. However, while in Riordan and Sappington [10] and Gary-Bobo and Spiegel [4] a unique signal suffices, reference to two signals is necessary in our environment with countervailing incentives[5]. For

[5] In particular, Gary-Bobo and Spiegel [4] take the highest possible signal to be most likely drawn from higher types, for all types. In their setting, it suffices to assume the first two conditions in our Assumption 2 together with (5) in our Assumption 3.

this reason, unlike in Gary-Bobo and Spiegel [4], the Maxmin scheme turns out to be "region-specific", *i.e.* it is specifically characterized over different cost ranges.

Proposition 1 *Suppose*

$$L \geq \left[q_i^{fb}(c_i) + K'(c_i)\right] \frac{f(c_{jr}(c_i)|c_i)}{df(c_{jr}(c_i)|c_i)/dc_i}, \ \forall c_i \in [\underline{c}, \overline{c}], \tag{7}$$

with $c_{jr}(c_i) = \overline{c}$ *for* $c_i < \widehat{c}$ *and* $c_{jr}(c_i) = \underline{c}$ *for* $c_i < \widehat{c}$. *Under Assumption 1 - 3, the first-best outcome is implemented with ex post payoffs (3) and (4) if*

$$K"(c_i) \geq 0, \ \forall c_i \in [\underline{c}, \overline{c}]. \tag{8}$$

The conditions reported in Proposition 1 are explained as follows. Condition (7) follows from the circumstance that agents cannot bear unbounded losses. The solution to (Γ) that is picked by the Maxmin scheme does not implement FB unless (7) is satisfied. This condition is similar to that in Proposition 2 of Gary-Bobo and Spiegel [4], although it specifies differently according to whether $c_i < \widehat{c}$ or $c_i > \widehat{c}$. Condition (8) suffices for the Maxmin payoff profile to be globally incentive compatible in (Γ). It requires that the fixed cost function be (weakly) convex in the marginal cost.

Let us illustrate the intuition behind (8). The transfer an agent of type c_i receives when he reports r_i and the other agent reports his true type c_j is given by

$$t_i(r_i, c_j) = r_i q_i^{fb}(r_i) + K(r_i) + \pi_i(r_i, c_j).$$

This transfer is composed of two elements. The first element, namely $r_i q_i^{fb}(r_i) + K(r_i)$, is a fixed payment equal to the total cost the agent would bear if he were of type r_i. The second element, namely $\pi_i(r_i, c_j)$, is an uncertain payment whose value depends on the realization of c_j. Because this realization is unknown to A_i, he faces a lottery with expected value

$$\begin{aligned} \int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|c_i) dc_j &= -\left[\pi_{i,r}^{fb}(r_i) - \pi_{i,p}^{fb}(r_i)\right] [f(c_{jr}(r_i)|r_i) - f(c_{jr}(r_i)|c_i)] \\ &= -\frac{q_i^{fb}(r_i) + K'(r_i)}{df(c_{jr}(r_i)|r_i)/dr_i} [f(c_{jr}(r_i)|r_i) - f(c_{jr}(r_i)|c_i)]. \end{aligned}$$

The introduction of this lottery is meant to offset the benefit A_i 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 for mimicking types. This requires that the wedge between the reward and the loss designed for type r_i, as expressed by the ratio $\frac{q_i^{fb}(r_i)+K'(r_i)}{df(c_{jr}(r_i)|r_i)/dr_i}$, be large enough. Indeed, this allows P to exploit the correlation between types, as represented by the difference $[f(c_{jr}(r_i)|r_i) - f(c_{jr}(r_i)|c_i)]$, to extract surplus. Recall however that, under the Maxmin scheme, the wedge $\pi_{i,r}^{fb}(r_i) - \pi_{i,p}^{fb}(r_i)$ is set at the minimum feasible level for each r_i and, in particular, it equals zero for $r_i = \widehat{c}$. Thus, whenever $\widehat{c}$ is reported, the lottery disappears. Under this circumstance, type $c_i \neq \widehat{c}$ is discouraged from reporting $\widehat{c}$ if

$$\int_{c_i}^{\widehat{c}} [K'(y_i) - K'(\widehat{c})]\, dy_i \leq 0,\ \forall c_i \in [\underline{c}, \overline{c}],$$

which explains (8).

The literature has shown that some restriction on the properties of the cost function is required for FB implementation also in the absence of countervailing incentives. From Riordan and Sappington [10], we learn that, when the signal space is smaller than the type space, together with the conditions on the likelihood function of the relevant signal (Assumption 2 and 3), FB enforcement calls for restrictions on the shape of the agent's cost function. It is thus not surprising that a lower bound on the concavity of K appears also in our setting, where P artificially reduces the space of relevant signals (the second agent's types) in the seek for an incentive scheme that imposes as small penalties as feasible. If agents were not protected by limited liability, then P would not need to use the specific *ex post* payoffs that belong to the Maxmin scheme and Remark 6 of McAfee and Reny [9], which we mentioned above, would apply. That is, Assumption 2 would ensure that P could find a payoff profile that extracts surplus (almost) entirely, independently of the properties of the agents' cost function.

The restriction imposed by (8) on the fixed cost function is tighter than the condition identified by Riordan and Sappington [10]. The latter only requires that the agent's cost function be less concave in type than the conditional likelihood function at the relevant signal. As stated in the corollary below, a similar result would entail in our model only in case agents were to display a systematic incentive either to overstate or to understate type, whatever the cost realization.

Corollary 1 *If there exists no* $\widehat{c} \in [\underline{c}, \overline{c}]$ *such that* $q_i^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$*, then,* $\forall c_i \in [\underline{c}, \overline{c}]$, *(8) is replaced by*

$$K"(c_i) \geq \left[q_i^{fb}(c_i) + K'(c_i)\right] \frac{d^2 f(c_{jr}(c_i)|c_i)/dc_i^2}{df(c_{jr}(c_i)|c_i)/dc_i}, \tag{9}$$

with $c_{jr}(c_i) = \overline{c}$ *when* $q_i^{fb}(c_i) + K'(c_i) > 0$ *and* $c_{jr}(c_i) = \underline{c}$ *when* $q_i^{fb}(c_i) + K'(c_i) < 0$.

The corollary above emphasizes that the presence of countervailing incentives exacerbates the requirement on the properties of the cost function.

It should by now be clear that the concavity restriction appears because P adopts the payoff profile that allows her to minimize the largest potential deficit. As already illustrated, when this scheme is designed for agents displaying countervailing incentives, the lottery tends to vanish as r_i approaches $\widehat{c}$. To circumvent this problem, P should design a different mechanism under which (1) all agent's types do face an effective lottery whatever their report and (2) (LL) is still satisfied. We hereafter describe how a scheme with these characteristics can be constructed.

3.2 An alternative scheme

Suppose (8) holds for all possible types, meaning that limited liability does not (necessarily) compromise FB implementation. As already explained, to induce information release at no agency cost, the expected value of the lottery is to be low enough.

Lemma 3 *The ex post profits that minimize the expected value of the lottery and, at the same time, satisfy (LIC), (PC) and (LL) are such that A_i is rewarded for one sole realization of $c_j \in [\underline{c}, \overline{c}]$, $\forall i \neq j \in \{1, 2\}$, and incurs the highest admissible loss $(-L)$ for all other realizations.*

The scheme presented in the lemma is similar to the Maxmin scheme in that it includes only one reward and equal losses. Yet, losses are here fixed at the largest feasible level so as to minimize the incentive to misreport type for any given c_i. Taken together, Lemma 1 and 3 evidence that spreading losses over as many realizations of c_j as possible is beneficial to P in two different ways. First, when the Maxmin scheme is adopted, spreading losses and minimizing the reward-loss wedge for the type c_j at which this wedge is maximum allow P to minimize the largest deficit that agents could be required to incur. Second, under the alternative scheme, spreading losses and maximizing the reward-loss wedge for *each* possible realization of c_j allow P to minimize the expected value of the lottery that agents are called to face.

Assumption 3 ensures that the type c_j for which A_i is rewarded under the alternative scheme remains the same as under the Maxmin scheme, namely $c_{jr}(c_i) = \overline{c}$ if $c_i < \widehat{c}$ and $c_{jr}(c_i) = \underline{c}$ if $c_i > \widehat{c}$. The payoff profile is thus written

$$\overline{\pi}_{i,r}^{fb}(c_i, L) = \frac{1 - f(c_{jr}(c_i) | c_i)}{f(c_{jr}(c_i) | c_i)} L, \ \forall c_i \in [\underline{c}, \overline{c}] \quad (10)$$

$$\overline{\pi}_{i,p}^{fb}(c_i, L) = -L, \ \forall c_i \in [\underline{c}, \overline{c}], \ \forall c_j \in [\underline{c}, \overline{c}], \ c_j \neq c_{jr}(c_i), \quad (11)$$

so that the expected value of the lottery is given by

$$\int_{\underline{c}}^{\overline{c}} \pi_i (r_i, c_j) f (c_j |c_i) dc_j = -L \left[1 - \frac{f (c_{jr} (r_i) |c_i)}{f (c_{jr} (r_i) |r_i)}\right].$$

This lottery is actually more effective at extracting surplus from agents, as compared to the one associated with the Maxmin scheme, because it allows P to take better advantage of type correlation. As a result, FB is enforced under milder conditions.

Proposition 2 *Suppose condition (7) holds. Under Assumption 1 - 3, the first-best outcome is implemented with ex post payoffs (10) and (11) if,* $\forall c_i \in [\underline{c}, \overline{c}]$ *and* $\forall i \neq j \in \{1, 2\}$,

$$\begin{aligned} K" (c_i) \;\; &\geq \;\; \left[q_i^{fb} (c_i) + K' (c_i) - \frac{df (c_{jr} (c_i) |c_i) /dc_i}{f (c_{jr} (c_i) |c_i)} L\right] \\ &+ \frac{d^2 f (c_{jr} (c_i) |c_i) /dc_i^2}{f (c_{jr} (c_i) |c_i)} L, \end{aligned} \tag{12}$$

with $c_{jr} (c_i) = \overline{c}$ *for* $c_i < \widehat{c}$ *and* $c_{jr} (c_i) = \underline{c}$ *for* $c_i < \widehat{c}$.

This proposition states that, whenever the lowest loss that is compatible with FB implementation is smaller than the largest deficit agents can sustain, FB is enforced with payoffs (10) and (11) provided that K'' does not fall below the lower bound imposed by (12). Condition (7) and Assumption 2 ensure that this bound is negative, showing that the requirement on the curvature of K is now relaxed as compared to (8).

A clear message ensues from our analysis. FB is at hand also when K is concave, provided that P is available to abandon the Maxmin scheme and opt for a mechanism that possibly inflicts more important (though still feasible) penalties to agents. Gary-Bobo and Spiegel [4] emphasize that resorting to the Maxmin scheme, rather than offering a payoff profile that entails larger deficits, can be especially convenient for a principal. Typically, regulators prefer to minimize the financial difficulties of the regulated firms both to avoid activity interruptions and because this is embarrassing for themselves. Our investigation evidences that, when the Maxmin scheme is adopted in environments with countervailing incentives, the loss it yields might result excessively low, on the opposite. This is actually the case when types are very intensely turned between the desire to over-report and that to under-report, *i.e.* when condition (8) is not met.

In the sequel of the analysis, we take (12) to be satisfied. We thus neglect the possibility that FB does not attain because fixed costs are too concave. We rather focus on the more interesting case in which FB implementation is beyond reach because the limits on liability are particularly stringent.

4 The optimal contract with tight limited liability

In this section, we explore the situation in which (LL) is so tight that condition (7) in Proposition 1 fails to hold. Under this circumstance, P cannot find a profit profile that implements FB. She thus designs a second-best (SB hereafter) contract, which is to be characterized in the sequel of the analysis.

To begin with, notice that, in fact, (7) is not violated for all feasible values of c_i.

Lemma 4 *Under Assumption 1 - 3, for any* $L \geq 0$, *at the solution to* (Γ), *there exists a unique range of types* $[c_{i2}, c_{i3}] \subseteq [\underline{c}, \overline{c}]$, *such that* $\widehat{c} \in [c_{i2}, c_{i3}]$, *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, (7) is surely satisfied since $q_i^{fb}(\widehat{c}) + K'(\widehat{c}) = 0$. Furthermore, (7) holds for the types that lie in a neighborhood of $\widehat{c}$, *i.e.* for all values of c_i for which the absolute value of $q_i^{fb}(c_i) + K'(c_i)$ is sufficiently low. (LL) is more and more likely to be binding as c_i diverges from $\widehat{c}$.

Lemma 5 *Under Assumption 1 - 3, there exists (at most) one cost value* $c_{i1} \in (\underline{c}, c_{i2})$ *(resp.* $c_{i4} \in (c_{i3}, \overline{c})$*) such that, at the solution to* (Γ), (PC) *is slack* $\forall c_i \in [\underline{c}, c_{i1})$ *(resp.* $\forall c_i \in (c_{i4}, \overline{c}]$*) and binding* $\forall c_i \in [c_{i1}, c_{i2}]$, *(resp.* $\forall c_i \in [c_{i3}, c_{i4}]$*). When no such cost value exists,* (PC) *is slack* $\forall c_i \in [\underline{c}, c_{i2})$ *(resp.* $\forall c_i \in (c_{i3}, \overline{c}]$*).*

At the solution to (Γ) under tight limited liability, not only P enforces FB for all types in $[c_{i2}, c_{i3}]$. She is also able to extract all surplus from some types below c_{i2} and some types above c_{i3}[6]. In Appendix, we show that the range of types below c_{i2} (resp. above c_{i3}) from which surplus is fully retained spans to the whole set $[\underline{c}, c_{i2}]$ (resp. $[c_{i3}, \overline{c}]$) if the likelihood that the highest (resp. lowest) possible value of c_j be drawn, conditional of c_i taking that same value, raises enough with c_i, *i.e.* if $\left.\frac{df(\overline{c}|c_i)}{dc_i}\right|_{\underline{c}}$ (resp. $\left.\frac{df(\underline{c}|c_i)}{dc_i}\right|_{\overline{c}}$) is sufficiently large (resp. small). In the converse case, surplus extraction is feasible only for sufficiently high (resp. low) cost values. All remaining types are assigned a positive *interim* payoff (*i.e.* an information rent).

The outcome above compares with the SB solution in Gary-Bobo and Spiegel [4]. They show that a sufficient condition for the agent's participation constraint to hold as an equality for the least efficient type and strictly for all other types is that

[6] Although surplus is retained, FB is not implemented for these types because quantities are distorted away from the efficient level, as will become clear in the sequel of the analysis.

the derivative of the conditional likelihood function of the highest possible signal be small enough[7]. In our framework, a milder condition is required on the rate of change of the conditional f. Indeed, this is only relevant at the extreme values of c_i, as previously explained. Moreover, the possibility that (PC) be slack for all types but one is ruled out, independently of the magnitude of that rate of change. The divergence of our finding from that of Gary-Bobo and Spiegel [4] is due to the presence of fixed costs that decrease with type, which weakens the incentives to cheat of types sufficiently close to $\widehat{c}$ and thus facilitates surplus extraction. Yet, the properties of the function f, jointly with those of the function K, determine how wide is the range of types that get no rent. On one side, all else equal, P retains surplus more easily when the probability of drawing the highest (resp. lowest) value for c_j increases (resp. decreases) much with c_i, *i.e.* when agents' types are especially informative signals. On the other side, P is more likely to induce truthtelling at zero rent when the fixed cost function is less concave[8]. This is in line with our previous conclusion that cost convexity facilitates P's task, which will be further confirmed in Proposition 3 below.

Suppose the aforementioned requirements about the conditional likelihood are met indeed, so that the cost values c_{i1} and c_{i4} do exist. We shall now see how the SB output is characterized in this situation. Consider that the incentive to overstate (resp. understate) type that an agent with $c_i < \widehat{c}$ (resp. $c_i > \widehat{c}$) would display if he were to receive the sole fixed payment to produce the FB quantity gets increasingly more intense as c_i approaches $\underline{c}$ (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 rewards and losses as c_i moves away from $\widehat{c}$. Nevertheless, (LL) imposes a bound on how large losses can be set, for FB does not attain when $c_i \notin [c_{i2}, c_{i3}]$. Without quantity distortions, P could solicit information revelation only by raising the reward sufficiently, which would yield an information rent to agents. 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 $[c_{i1}, c_{i2})$ and $(c_{i3}, c_{i4}]$, P distorts output till all surplus is extracted. This further clarifies why, over these cost ranges, participation constraints are saturated, as we said above. For types with more intense incentives to misreport, namely those in $[\underline{c}, c_{i1}]$ and $[c_{i4}, \overline{c}]$, P distorts output to contain the rent, but it would be too costly to remove the latter entirely.

[7] Compare page 5 of the technical appendix in Gary-Bobo and Spiegel [4].

[8] One can check that the more (resp. less) negative K' is, the higher $\left.\frac{df(\overline{c}|c_i)}{dc_i}\right|_{c_i=\underline{c}}$ (resp. the lower $\left.\frac{df(\underline{c}|c_i)}{dc_i}\right|_{c_i=\overline{c}}$) is to be for (PC) to be binding.

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 4 *The conditional likelihood and cumulative distribution function satisfy the following properties:*

$$\frac{d}{dc_i}\left(\frac{F(c_i|\overline{c})}{f(c_i|\overline{c})}\right) \geq 0, \ \forall c_i \in [\underline{c},\overline{c}], \ \forall i \in \{1,2\} \tag{13}$$

$$\frac{d}{dc_i}\left(\frac{1-F(c_i|\underline{c})}{f(c_i|\underline{c})}\right) \leq 0, \ \forall c_i \in [\underline{c},\overline{c}], \ \forall i \in \{1,2\}. \tag{14}$$

This assumption states the monotonicity of the conditional hazard rates $\frac{F(c_i|\overline{c})}{f(c_i|\overline{c})}$ and $\frac{1-F(c_i|\underline{c})}{f(c_i|\underline{c})}$ with respect to c_i. According to (13), once types between $\underline{c}$ and c_i have been drawn, it becomes more and more likely that a less efficient type is drawn for A_i as c_i raises, provided that the highest marginal cost is drawn for A_j. According to (14), once types between c_i and $\overline{c}$ have been drawn, it is less and less likely that a less efficient type be drawn for A_i as c_i raises, provided that the lowest marginal cost is drawn for A_j.

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

Lemma 6 *Suppose condition (7) does not hold. Under Assumption 1 - 4, at the solution to* (Γ)*, quantities are characterized as follows:*

$$S'\left(q_i^I(c_i)\right) = c_i + (1-\alpha)\frac{F(c_i|\overline{c})}{f(c_i|\overline{c})}, \ \forall c_i \in [\underline{c}, c_{i1}] \tag{15}$$

$$q_i^{II}(c_i) = \frac{df(\overline{c}|c_i)/dc_i}{f(\overline{c}|c_i)}L - K'(c_i), \ \forall c_i \in [c_{i1}, c_{i2}] \tag{16}$$

$$q_i^{III}(c_i) = q_i^{fb}(c_i), \ \forall c_i \in [c_{i2}, c_{i3}] \tag{17}$$

$$q_i^{IV}(c_i) = \frac{df(\underline{c}|c_i)/dc_i}{f(\underline{c}|c_i)}L - K'(c_i), \ \forall c_i \in [c_{i3}, c_{i4}] \tag{18}$$

$$S'\left(q_i^V(c_i)\right) = c_i - (1-\alpha)\frac{1-F(c_i|\underline{c})}{f(c_i|\underline{c})}, \ \forall c_i \in [c_{i4}, \overline{c}]. \tag{19}$$

Moreover, interim rents are given by

$$E_{c_j}\left[\pi_i^I(c_i,c_j)\right]=f(\overline{c}\,|c_i)\int_{c_i}^{c_{i1}}\frac{q_i^I(y_i)+K'(y_i)}{f(\overline{c}|y_i)}dy_i-\left[1-\frac{f(\overline{c}|c_i)}{f(\overline{c}|c_{i1})}\right]L,\quad \forall c_i\in[\underline{c},c_{i1}] \tag{20}$$

$$E_{c_j}\left[\pi_i^k(c_i,c_j)\right]=0,\ \forall c_i\in[c_{i1},c_{i2}],[c_{i2},c_{i3}],[c_{i3},c_{i4}],\quad \forall k\in\{II,III,IV\} \tag{21}$$

$$E_{c_j}\left[\pi_i^V(c_i,c_j)\right]=-f(\underline{c}\,|c_i)\int_{c_{i4}}^{c_i}\frac{q_i^V(y_i)+K'(y_i)}{f(\underline{c}|y_i)}dy_i-\left[1-\frac{f(\underline{c}|c_i)}{f(\underline{c}|c_{i4})}\right]L,\quad \forall c_i\in[c_{i4},\overline{c}]. \tag{22}$$

To begin with, (17) confirms that output is still efficiently set as long as $c_i\in[c_{i2},c_{i3}]$. According to (15) and (19), the same occurs at both the lowest and the highest marginal cost realization. (15) further highlights that output is downward distorted for all types in $(\underline{c},c_{i1}]$, which allows to contain the rent in (20). Moreover, under the first part of Assumption 4, q_i^I decreases with c_i all over this set. Observe that the SB quantity solution in Gary-Bobo and Spiegel [4] is characterized precisely as in (15) for all possible agent's types. This occurs because, in their context, as in any standard adverse selection problem, the agent displays a systematic incentive to overstate type. (19) further evidences that output is upward distorted for all types in $[c_{i4},\overline{c})$, which helps limit the rent in (22). Under the second part of Assumption 4, also q_i^V decreases with type $\forall c_i\in[c_{i4},\overline{c})$. Lastly, (16) and (18) 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 the lemmas above but not yet characterized.

Lemma 7 *Suppose condition (7) does not hold. Under Assumption 1 - 4, at the solution to* (Γ), *the cost values* c_{i1}, c_{i2}, c_{i3} *and* c_{i4}, *are defined as follows:*

$$q_i^I(c_{i1})+K'(c_{i1})=\frac{df(\overline{c}\,|c_{i1})/dc_{i1}}{f(\overline{c}\,|c_{i1})}L \tag{23}$$

$$q_i^{II}(c_{i2})=q_i^{fb}(c_{i2}) \tag{24}$$

$$q_i^{IV}(c_{i3})=q_i^{fb}(c_{i3}) \tag{25}$$

$$q_i^V(c_{i4})+K'(c_{i4})=\frac{df(\underline{c}\,|c_{i4})/dc_{i4}}{f(\underline{c}\,|c_{i4})}L. \tag{26}$$

Interpreting Lemma 7 together with the results previously presented, it should be clear that c_{i1} is the cost value at which P retains all surplus from A_i by sufficiently deflating output q_i^I below the FB level, c_{i2} is the value at which P retains all surplus by keeping output q_i^{II} at the FB level and similarly for c_{i3} and c_{i4}.

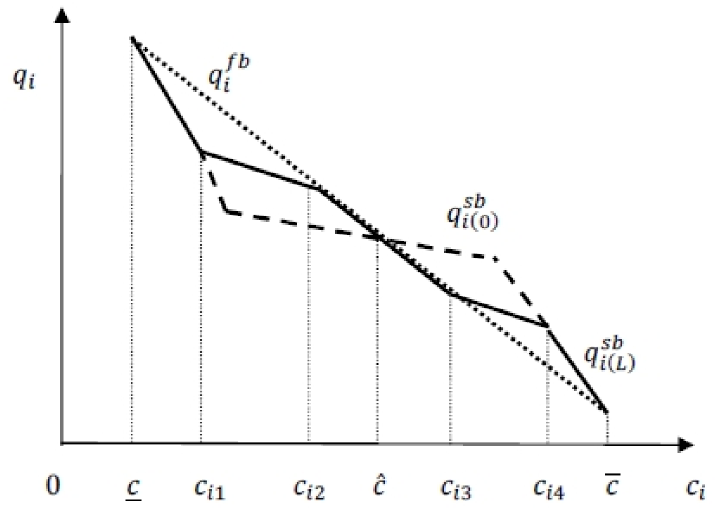

Figure 1: The FB output profile (q_i^{fb}; dotted line) and the output profile in the SB contract with tight limited liability for $L > 0$ ($q_{i(L)}^{sb}$; thick line) and $L = 0$ ($q_{i(0)}^{sb}$; dashed line).

A graphical illustration of the full profile of quantities is provided in Figure 1 with regards to both FB implementation and the SB contract defined by (15) to (19). 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 and would collapse onto the singleton $\{\widehat{c}\}$ in the extreme case in which $L = 0$. The graph further shows that the SB quantity decreases with c_i all over the support, *i.e.* $\frac{dq_i^k(c_i)}{dc_i} \leq 0$ $\forall k \in \{I, II, III, IV, V\}$, $\forall c_i \in [\underline{c}, \overline{c}]$, with a rate of decrease that is specific to each cost interval[9]. In particular, it is $\frac{dq_i^I(c_i)}{dc_i} < \frac{dq_i^{fb}(c_i)}{dc_i} < \frac{dq_i^{II}(c_i)}{dc_i}$ and $\frac{dq_i^V(c_i)}{dc_i} < \frac{dq_i^{fb}(c_i)}{dc_i} < \frac{dq_i^{IV}(c_i)}{dc_i}$.

The following proposition lists the conditions under which the SB solution previously characterized is globally incentive compatible.

Proposition 3 *Suppose condition (7) does not hold. Under Assumption 1 - 4, the quantity profile (15) - (19) is implemented as the solution to* (Γ) *if*

$$\frac{dq_i^I(c_i)}{dc_i} \leq -\left[q_i^I(c_i) + K'(c_i)\right] \frac{df(\overline{c}|c_i)/dc_i}{f(\overline{c}|c_i)}, \ \forall c_i \in [\underline{c}, c_{i1}] \tag{27}$$

$$\frac{dq_i^V(c_i)}{dc_i} \leq -\left[q_i^V(c_i) + K'(c_i)\right] \frac{df(\underline{c}|c_i)/dc_i}{f(\underline{c}|c_i)}, \ \forall c_i \in [c_{i4}, \overline{c}] \tag{28}$$

$$K\text{''}(c_i) \geq \frac{d^2 f(c_{jr}(c_i)|c_i)/dc_i^2}{f(c_{jr}(c_i)|c_i)} L, \ \forall c_i \in [\underline{c}, \overline{c}], \tag{29}$$

with $c_{jr}(c_i) = \overline{c}$ *for* $c_i < \widehat{c}$ *and* $c_{jr}(c_i) = \underline{c}$ *for* $c_i < \widehat{c}$.

We have previously explained that, under Assumption 4, quantities q_i^I and q_i^V decrease with type. Proposition 3 further evidence that, for the contract presented

[9] That (16) and (18) decrease with c_i is ensured by condition (29) in Proposition 3 below.

in Lemma 6 to be globally incentive compatible, it suffices that those quantities decrease sufficiently fast over the respective cost ranges (see Figure 1 again). According to condition (27) and (28), how fast q_i^I and q_i^V should decrease depends on 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 (27), keeping in mind that analogous reasoning applies to (28), *mutatis mutandis*. Take $c_i \in [\underline{c}, c_{i1})$. As the report r_i is raised above the true type c_i, under Assumption 2, the probability of reward increases. Since the loss that A_i might bear equals $-L$ whatever the report, over-reporting yields a higher *interim* payoff, as compared to truthtelling, unless the quantity is diminished sufficiently. The incentive to over-report is removed if $q_i^I(c_i)$ decreases as fast as (27) dictates. Perfectly analogous to (27) would be the sufficient condition for global incentive compatibility in Gary-Bobo and Spiegel [4] 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[10].

Condition (29) tells that the contract described in Lemma 4 - 6 is optimal if, for all possible types, the curvature of the fixed cost function does not fall below some given bound that depends on both the conditional likelihood and L. In fact, (29) is the counterpart of (12) in the FB framework previously explored and can be interpreted in a similar fashion, *mutatis mutandis*. Yet, (29) is more requiring as compared to (12). This further reflects the circumstance that, all else equal, it is harder to induce information release when the limits on agents' liability are stringent.

Corollary 2 *Take $L = 0$ and $K" = 0$. Suppose condition (7) does not hold, whereas (27) and (28) are satisfied. At the solution to (Γ), $q_i^{sb}(c_i) = q_i^{fb}(\widehat{c})$, $\forall c_i \in [c_{i1}, c_{i4}]$.*

The corollary refers to the specific situation in which agents can bear no deficit *ex post* and fixed costs are linear in type. In that case, the range of types for which FB is enforced collapses onto the singleton $\{\widehat{c}\}$. To see this, recall that c_{i2} and c_{i3} are defined by $q_i^{II}(c_{i2}) = q_i^{fb}(c_{i2})$ and $q_i^{II}(c_{i3}) = q_i^{fb}(c_{i3})$ respectively. Moreover, with $L = 0$, $q_i^{II}(c_i) = -K'(c_i)$ and $q_i^{IV}(c_i) = -K'(c_i)$. Remembering also the definition of $\widehat{c}$, it is immediate to conclude that $c_{i2} \equiv \widehat{c} \equiv c_{i3}$ when $L = 0$. Further observe that quantities $q_i^{II}(c_i)$ and $q_i^{IV}(c_i)$ are constant over types when so is $K'(c_i)$. Hence, all types within the set $[c_{i1}, c_{i4}]$, from which surplus is entirely extracted, are required to produce the same amount of output, *i.e.* the optimal contract entails pooling at $q_i^{fb}(\widehat{c})$ in a neighborhood of $\widehat{c}$.

[10] Compare the inequality at the end of page 5 in the technical appendix of Gary-Bobo and Spiegel [4] with (71) in the proof of (27) in our appendix.

The outcome in Corollary 2 is reminiscent of that Maggi and Rodriguez-Clare [8] find in a single-agent setting. 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[11]. The case of $K" = 0$ in our model is the counterpart for the linear reservation utility in Maggi and Rodriguez-Clare [8]. Corollary 2 evidences that, when the case of $K" = 0$ arises, the optimal contract exhibits analogous features (namely, pooling and no rent in a neighborhood of $\widehat{c}$) in the multi-agent framework as soon as agents cannot be punished *ex post*. This is explained by considering that having $L = 0$ with more than one agent is tantamount to assuming that agents are to break even *ex post* (rather than only at *interim*), while *ex post* and *interim* participation are equivalent in the single-agent setting without correlated information. Observe however that, despite the analogy in terms of structure, the optimal multi-agent contract is not simply twice a replica of the single-agent contract. Indeed, improvements are available, as usual with correlated types. First, the range of types for which bunching arises is less wide. Second, the quantity distortions induced for low and high types are smaller. Third, the expected rents that accrue to those same types are lower than the rents P assigns when she deals with a single agent.

4.1 "Very concave" fixed costs

As previously said, condition (12) in Proposition 2 is taken to be satisfied all along the analysis. Even under this assumption, it is not necessarily the case that condition (29) in Proposition 3 is satisfied in turn. In what follows, we consider the situation in which (29) is violated. The following proposition describes the optimal contract under this circumstance.

Proposition 4 *Suppose neither condition (7) nor condition (29) holds. Under Assumption 1 - 4, the quantity solution to* (Γ) *is given by*

$$\begin{aligned} q_i^{sb}(c_i) &= q_i^{I}(c_i), \ \forall c_i \in \left[\underline{c}, c_i^{-}\right) \\ q_i^{sb}(c_i) &= q_i^{fb}(\widehat{c}), \ \ \forall c_i \in \left[c_i^{-}, c_i^{+}\right] \\ q_i^{sb}(c_i) &= q_i^{V}(c_i), \ \forall c_i \in \left(c_i^{+}, \overline{c}\right], \end{aligned}$$

[11] An environment with reservation utility linear in type is analysed also in other works, such as that of Brainard and Martimort [1], with analogous result.

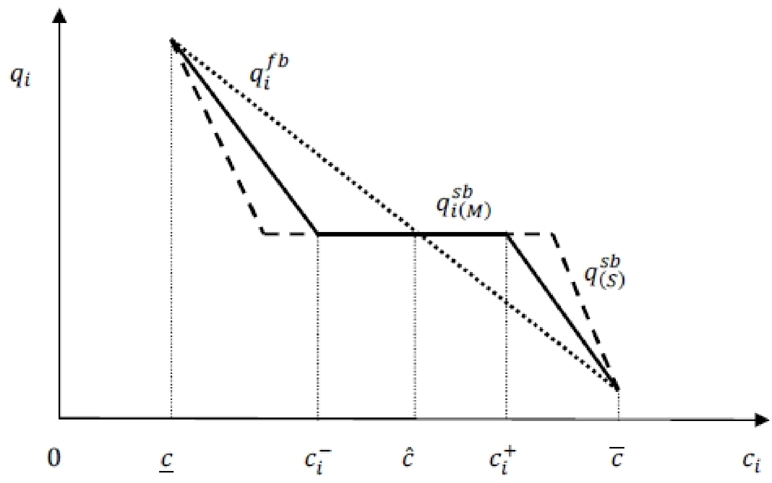

Figure 2: The FB output profile (q_i^{fb}; dotted line), the output profile in the SB $m$ulti-agent contract ($q_{i(M)}^{sb}$; thick line) and the output profile in the $s$ingle-agent contract ($q_{(S)}^{sb}$; dashed line) with K "very" concave.

where c_i^- and c_i^+ are such that

$$\begin{aligned} q_i^I\left(c_i^-\right) &= q_i^{fb}\left(\widehat{c}\right) \\ q_i^V\left(c_i^+\right) &= q_i^{fb}\left(\widehat{c}\right). \end{aligned}$$

Moreover, (PC) is binding only for type $\widehat{c}$.

The contract described in the proposition entails a pooling equilibrium in a neighborhood of $\widehat{c}$. This is the sole type from which P is able to retain all surplus when (29) is violated. The contract is reminiscent of that characterized by Lewis and Sappington [6]. They study countervailing incentives in a single-agent setting without correlated signals, focusing on the case in which the agent's fixed cost function is concave in type[12]. Yet, in our environment, pooling concerns a smaller range of types, a benefit that follows from the presence of information correlation. This is shown by the graph in Figure 2, which compares the optimal output profile in the two situations.

Having (29) violated means that, as long as $L > 0$, fixed costs must be sufficiently concave in type for the contract illustrated in Proposition 4 to be SB optimal. By contrast, the single-agent optimal contract exhibits the structure aforementioned even with slightly concave fixed costs. This shows that, whenever losses can be inflicted to agents *ex post*, the presence of information correlation yields an additional benefit. That is, it also enlarges the class of environments in which separating

[12] Maggi and Rodriguez-Clare [8] obtain the same outcome in the equivalent situation in which the agent's reservation utility is concave in type.

equilibria arise.

Further observe that the structure of the contract in Proposition 4 is similar to that in Corollary 2, except that, in the former, the range of types with no rent degenerates onto a singleton. This follows from the circumstance that, as already illustrated, incentives to over/under-report are especially strong when K is very concave. In that case, information release is not induced unless even the types around $\widehat{c}$ are given up a rent. To interpret this point in a unified way with the rest of our SB results, it is useful to recall that, in single-agent relationships, the linear fixed costs (or, equivalently, linear reservation utility) case with pooling and no rent for some type range can be seen as a "knife-edge" situation: pooling is removed as soon as K becomes convex; all types but one obtain a rent as soon as K becomes concave (compare Maggi and Rodriguez-Clare [8]). From our analysis, it emerges that the linear case remains a "knife-edge" situation in multi-agent hierarchies insofar as *ex post* deficits are unfeasible (recall the explanation after Corollary 2). In correlated information frameworks, the relevant "knife-edge" situation becomes condition (29) as soon as agents can be exposed to (bounded) losses under *interim* participation.

5 Concluding remarks

In this article, we have studied centralized contracting between a principal and two agents who have countervailing incentives to misreport their types, which are correlated. We have focused on the realistic case in which agents are limitedly liable. As an example of the situations we have represented, one may consider centralized regulation of utilities such as electricity and water and sewage.

Our analysis predicts that, as long as agents' pockets are sufficiently deep, the first-best outcome is implemented by the incentive scheme that yields the smallest feasible *ex post* loss to agents (the Maxmin scheme), if the latter's fixed costs are either linear or convex in type. However, the first-best outcome is unfeasible if the agents' technology does not display this property, unless the principal offers a contract that imposes higher deficit to agents. We show that, in the presence of countervailing incentives, the contract that yields the highest sustainable loss to agents expands at maximum the range of cost functions that support first best.

Our analysis further predicts that, if the agents' fixed costs are not very concave in type (so that incentives to over and under-report are not too intense), the optimal incentive scheme is a separating contract under which, thanks to the presence of countervailing incentives, the first-best outcome can still be effected for some range of types even when agents do not have especially deep pockets. Otherwise, the optimal contract entails pooling of quantities. However, the concavity threshold between separating and pooling contracts does depend on the agents' liability. As the latter raises (though not to the point that first best can be enforced for any type), increasingly more concave cost functions, *i.e.* a wider class of possible technologies, sustain the separating contract.

The results of our study provide further scope for resorting to centralized incentive schemes in correlated information settings in which agents display countervailing incentives and can be exposed to some deficit *ex post*. This seems to be especially relevant with regards to contexts, such as energy sectors, in which there is a rich variety of technologies that can be utilized for activity. Our findings further suggest that, when the characteristics of the technologies agents use make their incentives to lie especially strong, improving contractual efficiency may require to transfer as much uncertainty as feasible to agents and, hence, to raise their losses. This contrasts with the usual attitude of regulators not to aggravate the financial burden of the regulated firms.

All along our work, we have taken agents to behave non-cooperatively. The analysis could be extended to assess how collusion would affect the principal's strategy and achievements. This would be useful because, in some cases, collusion represents a concern for regulators in the industrial contexts our model stylizes.

References

[1] 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

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

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

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

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

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

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

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

[9] McAfee, R.P., and P.J. Reny (1992), "Correlated Information and Mechanism Design", *Econometrica*, 60(2), 395-421

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

[11] Vinella, A. (2008), "Countervailing incentives in multi-agent environments", *mimeo*

A First-best implementation

A.1 Local incentive constraint (LIC)

Let $\widetilde{\pi}_i(r_i, c_j)$ the *ex post* profit of A_i when he has type c_i and reports r_i. His *interim* profit is written

$$\begin{aligned} E_{c_j}\left[\widetilde{\pi}_i(r_i, c_j)\right] &\equiv \int_{\underline{c}}^{\overline{c}} \widetilde{\pi}_i(r_i, c_j) f(c_j | c_i) \, dc_j \\ &= \int_{\underline{c}}^{\overline{c}} \{t_i(r_i, c_j) - c_i q_i(r_i, c_j) - K(c_i)\} f(c_j | c_i) \, dc_j. \end{aligned} \quad (30)$$

From (30), the first order-condition of A_i's programme, evaluated at $r_i = c_i$, is given by

$$\int_{\underline{c}}^{\overline{c}} \left[\frac{dt_i(c_i, c_j)}{dc_i} - c_i \frac{dq_i(c_i, c_j)}{dc_i}\right] f(c_j | c_i) \, dc_j = 0. \quad (31)$$

From $(2a)$, we can compute

$$\frac{dt_i(c_i, c_j)}{dc_i} = \frac{d\pi_i(c_i, c_j)}{dc_i} + c_i \frac{dq_i(c_i, c_j)}{dc_i} + q_i(c_i, c_j) + K'(c_i). \quad (32)$$

Replacing (32) into (31), we have

$$\int_{\underline{c}}^{\overline{c}} \left\{\frac{d\pi_i(c_i, c_j)}{dc_i} + [q_i(c_i, c_j) + K'(c_i)]\right\} f(c_j | c_i) \, dc_j = 0. \quad (33)$$

Since (PC) is binding for all c_i at FB, so that

$$\int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) f(c_j | c_i) \, dc_j = 0, \quad (34)$$

both $E_{c_j}[\pi_i(c_i, c_j)]$ and $dE_{c_j}[\pi_i(c_i, c_j)]/dc_i$ are equal to zero, so that we obtain the local incentive constraint (LIC).

A.2 Proof of Lemma 1

We develop the proof in four steps. We begin by rewriting the local incentive constraint (LIC) in a way that takes into acount the binding (PC). We use it to show that, when this constraint is satisfied, all rewards are equal and all losses are equal under the Maxmin scheme. We then determine the *ex post* profits. We lastly prove that it is optimal to reward each agent for only one type c_j and let him bear a loss for all other types.

A.2.1 Local incentive compatibility rewritten

Using (LIC) in (33), together with $q_i(c_i, c_j) = q_i^{fb}(c_i)$, we obtain

$$q_i^{fb}(c_i) + K'(c_i) = \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) \frac{df(c_j|c_i)}{dc_i} dc_j \tag{35}$$

A.2.2 All rewards are equal and all losses are equal under the Maxmin scheme

Take three different feasible types of A_j, namely c_{j1}, c_{j2}, $c_{j3} \in [\underline{c}, \overline{c}]$ and suppose that $\pi_i(c_i, c_{j1}) > \pi_i(c_i, c_{j2}) > \pi_i(c_i, c_{j3})$, with $\pi_i(c_i, c_{j3})$ the highest loss at the solution to (Maxmin). Because (PC) is binding at FB implementation, we can write

$$\begin{aligned}
& \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) \frac{df(c_j|c_i)}{dc_i} dc_j \\
= & \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) \frac{df(c_j|c_i)}{dc_i} dc_j - \frac{df(c_{j1}|c_i)/dc_i}{f(c_{j1}|c_i)} \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) f(c_j|c_i) dc_j \\
= & \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) f(c_j|c_i) \gamma(c_j, c_{j1}) dc_j,
\end{aligned}$$

with

$$\gamma(c_j, c_{j1}) \equiv \frac{\frac{df(c_j|c_i)}{dc_i}}{f(c_j|c_i)} - \frac{\frac{df(c_{j1}|c_i)}{dc_i}}{f(c_{j1}|c_i)}.$$

We use this expression to rewrite (35) as

$$q_i^{fb}(c_i) + K'(c_i) = \int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) f(c_j|c_i) \gamma(c_j, c_{j1}) dc_j. \tag{36}$$

The left-hand side of (36) (LHS hereafter) is independent of $\pi_i(c_i, c_{j1})$, $\pi_i(c_i, c_{j2})$ and $\pi_i(c_i, c_{j3})$. The right-hand side of (36) (RHS hereafter) is independent of $\pi_i(c_i, c_{j1})$ because $\gamma(c_{j1}, c_{j1}) = 0$. Assume that $\pi_i(c_i, c_{j3})$ is raised by some $\varepsilon > 0$. Keeping all other profits constant, $\pi_i(c_i, c_{j2})$ must vary by some amount $\Delta\pi_i(c_i, c_{j2})$ for the RHS of (36) to remain unchanged. Thus, it is

$$\varepsilon f(c_{j3}|c_i) \gamma(c_{j3}, c_{j1}) = -\Delta\pi_i(c_i, c_{j2}) f(c_{j2}|c_i) \gamma(c_j, c_{j2})$$

or, equivalently,

$$\Delta\pi_i(c_i, c_{j2}) = -\varepsilon \frac{f(c_{j3}|c_i) \gamma(c_{j3}, c_{j1})}{f(c_{j2}|c_i) \gamma(c_j, c_{j2})}. \tag{37}$$

At the same time, to ensure that (PC) is binding, $\pi_i(c_i, c_{j1})$ must also change by some amount $\Delta\pi_i(c_i, c_{j1})$, so that

$$\int_{\underline{c}}^{\overline{c}} \pi_i(c_i, c_j) f(c_j|c_i) dc_j + \varepsilon f(c_{j3}|c_i)$$
$$+\Delta\pi_i(c_i, c_{j2}) f(c_{j2}|c_i) + \Delta\pi_i(c_i, c_{j1}) f(c_{j1}|c_i) = 0.$$

From this equality and (34),

$$\varepsilon f(c_{j3}|c_i) + \Delta\pi_i(c_i, c_{j2}) f(c_{j2}|c_i) + \Delta\pi_i(c_i, c_{j1}) f(c_{j1}|c_i) = 0.$$

Substituting $\Delta\pi_i(c_i, c_{j2})$ from (37) in the above expression we then find

$$\Delta\pi_i(c_i, c_{j1}) = -\varepsilon \frac{f(c_{j3}|c_i)}{f(c_{j1}|c_i)} \left[1 - \frac{\gamma(c_{j3}, c_{j1})}{\gamma(c_{j2}, c_{j1})}\right]. \tag{38}$$

From (37) and (38), we deduce that if we choose $\varepsilon > 0$ to be sufficiently small, $\pi_i(c_i, c_{j3})$ can be raised without changing the initial ranking of the three profits. This result contradicts the assumption that $\pi_i(c_i, c_{j3})$ is the biggest loss at the solution to (Maxmin). Hence, it must be the case that either $\pi_i(c_i, c_{j1}) = \pi_i(c_i, c_{j2})$ or $\pi_i(c_i, c_{j2}) = \pi_i(c_i, c_{j3})$.

A.2.3 *Ex post* profits

Using the previous result we denote the ex post profits assigned to A_i as $\pi_{i,r}$ (reward) and $\pi_{i,p}$ (punishment). We further define a function $p(c_j|c_i)$ such that $p(c_j|c_i) = f(c_j|c_i)$ if A_i is rewarded and $p(c_j|c_i) = 0$ otherwise. Using this notation, (34) is rewritten

$$\pi_{i,p} = -\frac{\int_{\underline{c}}^{\overline{c}} p(c_j|c_i) dc_j}{1 - \int_{\underline{c}}^{\overline{c}} p(c_j|c_i) dc_j} \pi_{i,r}. \tag{39}$$

Using the above notations we also rewrite (35) as

$$q_i^{fb}(c_i) + K'(c_i) = \pi_{i,r} \int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i} dc_j - \pi_{i,p} \int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i} dc_j. \tag{40}$$

From (39) and (40), we obtain the *ex post* profits

$$\pi_{i,r} = \left[q_i^{fb}(c_i) + K'(c_i)\right] \frac{1 - \int_{\underline{c}}^{\overline{c}} p(c_j|c_i) dc_j}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i} dc_j} \tag{41}$$

$$\pi_{i,p} = \left[q_i^{fb}(c_i) + K'(c_i)\right] \frac{-\int_{\underline{c}}^{\overline{c}} p(c_j|c_i) dc_j}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i} dc_j}. \tag{42}$$

A.2.4 One reward is optimal

At the solution to (Maxmin), the punishment $\pi_{i,p}$ takes the maximum feasible value. This requires that the ratio $\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j \Big/ \int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j$ be minimized with respect to c_j. The ratio depends on the number of rewards. Assume that, starting from a certain number of rewards, a new one is added for some realization $\widetilde{c}_j$. For this to enhance P's problem, it must be the case that the difference

$$\begin{aligned} D &= \frac{\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j + f(\widetilde{c}_j|c_i)}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j + \frac{df(\widetilde{c}_j|c_i)}{dc_i}} - \frac{\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j} \\ &= \frac{f(\widetilde{c}_j|c_i)\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j - \frac{df(\widetilde{c}_j|c_i)}{dc_i}\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j}{\left[\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j + \frac{df(\widetilde{c}_j|c_i)}{dc_i}\right]\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j} \end{aligned}$$

be negative. Notice that, from (41) and (42), reward states must be such that $\left(\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j\right) > 0$. Thus, for the new reward to be conveniently added for some realization $\widetilde{c}_j$, one needs to have $\left(\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j + \frac{df(\widetilde{c}_j|c_i)}{dc_i}\right) > 0$. For this reason, the sign of D coincides with the sign of its numerator, which is rewritten

$$N(D) = \frac{df(\widetilde{c}_j|c_i)}{dc_i}\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j \left[\frac{f(\widetilde{c}_j|c_i)}{\frac{df(\widetilde{c}_j|c_i)}{dc_i}} - \frac{\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j}\right].$$

If $(df(\widetilde{c}_j|c_i)/dc_i) < 0$, then A_i cannot be rewarded when A_j has type $\widetilde{c}_j$ because $N(D) > 0$. If $(df(\widetilde{c}_j|c_i)/dc_i) > 0$, then A_i is rewarded if and only if

$$\frac{f(\widetilde{c}_j|c_i)}{\frac{df(\widetilde{c}_j|c_i)}{dc_i}} - \frac{\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j}{\int_{\underline{c}}^{\overline{c}} \frac{dp(c_j|c_i)}{dc_i}dc_j} < 0,$$

in which case $N(D) < 0$ and so $D < 0$. For some given c_i, take $c_{jr}(c_i)$ to be the type of A_j for which the ratio $\frac{dp(c_j|c_i)/dc_i}{p(c_j|c_i)}$ is highest. Then, $\gamma(\widetilde{c}_j, c_{jr}(c_i)) < 0$, $\forall \widetilde{c}_j \in [\underline{c}, \overline{c}]$. Using this in the above condition, it follows that it is optimal to assign only one reward, when A_j has type $c_{jr}(c_i)$, and a loss when A_j has any type $\widetilde{c}_j \neq c_{jr}(c_i)$.

A.3 Determination of (3) and (4)

From the definition of $p(c_j|c_i)$ in the proof of Lemma 1 and from Lemma 1, it is $\int_{\underline{c}}^{\overline{c}} p(c_j|c_i)\,dc_j = f(c_{jr}(c_i)|c_i)$. Using this in (41) and (42), (3) and (4) are determined.

A.4 Proof of Lemma 2

Suppose $c_{jr}(c_i) = c$, $\forall c_i \in [\underline{c}, \overline{c}]$, with c some constant from $[\underline{c}, \overline{c}]$. Take also $df(c|c_i)/dc_i > 0$. For $c_i < \widehat{c}$, the punishment is as from (4), *i.e.* $\pi_{i,p}^{fb}(c_i) < 0$, with

$c_{jr}(c_i) = c$. Similarly, for $c_i > \widehat{c}$, the punishment is as from (3), *i.e.* $\pi_{i,r}^{fb}(c_i) < 0$, with $c_{jr}(c_i) = c$. Furthermore, the Maxmin scheme must be such that c maximizes both $\pi_{i,p}^{fb}(c_i)$ for $c_i < \widehat{c}$ and $\pi_{i,r}^{fb}(c_i)$ for $c_i > \widehat{c}$ at once. The former requires that $\frac{d}{dc_i}\left(\frac{f(c_j|c_i)}{f(c|c_i)}\right) < 0$, $\forall c_j \neq c$, the latter that $\frac{d}{dc_i}\left(\frac{1-f(c_j|c_i)}{f(c|c_i)}\right) > 0$, $\forall c_j \neq c$. Assume $\frac{d}{dc_i}\left(\frac{f(c_j|c_i)}{f(c|c_i)}\right) < 0$ holds true. Together with $df(c|c_i)/dc_i > 0$, this involves that $\frac{d}{dc_i}\left(\frac{1-f(c_j|c_i)}{f(c|c_i)}\right) \leq 0$, contradicting the hypothesis that $c_{jr}(c_i) = c$.

The proof proceeds similarly for $df(c|c_i)/dc_i < 0$.

A.5 Proof of Proposition 1

The profits (3) and (4) are such that the local incentive constraint (35) is satisfied (recall the proof of Lemma 1 and 2). In what follows, we first find the condition for global incentive compatibility and then provide the proof of Proposition 1.

A.5.1 Global incentive compatibility

The *interim* profit when A_i is of type c_i and reports r_i is written

$$\begin{aligned}
E_{c_j}\left[\widetilde{\pi}_i(r_i, c_j)\right] &\equiv \int_{\underline{c}}^{\overline{c}} \widetilde{\pi}_i(r_i, c_j) f(c_j|c_i)\,dc_j \\
&= q_i^{fb}(r_i)(r_i - c_i) + K(r_i) - K(c_i) \\
&\quad + \int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|c_i)\,dc_j \\
&= q_i^{fb}(r_i)(r_i - c_i) + K(r_i) - K(c_i) + \int_{\underline{c}}^{\overline{c}} \pi_{i,p}^{fb}(r_i) f(c_j|c_i)\,dc_j \\
&\quad + \left[\pi_{i,r}^{fb}(r_i) - \pi_{i,p}^{fb}(r_i)\right] f(c_{jr}(r_i)|c_i)
\end{aligned} \tag{43}$$

Substituting the values from (3) and (4), we rewrite (43) as

$$\begin{aligned}
E_{c_j}\left[\widetilde{\pi}_i(r_i, c_j)\right] &= \int_{ci}^{ri} \left\{\left[q_i^{fb}(r_i) + K'(r_i)\right]\left[1 - \frac{df(c_{jr}(r_i)|x)/dx}{df(c_{jr}(r_i)|r_i)/dr_i}\right]\right. \\
&\quad \left. + K'(x) - K'(r_i)\right\} dx.
\end{aligned} \tag{44}$$

From (44), from the condition for global incentive compatibility $E\left[\widetilde{\pi}_i(r_i, c_j)\right] \leq 0$ and taking into account that $c_{jr}(r_i) = \overline{c}$ if $r_i < \widehat{c}$ and $c_{jr}(r_i) = \underline{c}$ if $r_i > \widehat{c}$, we deduce the following conditions:

$$K'(r_i) - K'(c_i) \geq \left[q_i^{fb}(r_i) + K'(r_i)\right]\left[1 - \frac{df(c_{jr}(r_i)|c_i)/dc_i}{df(c_{jr}(r_i)|r_i)/dr_i}\right], \text{ if } r_i \geq c_i \tag{45}$$

$$K'(r_i) - K'(c_i) \leq \left[q_i^{fb}(r_i) + K'(r_i)\right]\left[1 - \frac{df(c_{jr}(r_i)|c_i)/dc_i}{df(c_{jr}(r_i)|r_i)/dr_i}\right], \text{ if } r_i \leq c_i \tag{46}$$

These conditions are satisfied if $K"(c_i) \geq 0$.

A.5.2 Proposition 1

From (3) and (4), $\pi_{i,p}^{fb}(c_i, c_{jr}(r_i)) < 0 < \pi_{i,r}^{fb}(c_i, c_{jr}(r_i))$. Hence, (LL) is satisfied as long as

$$\left[q_i^{fb}(c_i) + K'(c_i)\right] \frac{f(c_{jr}(c_i)|c_i)}{df(c_{jr}(c_i)|c_i)/dc_i} \leq L. \quad (47)$$

Because (3) and (4) are locally incentive compatible (see proof of Lemma 1 and 2) and since they belong to the Maxmin scheme, FB is implemented if and only if (47) is satisfied, provided that $K"(c_i) \geq 0$. Moreover, under Assumption (1) and (3), $c_{jr} = \overline{c}$ if $c_i \leq \widehat{c}$ and $c_{jr} = \underline{c}$ otherwise.

A.6 Proof of Corollary 1

It follows immediately from the global incentive constraints (45) and (46), with $q_i^{fb}(r_i) + K'(r_i) \neq 0 \ \forall r_i \in [\underline{c}, \overline{c}]$.

A.7 Proof of Lemma 3

We hereafter show that the expected value of the lottery, $\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|c_i) dc_j$, is minimized (with (LIC), (PC) and (LL) all satisfied) when P assigns one reward and losses that are all equal to $-L$. We proceed as follows. We first calculate the expected value of the lottery with one reward and punishments all equal to $-L$. We then calculate the expected value of the lottery with three distinct profits, the smallest of which equal to $-L$. We finally 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 A_i is of type c_i and reports r_i, he receives a reward $\pi_{i,r}(r_i)$ for $c_j = c_{jr}(r_i)$ and a loss $\pi_{i,p}(r_i) = -L$ for all other values of c_j. P seeks to minimize

$$\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|c_i) dc_j = \pi_{i,p}(r_i) + [\pi_{i,r}(r_i) - \pi_{i,p}(r_i)] f(c_{jr}(r_i)|c_i). \quad (48)$$

With (PC) binding for type r_i, we have

$$\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|r_i) dc_j = 0,$$

which is rewritten

$$\pi_{i,r}(r_i) = -\frac{1 - f(c_{jr}(r_i)|r_i)}{f(c_{jr}(r_i)|r_i)} \pi_{i,p}(r_i).$$

Replacing this expression into (48) together with $\pi_{i,p}(c_i) = -L$, we get

$$\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j|c_i) dc_j = -L\left[1 - \frac{f(c_{jr}(r_i)|c_i)}{f(c_{jr}(r_i)|r_i)}\right] \equiv \Omega. \quad (49)$$

Assume next that P implements FB with three distinct profits $\pi_{i,p}(c_i)$, $\pi_1(c_i)$ and $\pi_{i,r}(c_i)$, such that $\pi_{i,p}(c_i) = -L$ and $\pi_{i,p}(c_i) < \pi_1(c_i) < \pi_{i,r}(c_i)$. Profit $\pi_1(c_i)$ is assigned when $c_j = c_{j1}$ and $\pi_{i,r}(c_i)$ when $c_j = c_{jr}(c_i)$. The expected value of the lottery becomes

$$\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j | c_i) dc_j = [\pi_{i,r}(r_i) - \pi_{i,p}(r_i)] f(c_{jr}(r_i) | c_i) + \pi_{i,p}(r_i) + [\pi_1(r_i) - \pi_{i,p}(r_i)] f(c_{j1} | c_i), \quad (50)$$

whereas the binding (PC) is now written

$$\pi_{i,r}(r_i) - \pi_{i,p}(r_i) = -\frac{\pi_{i,p}(r_i) + [\pi_1(r_i) - \pi_{i,p}(r_i)] f(c_{j1} | r_i)}{f(c_{jr}(r_i) | r_i)}.$$

Replacing this expression into (50), together with $\pi_{i,p}(r_i) = -L$, we obtain

$$\begin{aligned}\int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j | c_i) dc_j &= -L\left[1 - \frac{f(c_{jr}(r_i) | c_i)}{f(c_{jr}(r_i) | r_i)}\right] \\ &\quad + [\pi_1(r_i) + L]\left[1 - \frac{f(c_{j1} | r_i) f(c_{jr}(r_i) | c_i)}{f(c_{j1} | c_i) f(c_{jr}(r_i) | r_i)}\right] f(c_{j1} | c_i) \\ &\equiv \Psi \end{aligned} \quad (51)$$

We are now left with comparing (49) with (51). We calculate

$$\Psi - \Omega = [\pi_1(r_i) + L]\left[-\frac{f(c_{j1} | r_i) f(c_{jr}(r_i) | c_i)}{f(c_{j1} | c_i) f(c_{jr}(r_i) | r_i)} + 1\right] f(c_{j1} | c_i).$$

From Proposition 2, $c_{jr}(r_i) = \overline{c}$ if $r_i < \widehat{c}$ and $c_{jr}(r_i) = \underline{c}$ if $r_i > \widehat{c}$. Under Assumption 3 and because $\pi_1(r_i) + L > 0$, we have $\Psi - \Omega > 0$. Hence, the expected value of the lottery is higher with any profit triplet $\{\pi_{i,p}(c_i), \pi_1(c_i), \pi_{i,r}(c_i)\}$ such that $\pi_{i,p}(c_i) = -L$ and $\pi_{i,p}(c_i) < \pi_1(c_i) < \pi_{i,r}(c_i)$ than it is with the profit pair $\{\pi_{i,p}(c_i), \pi_{i,r}(c_i)\}$ such that $\pi_{i,p}(c_i) = -L$.

A.8 Proof of Proposition 2

Since $\pi_{i,p}(c_i) = -L$ and (PC) is binding, the reward is given by $\pi_{i,r}(r_i) = \frac{1 - f(c_{jr}(r_i)|c_i)}{f(c_{jr}(r_i)|c_i)} L$, with $c_{jr}(r_i)$ the type c_j for which A_i is rewarded whenever he reports r_i. Then, the payoff $E_{c_j}[\widetilde{\pi}_i(r_i, c_j)]$ described by (44) is rewritten

$$\begin{aligned} E_{c_j}[\widetilde{\pi}_i(r_i, c_j)] &= q_i^{fb}(r_i)(r_i - c_i) + K(r_i) - K(c_i) \\ &\quad + \int_{\underline{c}}^{\overline{c}} \pi_{i,p}^{fb}(r_i, c_{jr}(r_i)) f(c_j | c_i) dc_j \\ &\quad + \left[\pi_{i,r}^{fb}(r_i, c_{jr}(r_i)) - \pi_{i,p}^{fb}(r_i, c_{jr}(r_i))\right] f(c_{jr}(r_i) | c_i) \\ &= \int_{ci}^{ri} \left[q_i^{fb}(r_i) + K'(r_i) - L\frac{df(c_{jr}(r_i) | y_i)/dy_i}{f(c_{jr}(r_i) | r_i)} \right. \\ &\quad \left. + K'(y_i) - K'(r_i)\right] dy_i \end{aligned}$$

$E_{c_j}[\widetilde{\pi}_i(r_i, c_j)] \leq 0$ for all $y_i < r_i$ under condition (12).

To see which value $c_{jr}(r_i)$ takes, suppose $c_i < r_i$. Under Assumption 3, the ratio $\frac{f(c_{jr}(r_i)|c_i)}{f(c_{jr}(r_i)|r_i)}$ in the expression here above is minimized. Hence, (12) is least stringent with $c_{jr}(r_i) = \overline{c}$ when $r_i < \widehat{c}$ and $c_{jr}(r_i) = \underline{c}$ when $r_i > \widehat{c}$. The same reasoning applies for $c_i > r_i$.

B The optimal contract with tight limited liability

B.1 Proof of Lemma 4

From the definition of $\widehat{c}$ in Lemma 2, (7) holds for $c_i = \widehat{c}$, $\forall L \geq 0$. Take now $c_i < \widehat{c}$ and suppose that (7) is violated for c_i, so that

$$q_i^{fb}(c_i) + K'(c_i) > L\frac{df(\overline{c}|c_i)/dc_i}{f(\overline{c}|c_i)}. \qquad (52)$$

(i) Suppose that

$$\frac{dq_i^{fb}(c_i)}{dc_i} + K^{"}(c_i) < \frac{L}{f(\overline{c}|c_i)}\left[\frac{d^2f(\overline{c}|c_i)}{dc_i^2} - \frac{(df(\overline{c}|c_i)/dc_i)^2}{f(\overline{c}|c_i)}\right]. \qquad (53)$$

meaning that, as c_i raises, the LHS of (52) (which is negative by Assumption 1) decreases faster than the RHS. Since (52) does not hold for $c_i = \widehat{c}$, there is at most one value $c_{i2} \in [\underline{c}, \widehat{c}]$ such that (52) does not hold if $c_i \in [\underline{c}, c_{i2})$ and holds if $c_i \in [c_{i2}, \widehat{c}]$. This value exists if (52) holds for $c_i = \underline{c}$.

(ii) Next suppose that (53) is not satisfied, so that, as c_i raises, the LHS of (52) decreases less fast than the RHS. Hence, if (52) does not hold for $c_i = \underline{c}$, then it does not hold for any $c_i \in [\underline{c}, \widehat{c}]$, in which case there is no $c_i \in [\underline{c}, \widehat{c}]$ for which (7) is violated. If (52) holds for $c_i = \underline{c}$, then it must hold for any $c_i \in [\underline{c}, \widehat{c}]$, involving that (7) is violated for all types within this interval. This contradicts the definition of $\widehat{c}$, under which (7) is satisfied for $c_i = \widehat{c}$. Therefore, (52) does not hold for $c_i = \underline{c}$, so that (7) is satisfied for all $c_i \in [\underline{c}, \widehat{c}]$.

Considering (i) and (ii) altogether, we deduce that there exists at most one subset $[\underline{c}, c_{i2}) \subseteq [\underline{c}, \widehat{c}]$ over which (7) is violated, with $c_{i2} \in [\underline{c}, \widehat{c}]$. This value exists whenever (7) is violated for $c_i = \underline{c}$.

A similar reasoning applies when $c_i > \widehat{c}$, so that there exists at most one subset $(c_{i3}, \overline{c}] \subseteq [\widehat{c}, \overline{c}]$, with $c_{i3} \in [\widehat{c}, \overline{c}]$, for which (7) is violated.

B.2 Proof of Lemma 5

From the proof of Lemma 4, for any $c_i \in [\underline{c}, \widehat{c}]$ to which a rent accrues, the SB quantity is given by $q_i^I(c_i)$ as defined by (15). For any $c_i \in [\underline{c}, \widehat{c}]$ to which no rent accrues, the SB quantity is given by $q_i^{II}(c_i)$ as defined by (16). A rent is left to type $c_i \in [\underline{c}, \widehat{c}]$ if and only if

$$q_i^I(c_i) + K'(c_i) > L\frac{df(\overline{c}|c_i)/dc_i}{f(\overline{c}|c_i)}. \qquad (54)$$

For the types for which (54) is violated, P does better by choosing the quantity $q_i^{II}(c_i) \geq q_i^I(c_i)$ so as to extract all surplus. Using (54) together with the condition $K"(c_i) < -\frac{dq_i^I(c_i)}{dc_i}$ (from Assumption 1 and $\frac{dq_i^I(c_i)}{dc_i} < \frac{dq_i^{fb}(c_i)}{dc_i}$), the sequel of the proof proceeds identically to that of Lemma 4 once $q_i^{fb}(c_i)$ is replaced with $q_i^I(c_i)$, c_{i2} with c_{i1}, $\widehat{c}$ with c_{i2} and (52) with (54). The cost value c_{i1} exists whenever (54) is satisfied for $c_i = \underline{c}$, *i.e.* whenever a rent is given up to type $\underline{c}$.

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

B.3 Proof of Lemma 6

B.3.1 Expected welfare

Define

$$\widetilde{W}(a,b) \equiv \sum_{i \neq j} \int_a^b \int_{\underline{c}}^{\overline{c}} [V(q_i(c_i,c_j)) + \alpha \pi_i(c_i,c_j)] f(c_j | c_i) f_i(c_i) dc_j dc_i, \tag{55}$$

so that the objective function in (Γ) is rewritten

$$\widetilde{W} = \sum_{i \neq j} \left[\widetilde{W}(\underline{c}, c_{i1}) + \widetilde{W}(c_{i1}, c_{i2}) + \widetilde{W}(c_{i2}, c_{i3}) + \widetilde{W}(c_{i3}, c_{i4}) + \widetilde{W}(c_{i4}, \overline{c}) \right].$$

Since the maximization of expected welfare 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 $\forall c_i \in [c_{i2}, c_{i3}]$ (Lemma 4) and we shall not come back to this case.

B.3.2 The solution for $c_i \in [\underline{c}, c_{i1})$

We first calculate the *ex post* transfer, then the expected transfer for $c_i \in [\underline{c}, c_{i1})$, namely $E[t_i(c_i) | c_i < c_{i1}]$. We finally replace it into the expression of $\widetilde{W}(\underline{c}, c_{i1})$ and optimize with respect to quantity.

The *ex post* transfer when $c_i \in [\underline{c}, c_{i1})$ It is useful to define $t_i(c_i, c_{jr}(c_i)) \equiv g_i(c_i)$ the transfer A_i receives when he is rewarded and $t_i(c_i, c_j) \equiv h_i(c_i, c_j)$ the transfer he receives when he is punished. For sake of simplicity, $h_i(c_i, c_j)$ is defined for any $c_j \in [\underline{c}, \overline{c}]$, although in reality $h_i(c_i, c_{jr}(c_i))$ does not exist (A_i is not actually punished in state $c_{jr}(c_i)$). Replacing into (31) and rearranging, we get

$$\begin{aligned} g_i(c_i) &= \int_{\underline{c}}^{\overline{c}} c_i \frac{dq_i(c_i,c_j)}{dc_i} \frac{f(c_j | c_i)}{f(c_{jr}(c_i) | c_i)} dc_j \\ &\quad - \int_{\underline{c}}^{\overline{c}} \frac{dh_i(c_i,c_j)}{dc_i} \frac{f(c_j | c_i)}{f(c_{jr}(c_i) | c_i)} dc_j + \frac{dh_i(c_i, c_{jr}(c_i))}{dc_i} \end{aligned}$$

Define $c_{ik} \in \{c_{i1}, c_{i4}\}$ any type c_i for which $E_{c_j}[\pi_i(c_{ik}, c_j)] = 0$. Integrating all terms above from c_i to c_{ik} we obtain

$$\begin{aligned} g_i(c_i) &= g_i(c_{ik}) - \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} y_i \frac{dq_i(y_i, c_j)}{dy_i} \frac{f(c_j|y_i)}{f(c_{jr}(c_i)|y_i)} dc_j dy_i \\ &\quad + \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \frac{dh_i(c_i, c_j)}{dy_i} \frac{f(c_j|y_i)}{f(c_{jr}(c_i)|y_i)} dc_j dy_i - h_i(c_{ik}, c_{jr}(c_i)) \\ &\quad + h_i(c_i, c_{jr}(c_i)). \end{aligned} \tag{56}$$

Integrating by parts the second and the third term in the RHS of (56), we can write

$$\begin{aligned} g_i(c_i) &= g_i(c_{ik}) - h_i(c_{ik}, c_{jr}(c_i)) + h_i(c_i, c_{jr}(c_i)) \\ &\quad - \int_{\underline{c}}^{\overline{c}} c_{ik} q_i(c_{ik}, c_j) \frac{f(c_j|c_{ik})}{f(c_{jr}(c_i)|c_{ik})} dc_j + \int_{\underline{c}}^{\overline{c}} h_i(c_{ik}, c_j) \frac{f(c_j|c_{ik})}{f(c_{jr}(c_i)|c_{ik})} dc_j \\ &\quad + \int_{\underline{c}}^{\overline{c}} q_i(c_i, c_j) c_i \frac{f(c_j|c_i)}{f(c_{jr}(c_i)|c_i)} dc_j - \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) \frac{f(c_j|c_i)}{f(c_{jr}(c_i)|c_i)} dc_j \\ &\quad + \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} q_i(y_i, c_j) \frac{d}{dy_i} \left[y_i \frac{f(c_j|y_i)}{f(c_{jr}(c_i)|y_i)} \right] dc_j dy_i \\ &\quad - \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} h_i(y_i, c_j) \frac{d}{dy_i} \left[\frac{f(c_j|y_i)}{f(c_{jr}(c_i)|y_i)} \right] dc_j dy_i.. \end{aligned} \tag{57}$$

Denote

$$\psi(c_i, c_j) \equiv \frac{f(c_j|c_i)}{f(c_{jr}(c_i)|c_i)}, \tag{58}$$

so that it is $\psi(c_i, c_j) \equiv \frac{f(c_j|c_i)}{f(\overline{c}|c_i)}$ if $c_i < \widehat{c}$ and $\psi(c_i, c_j) \equiv \frac{f(c_j|c_i)}{f(\underline{c}|c_i)}$ if $c_i > \widehat{c}$. Rearranging terms in $g_i(c_i)$, we obtain

$$\begin{aligned} g_i(c_i) &= \int_{\underline{c}}^{\overline{c}} [h_i(c_{ik}, c_j) - c_{ik} q_i(c_{ik}, c_j)] \psi(c_{ik}, c_j) dc_j + g_i(c_{ik}) \\ &\quad - h_i(c_{ik}, c_{jr}(c_i)) + \int_{\underline{c}}^{\overline{c}} q_i(c_i, c_j) c_i \psi(c_i, c_j) dc_j \\ &\quad + \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} q_i(y_i, c_j) \frac{d}{dy_i} [y_i \psi(y_i, c_j)] dc_j dy_i - \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) \psi(c_i, c_j) dc_j \\ &\quad - \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} h_i(y_i, c_j) \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i + h_i(c_i, c_{jr}(c_i)). \end{aligned} \tag{59}$$

Using $(2a)$ we can rewrite

$$\begin{aligned} &\int_{\underline{c}}^{\overline{c}} [h_i(c_{ik}, c_j) - c_{ik} q_i(c_{ik}, c_j)] \psi(c_{ik}, c_j) dc_j + r_i(c_{ik}) - h_i(c_{ik}, c_{jr}(c_{ik})) \\ = &\int_{\underline{c}}^{\overline{c}} [\pi_i(c_{ik}, c_j) + K(c_{ik})] \psi(c_{ik}, c_j) dc_j. \end{aligned}$$

Replacing into (59) returns

$$
\begin{aligned}
g_i(c_i) &= \int_{\underline{c}}^{\overline{c}} [\pi_i(c_{ik}, c_j) + K(c_{ik})] \psi(c_{ik}, c_j) dc_j - \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) \psi(c_i, c_j) dc_j \\
&+ \int_{\underline{c}}^{\overline{c}} q_i(c_i, c_j) c_i \psi(c_i, c_j) dc_j + h_i(c_i, c_{jr}(c_i)) \\
&+ \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} q_i(y_i, c_j) \frac{d}{dy_i} [y_i \psi(y_i, c_j)] dc_j dy_i \\
&- \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} h_i(y_i, c_j) \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i.
\end{aligned} \tag{60}
$$

Using $(2a)$ as well as $t_i(c_i, c_j) = h(c_i, c_j)$ for $c_j \neq c_{jr}(c_i)$ and letting $\pi_{i,p}^{sb}(c_i, c_j)$ the punishment, we have $\pi_{i,p}^{sb}(c_i, c_j) = h_i(c_i, c_j) - [c_i q_i(c_i, c_j) + K(c_i)]$. We use this to rewrite the expression

$$
\begin{aligned}
& \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} q_i(y_i, c_j) \frac{d}{dy_i} [y_i \psi(y_i, c_j)] dc_j dy_i - \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} h_i(y_i, c_j) \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i \\
= & \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \left[q_i(y_i, c_j) \psi(y_i, c_j) + [y_i q_i(y_i, c_j) - h_i(y_i, c_j)] \frac{d\psi(y_i, c_j)}{dy_i} \right] dc_j dy_i \\
= & \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \left[q_i(y_i, c_j) \psi(y_i, c_j) - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(y_i, c_j)}{dy_i} \right] dc_j dy_i.
\end{aligned}
$$

Replacing this into (60) yields the *ex post* transfer

$$
\begin{aligned}
g_i(c_i) &= \int_{\underline{c}}^{\overline{c}} [\pi_i(c_{ik}, c_j) + K(c_{ik})] \psi(c_{ik}, c_j) dc_j + \int_{\underline{c}}^{\overline{c}} q_i(c_i, c_j) c_i \psi(c_i, c_j) dc_j \\
&- \left[\int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) \psi(c_i, c_j) dc_j - h_i(c_i, c_{jr}(c_i)) \right] \\
&+ \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} q_i(y_i, c_j) \psi(y_i, c_j) - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i.
\end{aligned} \tag{61}
$$

The expected transfer for $c_i \in [\underline{c}, c_{i1})$ Using the notation $h_i(c_i, c_j)$ and $g_i(c_i)$ as defined in the proof of Lemma 5, the expected transfer of A_i when $c_i < c_{i1}$ is given by

$$
\begin{aligned}
E[t_i(c_i) | c_i < c_{i1}] &= \int_{\underline{c}}^{c_{i1}} \left\{ \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) f(c_j | c_i) dc_j \right. \\
&\quad \left. + [g_i(c_i) - h_i(c_i, \overline{c})] f(\overline{c} | c_i) \right\} f_i(c_i) dc_j dc_i \\
&= \int_{\underline{c}}^{c_{i1}} \left[\int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j) f(c_j | c_i) dc_j - h(c_i, \overline{c}) f(\overline{c} | c_i) \right] f_i(c_i) dc_i \\
&\quad + \int_{\underline{c}}^{c_{i1}} g_i(c_i) f(\overline{c} | c_i) f_i(c_i) c_i.
\end{aligned}
$$

Substitute $c_{jr}(c_i) = \overline{c}$ into (61) and then substitute $g_i(c_i)$ from (61) into the above expression, with $c_{ik} = c_{i1}$. This yields

$$\begin{aligned} E\left[t_i(c_i)\,|c_i < c_{i1}\right] = \int_{\underline{c}}^{c_{i1}} \Bigg\{ \Bigg[\int_{\underline{c}}^{\overline{c}} \left[\pi_i(c_{i1}, c_j) + K(c_{i1})\right] \psi(c_{i1}, c_j) \\ + \int_{\underline{c}}^{\overline{c}} c_i q_i(c_i, c_j)\, \psi(c_i, c_j) + \int_{\underline{c}}^{c_{i1}} \int_{\underline{c}}^{\overline{c}} \Big(q_i(y_i, c_j)\, \psi(y_i, c_j) \\ - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(y_i, c_j)}{dy_i} \Big)\, dy_i \Bigg] dc_j \Bigg\} f(\overline{c}\,|c_i)\, f_i(c_i)\, dc_i. \end{aligned}$$

Define

$$\phi(c_i) \equiv \int_{\underline{c}}^{c_i} f(\overline{c}\,|y_i)\, f_i(y_i)\, dy_i, \forall c_i \in [\underline{c}, c_{i1}) \tag{62}$$

for any $c_i \in [\underline{c}, c_{i1})$. We calculate

$$\begin{aligned} & \int_{\underline{c}}^{c_{i1}} \Bigg\{ \int_{c_i}^{c_{i1}} \Bigg[\int_{\underline{c}}^{\overline{c}} \Big(q_i(y_i, c_j)\, \psi(y_i, c_j) \\ & \quad - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(y_i, c_j)}{dy_i} \Big)\, dy_i \Bigg] dc_j \Bigg\} f(\overline{c}\,|c_i)\, f_i(c_i)\, dc_i \\ = & \int_{\underline{c}}^{c_{i1}} \Bigg[\int_{\underline{c}}^{\overline{c}} \Big(q_i(c_i, c_j)\, \psi(c_i, c_j) \\ & \quad - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(c_i, c_j)}{dc_i} \Big)\, dc_j \Bigg] \phi(c_i, \overline{c})\, dc_i. \end{aligned}$$

We thus find

$$\begin{aligned} E\left[t_i(c_i)\,|c_i < c_{i1}\right] = & \int_{\underline{c}}^{c_{i1}} \Big\{ \int_{\underline{c}}^{\overline{c}} \big[\left[\pi_i(c_{i1}, c_j) + K(c_{i1})\right] \psi(c_{i1}, c_j) \\ & + c_i q_i(c_i, c_j)\, \psi(c_i, c_j)\big]\, dc_j \Big\} f(\overline{c}\,|c_i)\, f_i(c_i)\, dc_i \\ & + \int_{\underline{c}}^{c_{i1}} \Bigg\{ \int_{\underline{c}}^{\overline{c}} \Big[q_i(c_i, c_j)\, \psi(c_i, c_j) \\ & - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(c_i, c_j)}{dc_i} \Big]\, dc_j \Bigg\} \phi(c_i, \overline{c})\, dc_i. \end{aligned} \tag{63}$$

The optimal output for $c_i \in [\underline{c}, c_{i1})$ Substituting (63) into (2a) and then (2a) into the expression of $\widetilde{W}(\underline{c}, c_{i1})$ from (55), we rewrite it as follows

$$
\begin{aligned}
\widetilde{W}(\underline{c}, c_{i1}) &= \int_{\underline{c}}^{c_{i1}} \int_{\underline{c}}^{\overline{c}} [S(q_i(c_i, c_j)) - \alpha c_i q_i(c_i, c_j) - \alpha K(c_i)] f(c_j | c_i) f_i(c_i) dc_j dc_i \qquad (64)\\
& - (1-\alpha) \int_{\underline{c}}^{c_{i1}} \left\{ \int_{\underline{c}}^{\overline{c}} [[\pi_i(c_{i1}, c_j) + K(c_{i1})] \psi(c_{i1}, c_j) \right.\\
& \left. + c_i q_i(c_i, c_j) \psi(c_i, c_j)] dc_j \right\} f(\overline{c} | c_i) f_i(c_i) dc_i\\
& - (1-\alpha) \int_{\underline{c}}^{c_{i1}} \left\{ \int_{\underline{c}}^{\overline{c}} \left[q_i(c_i, c_j) \psi(c_i, c_j) \right.\right.\\
& \left.\left. - \left[\pi_{i,p}^{sb}(c_i, c_j) + K(c_i)\right] \frac{d\psi(c_i, c_j)}{dc_i} \right] dc_j \right\} \phi(c_i, \overline{c}) dc_i.
\end{aligned}
$$

From the definition of c_{i1} (see Lemma 5), $E[\pi_i(c_{i1}, c_j)] = 0$. Also, because $\widetilde{W}(\underline{c}, c_{i1})$ decreases with $\pi_{i,p}^{sb}(c_i, c_j)$, it is optimal to set the latter at the lowest feasible value, *i.e.* $\pi_{i,p}^{sb}(c_i, c_j) = -L$. Replacing into $\widetilde{W}(\underline{c}, c_{i1})$, the first-order condition with respect to q_i, $\forall c_i \in [\underline{c}, c_{i1})$, is given by

$$
\begin{aligned}
& [S'(q_i(c_i, c_j)) - \alpha c_i] f(c_i, c_j)\\
= \; & (1-\alpha) [c_i \psi(c_i, c_j) f(\overline{c} | c_i) f_i(c_i) + \psi(c_i, c_j) \phi(c_i, \overline{c})].
\end{aligned}
$$

Denoting $q_i^I(c_i)$ the quantity that satisfies the condition above and using the equality $f(c_j | c_i) f_i(c_i) = f(c_i, c_j)$ together with (58) and (62), we can rewrite

$$
\begin{aligned}
S'\left(q_i^I(c_i)\right) &= \alpha c_i + (1-\alpha) \frac{\psi(c_i, c_j)}{f(c_i, c_j)} [c_i f(\overline{c} | c_i) f_i(c_i) + \phi(c_i, \overline{c})]\\
&= c_i + (1-\alpha) \frac{\int_{\underline{c}}^{c_i} f(\overline{c} | y_i) f_i(y_i) dy_i}{f(\overline{c} | c_i) f_i(c_i)}\\
&= c_i + (1-\alpha) \frac{F(c_i | \overline{c})}{f(c_i | \overline{c})},
\end{aligned}
$$

with $F(c_i | \overline{c}) = \frac{\int_{\underline{c}}^{c_i} f(\overline{c}|y_i) f_i(y_i) dy_i}{\int_{\underline{c}}^{\overline{c}} f(\overline{c}|c_i) f_i(c_i) dc_i}$ and $f(c_i | \overline{c}) = \frac{f(\overline{c}|c_i) f_i(c_i)}{\int_{\underline{c}}^{\overline{c}} f(\overline{c}|c_i) f_i(c_i) dc_i}$.

B.3.3 The solution for $c_i \in [c_{i1}, c_{i2})$

From Lemma 5 one has $E_{c_j}[\pi_i(c_i, c_j)] = 0$ whenever $c_i \in [c_{i1}, c_{i2})$. It means that the functional form of *ex post* profits $\pi_i(c_i, c_j)$ is similar to (3) and (4), except that $q_i^{fb}(c_i)$ is replaced by $q_i^{II}(c_i)$, whose value we need to determine. In particular, the punishment is

$$
\pi_{i,p}^{sb}(c_i, c_j) = -\left[q_i^{II}(c_i) + K'(c_i)\right] \frac{f(\overline{c} | c_i)}{df(\overline{c} | c_i) / dc_i}.
$$

Moreover, by Lemma 4 $\pi_{i,p}^{sb}(c_i, c_j) = -L, \forall c_i \in [\underline{c}, c_{i2})$. Using in the above equation, we find that $q_i^{II}(c_i)$ is defined by (16) in the lemma.

The proof is identical for $c_i \in (c_{i3}, c_{i4}]$.

B.3.4 The solution for $c_i \in (c_{i4}, \overline{c}]$

Define

$$\varphi(c_i) \equiv \int_{c_i}^{\overline{c}} f(\underline{c}\,|y_i) f_i(y_i)\, dy_i, \forall c_i \in (c_{i4}, \overline{c}].$$

Proceeding as for $c_i \in [\underline{c}, c_{i1})$, one finds the expected transfer

$$\begin{aligned} E[t_i(c_i)\,|c_i > c_{i4}] &= \int_{c_{i4}}^{\overline{c}} \Big\{ \int_{\underline{c}}^{\overline{c}} [[\pi_i(c_{i4}, c_j) + K(c_{i4})]\, \psi(c_{i4}, c_j) \\ &\quad + c_i q_i(c_i, c_j)\, \psi(c_i, c_j)]\, dc_j \Big\} f(\underline{c}\,|c_i) f_i(c_i)\, dc_i \\ &\quad + \int_{c_{i4}}^{\overline{c}} \Big\{ \int_{\underline{c}}^{\overline{c}} \Big[[\pi_i(c_i, c_j) + K(c_i)] \frac{d\psi(c_i, c_j)}{dy_i} \\ &\quad - q_i(c_i, c_j)\, \psi(c_i, c_j)]\, dc_j \Big\} \varphi(c_i)\, dc_i. \end{aligned} \tag{65}$$

Substituting (65) into $\widetilde{W}(c_{i4}, \overline{c})$, we can characterize the optimal output $q_i^V(c_i)$ as

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

with $[1 - F(c_i\,|\underline{c})] = \frac{\int_{c_i}^{\overline{c}} f(\underline{c}|y_i) f_i(y_i) dy_i}{\int_{\underline{c}}^{\overline{c}} f(\underline{c}|c_i) f_i(c_i) dc_i}$ and $f(c_i\,|\underline{c}) = \frac{f(\underline{c}|c_i) f_i(c_i)}{\int_{\underline{c}}^{\overline{c}} f(\underline{c}|c_i) f_i(c_i) dc_i}$.

B.4 Proof of Proposition 3

The *interim* profit of A_i when he reports r_i and A_j reports his true type is given by

$$E_{c_j}[\widetilde{\pi}_i(r_i, c_j)] = q_i^{sb}(r_i)(r_i - c_i) + K(r_i) - K(c_i) + \int_{\underline{c}}^{\overline{c}} \pi_i(r_i, c_j) f(c_j\,|c_i), \tag{66}$$

with q_i^{sb} the SB output in each of the cost intervals in Lemma 6. The *ex post* profit $\pi_i(r_i, c_j)$ that appears in (66) is calculated differently, according to the value the report r_i takes. We thus perform the analysis case by case.

B.4.1 Case $r_i \in [\underline{c}, c_{i1}]$

We will proceed as follows. We first calculate the *ex post* profit $\pi_i(r_i, c_j)$, $\forall r_i \in [\underline{c}, c_{i1}]$, $\forall c_j \in [\underline{c}, \overline{c}]$, $\forall i \neq j \in \{1, 2\}$. We replace into (66) so as to calculate $E_{c_j}[\widetilde{\pi}_i(r_i, c_j)]$. We finally state the global incentive condition $E_{c_j}[\widetilde{\pi}_i(r_i, c_j)] \leq E_{c_j}[\pi_i(c_i, c_j)]$ for any report $r_i \in [\underline{c}, c_{i1}]$. Two sub-cases are considered, namely $c_i \in [\underline{c}, c_{i1}]$ and $c_i \in [c_{i1}, \overline{c}]$.

The *ex post* profit $\pi_i(r_i, c_j)$ Recall that $c_{ik} \in \{c_{i1}, c_{i4}\}$ is a type c_i for which $E_{c_j}(\pi_i(c_{ik}, c_j)) = 0$. Using the definition of c_{ik} and replacing $\int_{\underline{c}}^{\overline{c}} \psi(c_{ik}, c_j)\, dc_j = \frac{1}{f(c_{jr}(c_{ik})|c_{ik})}$ (from (58)) into (61), we obtain

$$\begin{aligned}
g_i(c_i) &= \frac{K(c_{ik})}{f(c_{jr}(c_{ik})|c_{ik})} + c_i q_i^{sb}(c_i) \int_{\underline{c}}^{\overline{c}} \psi(c_i, c_j)\, dc_j - \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j)\, \psi(c_i, c_j)\, dc_j \\
&\quad + \int_{c_i}^{c_{ik}} \left\{ \int_{\underline{c}}^{\overline{c}} \left[q_i^{sb}(y_i)\, \psi(y_i, c_j) - \left[\pi_{i,p}^{sb}(y_i, c_j) + K(y_i)\right] \frac{d\psi(y_i, c_j)}{dy_i} \right] dc_j \right\} dy_i \\
&\quad + h_i(c_i, c_{jr}(c_i)).
\end{aligned}$$

We further calculate

$$\begin{aligned}
& \int_{\underline{c}}^{\overline{c}} c_i q_i^{sb}(c_i)\, \psi(c_i, c_j)\, dc_j - \int_{\underline{c}}^{\overline{c}} h_i(c_i, c_j)\, \psi(c_i, c_j)\, dc_j + h_i(c_i, c_{jr}(c_i)) \\
=& \int_{\underline{c}}^{\overline{c}} \left[c_i q_i^{sb}(c_i) - h(c_i, c_j)\right] \psi(c_i, c_j)\, dc_j + h_i(c_i, c_{jr}(c_i)) \\
=& \int_{\underline{c}}^{\overline{c}} \left[-K(c_i) - \pi_{i,p}^{sb}(c_i, c_j)\right] \psi(c_i, c_j)\, dc_j + \pi_{i,p}^{sb}(c_i, c_{jr}(c_i)) \\
& + c_i q_i^{sb}(c_i) + K(c_i) \\
=& \int_{\underline{c}}^{\overline{c}} \left[-K(c_i) + L\right] \psi(c_i, c_j)\, dc_j - L + c_i q_i^{sb}(c_i) + K(c_i) \\
=& \left[L - K(c_i)\right] \frac{1 - f(c_{jr}(c_i)|c_i)}{f(c_{jr}(c_i)|c_i)} + c_i q_i^{sb}(c_i)
\end{aligned}$$

and then substitute into the expression of $g_i(c_i)$ above. Rearranging yields

$$\begin{aligned}
g_i(c_i) &= c_i q_i^{sb}(c_i) + \int_{c_i}^{c_{ik}} q_i^{sb}(y_i) \int_{\underline{c}}^{\overline{c}} \psi(y_i, c_j)\, dc_j dy_i \\
&\quad + \frac{1 - f(c_{jr}(c_i)|c_i)}{f(c_{jr}(c_i)|c_i)} \left[L - K(c_i)\right] + L \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i \\
&\quad - \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} K(y_i) \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i + \frac{K(c_{ik})}{f(c_{jr}(c_i)|c_{ik})}.
\end{aligned}$$

Integrating by parts $\int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i$, where $\psi(y_i, c_j)$ is defined by (58), and $\int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} K(y_i) \frac{d\psi(y_i, c_j)}{dy_i} dc_j dy_i$ and then replacing into the above expression of $g_i(c_i)$, we find

$$\begin{aligned}
g_i(c_i) &= c_i q_i^{sb}(c_i) + K(c_i) + \frac{1 - f(c_{jr}(c_i)|c_{ik})}{f(c_{jr}(c_i)|c_{ik})} L \\
&\quad + \int_{c_i}^{c_{ik}} \int_{\underline{c}}^{\overline{c}} \left[q_i^{sb}(y_i) + K'(y_i)\right] \psi(y_i, c_j)\, dc_j dy_i.
\end{aligned} \tag{67}$$

Using (67) in $(2a)$ for $t_i(r_i, c_{jr}(r_i)) = g_i(r_i)$ (knowing that $g_i(r_i)$ is the transfer that corresponds to type $c_{jr}(r_i)$), the reward of A_i when he reports $r_i \in [\underline{c}, c_{i1}]$ and A_j has type $c_{jr}(r_i) = \overline{c}$ is written

$$\pi_i(r_i, \overline{c}) = \int_{r_i}^{c_{i1}} \frac{q_i^I(y_i) + K'(y_i)}{f(\overline{c}\,|y_i)} dy_i + \frac{1 - f(\overline{c}\,|c_{i1})}{f(\overline{c}\,|c_{i1})} L, \tag{68}$$

From the proof of Lemma 6, $\pi_i(r_i, c_j) = -L$ whenever $r_i \in [\underline{c}, c_{i1}]$ and $c_j \neq \overline{c}$.

The *interim* profit Using (68) and $\pi_i(r_i, c_j) = -L$ for $c_j \neq \overline{c}$ in (66), $E_{c_j}[\widetilde{\pi}_i(r_i, c_j)]$ is rewritten

$$\begin{aligned} E_{c_j}[\widetilde{\pi}_i(r_i, c_j)] &= -\int_{r_i}^{c_i} \left[q_i^I(r_i) + K'(y_i)\right] dy_i + f(\overline{c}\,|c_i) \int_{r_i}^{c_{i1}} \frac{q_i^I(y_i) + K'(y_i)}{f(\overline{c}\,|y_i)} dy_i \\ &\quad - \left[1 - \frac{f(\overline{c}\,|c_i)}{f(\overline{c}\,|c_{i1})}\right] L, \end{aligned} \tag{69}$$

whereas the *interim* profit from a truthful report $r_i = c_i$ is given by

$$E_{c_j}[\pi_i(c_i, c_j)] = f(\overline{c}\,|c_i) \int_{c_i}^{c_{i1}} \frac{q_i^I(y_i) + K'(y_i)}{f(\overline{c}\,|y_i)} dy_i - \left[1 - \frac{f(\overline{c}\,|c_i)}{f(\overline{c}\,|c_{i1})}\right] L. \tag{70}$$

Sub-case $c_i \in [\underline{c}, c_{i1}]$ Using (69) and (70), we have $E_{c_j}[\pi_i(c_i, c_j)] \geq E_{c_j}[\widetilde{\pi}_i(r_i, c_j)]$ if and only if

$$\int_{c_i}^{r_i} \left[q_i^I(y_i) + K'(y_i)\right] \left[1 - \frac{f(\overline{c}\,|c_i)}{f(\overline{c}\,|y_i)}\right] dy_i + \int_{c_i}^{r_i} \left[q_i^I(r_i) - q_i^1(y_i)\right] dy_i \leq 0. \tag{71}$$

This condition is satisfied whenever so is (27) in Proposition 3.

Sub-case $c_i \in [c_{i1}, \overline{c}]$ Assume that $r_i = c_{i1}$ and calculate

$$\begin{aligned} \frac{dE_{c_j}[\widetilde{\pi}_i(c_{i1}, c_j)]}{dc_i} &= -\left[q_i^I(c_{i1}) + K'(c_i)\right] + \frac{L}{f(\overline{c}\,|c_{i1})} \frac{df(\overline{c}\,|c_i)}{dc_i} \\ &= -\left[q_i^I(c_{i1}) + K'(c_{i1})\right] + \frac{L}{f(\overline{c}\,|c_{i1})} \frac{df(\overline{c}\,|c_{i1})}{dc_{i1}} \\ &\quad + K'(c_{i1}) - K'(c_i) + \frac{L}{f(\overline{c}\,|c_{i1})} \left[\frac{df(\overline{c}\,|c_i)}{dc_i} - \frac{df(\overline{c}\,|c_{i1})}{dc_{i1}}\right] \\ &= K'(c_{i1}) - K'(c_i) + \frac{L}{f(\overline{c}\,|c_{i1})} \left[\frac{df(\overline{c}\,|c_i)}{dc_i} - \frac{df(\overline{c}\,|c_{i1})}{dc_{i1}}\right]. \end{aligned}$$

With (29) from Proposition 3 satisfied for $c_{jr}(c_i) = \overline{c}$, $\frac{dE_{c_j}[\widetilde{\pi}_i(c_{i1}, c_j)]}{dc_i} \leq 0$. Moreover, if $c_i = c_{i1}$, then $E_{c_j}[\widetilde{\pi}_i(c_{i1}, c_j)] = E_{c_j}[\pi_i(c_{i1}, c_j)]$, which is zero by Lemma 5. Therefore, under (29), $E_{c_j}[\widetilde{\pi}_i(c_{i1}, c_j)] \leq 0$ whenever $c_i \in [c_{i1}, \overline{c}]$ and $r_i = c_{i1}$.

Take now $r_i \leq c_{i1}$ and calculate

$$\frac{dE_{c_j}\left[\widetilde{\pi}_i\left(r_i,c_j\right)\right]}{dr_i}=-\int_{r_i}^{c_i}\left\{\left[q_i^I\left(r_i\right)+K'\left(r_i\right)\right]\frac{df\left(\overline{c}\,|y_i\right)/dy_i}{f\left(\overline{c}\,|r_i\right)}+\frac{dq_i^I\left(r_i\right)}{dr_i}\right\}dy_i.$$

We look for the condition under which $\frac{dE_{c_j}[\widetilde{\pi}_i(r_i,c_j)]}{dr_i}\geq 0$. Because $c_i \geq c_{i1}$ and $r_i \leq c_{i1}$, this inequality holds if and only if

$$\frac{dq_i^I\left(r_i\right)}{dr_i}\leq-\left[q_i^I\left(r_i\right)+K'\left(r_i\right)\right]\frac{df\left(\overline{c}\,|y_i\right)/dy_i}{f\left(\overline{c}\,|r_i\right)},\forall r_i\in\left[\underline{c},c_{i1}\right]\text{ and }y_i\geq c_{i1},\qquad(72)$$

which is implied by (27) together with Assumption 2 and $y_i \geq r_i$. Since $\frac{dE_{c_j}[\widetilde{\pi}_i(r_i,c_j)]}{dr_i}\geq 0\ \forall r_i\in[\underline{c},c_{i1}]$ and $c_i\in[c_{i1},\overline{c}]$, whereas $E_{c_j}\left[\widetilde{\pi}_i\left(c_{i1},c_j\right)\right]\leq 0$ (as previously found), one has $E_{c_j}\left[\widetilde{\pi}_i\left(r_i,c_j\right)\right]\leq 0\ \forall r_i\in[\underline{c},c_{i1}]$ and $c_i\in[c_{i1},\overline{c}]$.

Overall, (27) and (29) ensure that A_i has no incentive to report $r_i\in[\underline{c},c_{i1}]$ such that $r_i\neq c_i$, whatever his real type.

B.4.2 Case $r_i \in [c_{i1}, c_{i2}]$

Since $\pi_i\left(r_i,c_j\right)=-L$ for $c_j\neq\overline{c}$ and $E\left[\pi_i\left(r_i,c_j\right)\right]=0,\forall r_i\in[c_{i1},c_{i2}]$ (by Lemma 4), we determine $\pi_i\left(r_i,\overline{c}\right)=\frac{1-f(\overline{c}|r_i)}{f(\overline{c}|r_i)}L$. Substituting these values of $\pi_i\left(r_i,c_j\right)$ into $E_{c_j}\left[\widetilde{\pi}_i\left(r_i,c_j\right)\right]$, together with $q_i^{II}\left(r_i\right)=\frac{df(\overline{c}|r_i)/dr_i}{f(\overline{c}|r_i)}L-K'\left(r_i\right)$ (from Lemma 6), for this interval (66) specifies as

$$E\left[\widetilde{\pi}_i^{II}\right]=\int_{c_i}^{r_i}\left\{\frac{L}{f\left(\overline{c}\,|r_i\right)}\left[\frac{df\left(\overline{c}\,|r_i\right)}{dr_i}-\frac{df\left(\overline{c}\,|y_i\right)}{dy_i}\right]+K'\left(y_i\right)-K'\left(r_i\right)\right\}dy_i.$$

From the expression above, condition (29) and $c_{jr}\left(c_i\right)=\overline{c}$, we can establish that $E\left[\widetilde{\pi}_i^{II}\right]\leq 0$.

B.4.3 Case $r_i \in [c_{i2}, c_{i3}]$

Proposition 2 shows that, in this case, the condition for global incentive compatibility is given by (12), which is implied by (29).

B.4.4 Case $r_i \in [c_{i3}, c_{i4}]$

Proceeding as we did for $r_i\in[c_{i1},c_{i2}]$, we find that the payoff of A_i when he reports r_i is written

$$\begin{aligned}E\left[\widetilde{\pi}_i^{IV}\right]&=\int_{c_i}^{r_i}\left[L\frac{df\left(\underline{c}\,|r_i\right)/dr_i}{f\left(\underline{c}\,|r_i\right)}-K'\left(r_i\right)+K'\left(y_i\right)-\frac{L}{f\left(\underline{c}\,|r_i\right)}\frac{df\left(\underline{c}\,|y_i\right)}{dy_i}\right]dy_i\\&=\int_{c_i}^{r_i}\left\{\frac{L}{f\left(\underline{c}\,|r_i\right)}\left[\frac{df\left(\underline{c}\,|r_i\right)}{dr_i}-\frac{df\left(\underline{c}\,|y_i\right)}{dy_i}\right]+K'\left(y_i\right)-K'\left(r_i\right)\right\}dy_i.\end{aligned}$$

Together with $c_{jr}\left(c_i\right)=\underline{c}$, (29) in Proposition 3 involves that $E\left[\widetilde{\pi}_i^{IV}\right]\leq 0$.

B.4.5 Case $r_i \in [c_{i4}, \overline{c}]$

We follow the same steps as we did for the very first case.

The *ex post* profit $\pi_i(r_i, c_j)$ Using (67) in $(2a)$ for $t_i(r_i, c_{jr}(r_i)) = g_i(r_i)$, the reward of A_i when he reports $r_i \in [c_{i4}, \overline{c}]$ and A_j has type $c_{jr}(r_i) = \underline{c}$ is written

$$\pi_i(r_i, \underline{c}) = -\int_{c_{i4}}^{r_i} \frac{q_i^V(y_i) + K'(y_i)}{f(\underline{c}|y_i)} dy_i + \frac{1 - f(\underline{c}|c_{i4})}{f(\underline{c}|c_{i4})} L, \tag{73}$$

From the proof of Lemma 6, $\pi_i(r_i, c_j) = -L$ whenever $r_i \in [\underline{c}, c_{i4}]$ and $c_j \neq \underline{c}$.

The *interim* profit The *interim* profit of A_i when he reports r_i is given by

$$\begin{aligned} E\left[\widetilde{\pi}_i^V\right] &= \int_{c_i}^{r_i} \left[q_i^V(r_i) + K'(y_i)\right] dy_i - \int_{c_{i4}}^{r_i} \left[q_i^V(y_i) + K'(y_i)\right] \frac{f(\underline{c}|c_i)}{f(\underline{c}|y_i)} dy_i \\ &\quad -L\left[1 - \frac{f(\underline{c}|c_i)}{f(\underline{c}|c_{i4})}\right], \end{aligned} \tag{74}$$

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

$$E_{c_j}[\pi_i(c_i, c_j)] = -\int_{c_{i4}}^{c_i} \left[q_i^V(y_i) + K'(y_i)\right] \frac{f(\underline{c}|c_i)}{f(\underline{c}|y_i)} dy_i - L\left[1 - \frac{f(\underline{c}|c_i)}{f(\underline{c}|c_{i4})}\right]. \tag{75}$$

Sub-case $c_i \in [c_{i4}, \overline{c}]$ Using (74) and (75), we have $E_{c_j}[\pi_i(c_i, c_j)] \geq E_{c_j}[\widetilde{\pi}_i(r_i, c_j)]$ if and only if

$$\int_{c_i}^{r_i} \left[q_i^V(y_i) + K'(y_i)\right] \left[1 - \frac{f(\underline{c}|c_i)}{f(\underline{c}|y_i)}\right] dy_i + \int_{c_i}^{r_i} \left[q_i^V(r_i) - q_i^V(y_i)\right] dy_i \leq 0.$$

This condition is satisfied whenever so is (28) in Proposition 3.

Sub-case $c_i \notin [c_{i4}, \overline{c}]$ Take first $r_i = c_{i4}$ and calculate

$$\begin{aligned} \frac{dE\left[\widetilde{\pi}_i^V\right]}{dc_i} &= -\left[q_i^V(c_{i4}) + K'(c_i)\right] + \frac{df(\underline{c}|c_i)/dc_i}{f(\underline{c}|c_{i4})} L \\ &= -\left[q_i^V(c_{i4}) + K'(c_{i4})\right] + \frac{df(\underline{c}|c_{i4})/dc_{i4}}{f(\underline{c}|c_{i4})} L \\ &\quad +K'(c_{i4}) - K'(c_i) + \frac{df(\underline{c}|c_i)/dc_i - df(\underline{c}|c_{i4})/dc_{i4}}{f(\underline{c}|c_{i4})} L \\ &= K'(c_{i4}) - K'(c_i) + \frac{df(\underline{c}|c_i)/dc_i - df(\underline{c}|c_{i4})/dc_{i4}}{f(\underline{c}|c_{i4})} L. \end{aligned}$$

One has $\frac{dE\left[\widetilde{\pi}_i^V\right]}{dc_i} \geq 0$ if (29) in Proposition 3 holds. Moreover, $E\left[\widetilde{\pi}_i^V\right] = 0$ if $c_i = r_i = c_{i4}$. This shows that any type $c_i \notin [c_{i4}, \overline{c}]$ that reports $r_i = c_{i4}$ obtains $E\left[\widetilde{\pi}_i^V\right] \leq 0$.

Furthermore,

$$\begin{aligned}\frac{dE\left[\widetilde{\pi}_i^V\right]}{dr_i} &= \left[q_i^V(r_i)+K'(r_i)\right]\left[1-\frac{f(\underline{c}\,|c_i)}{f(\underline{c}\,|r_i)}\right]+\int_{c_i}^{r_i}\frac{dq_i^V(r_i)}{dr_i}dy_i \\ &= \int_{c_i}^{r_i}\left\{\frac{dq_i^V(r_i)}{dr_i}+\left[q_i^V(r_i)+K'(r_i)\right]\frac{df(\underline{c}\,|y_i)/dy_i}{f(\underline{c}\,|r_i)}\right\}dy_i.\end{aligned}$$

$E\left[\widetilde{\pi}_i^V\right]\leq 0$ for any report $r_i\in[c_{i4},\overline{c}]$ if $\frac{dE\left[\widetilde{\pi}_i^V\right]}{dr_i}\leq 0$, which is implied by

$$\frac{dq_i^V(r_i)}{dr_i}\leq-\left[q_i^V(r_i)+K'(r_i)\right]\frac{df(\underline{c}\,|y_i)/dy_i}{f(\underline{c}\,|r_i)}.$$

In turn, this is implied by (28) in Proposition 3 together with Assumption 2 and $y_i\leq r_i$.

Overall, A_i has no incentive to report $r_i\in[c_{i4},\overline{c}]$ such that $r_i\neq c_i$ whenever (28) and (29) are satisfied.

C Proof of Proposition 4

We proceed as follows. We begin by showing that, whenever (29) is violated for any feasible c_i, at the SB solution, there exists no $c_i\neq\widehat{c}$ for which (PC) is binding. We then rewrite (Γ) for the situation in which (PC) is not binding $\forall c_i\neq\widehat{c}$ and show that there exists a unique cost range over which pooling arises.

Assume that (PC) is binding over some non empty interval $\left[c_i^L,c_i^H\right]$, with either $c_i^L\neq\widehat{c}$ or $c_i^H\neq\widehat{c}$ or both, $\forall i\in\{1,2\}$. Assume also that FB is not implementable over this interval at the solution to (Γ). From the proof of Lemma 6, the SB quantity would be $q_i^{II}(c_i),\forall c_i\in\left[c_i^L,c_i^H\right]$. Furthermore, the proof of Proposition 3 shows that the quantity $q_i^{II}(c_i)$ and the transfers that leave no rent to agents are not implementable when (29) is not satisfied for any feasible c_i. This contradicts the assumption that (PC) is binding and, at the same time, FB is not at hand for types in $\left[c_i^L,c_i^H\right]$.

Assume now that, at the solution to (Γ), (PC) is binding and FB is implemented $\forall c_i\in\left[c_i^L,c_i^H\right]$. From Lemma 5, it follows that there exist other cost values around $\left[c_i^L,c_i^H\right]$ for which (PC) is binding and the SB quantity is $q_i^{II}(c_i)$. Once again, since (29) is not satisfied for any feasible c_i, this contradicts also the assumption that FB is enforced $\forall c_i\in\left[c_i^L,c_i^H\right]$.

Overall, there exists no subset $\left[c_i^L,c_i^H\right]$, with either $c_i^L\neq\widehat{c}$ or $c_i^H\neq\widehat{c}$ or both, in which (PC) is binding. It follows that (PC) is slack for $\forall c_i\neq\widehat{c}$. Hence, the interval $[c_{i2},c_{i4}]$ defined by Lemma 4 reduces to the singleton $\{\widehat{c}\}$. From the proof of Proposition 3, the scheme is globally incentive compatible whenever

$$\frac{dq_i^{sb}(c_i)}{dc_i}\leq-\left[q_i^{sb}(c_i)+K'(c_i)\right]\frac{df(c_{jr}(c_i)\,|c_i)/dc_i}{f(c_{jr}(c_i)\,|c_i)},\ \forall c_i\in[\underline{c},\overline{c}],c_i\neq\widehat{c},\qquad(76)$$

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

$$\begin{aligned} \max_{q_i(c_i)} \widetilde{W} &\equiv \sum_{i\neq j} \left[\widetilde{W}(\underline{c},\widehat{c}) + \widetilde{W}(\widehat{c},\overline{c})\right] \\ & s.t. \ (76), \end{aligned}$$

where $\widetilde{W}(\underline{c},\widehat{c})$ and $\widetilde{W}(\widehat{c},\overline{c})$ as defined in the proof of Lemma 6 are such that, beside (76), all other relevant constraints are satisfied. In particular, $\widetilde{W}(\underline{c},\widehat{c})$ is defined by (55) with c_{i1} replaced by $\widehat{c}$ and $\pi_{i,p}^{sb}(c_i,c_j) = -L$. For low and high types, the optimal quantities are still $q_i^I(c_i)$ and $q_i^V(c_i)$ respectively, as characterized by (15) and (19) in Lemma 6. However, such quantities do not satisfy (76) in a neighborhood of $\widehat{c}$. To see this, rewrite (76) as

$$\begin{aligned} q_i^{sb}(c_i) &\geq q_i^{sb}(\widehat{c}) + \int_{c_i}^{\widehat{c}} \left[q_i^{sb}(y_i) + K'(y_i)\right] \frac{df(\overline{c}\,|y_i)/dy_i}{f(\overline{c}\,|y_i)}, \ \forall c_i \in [\underline{c},\widehat{c}) \\ q_i^{sb}(c_i) &\leq q_i^{sb}(\widehat{c}) - \int_{\widehat{c}}^{c_i} \left[q_i^{sb}(y_i) + K'(y_i)\right] \frac{df(\underline{c}\,|y_i)/dc_i}{f(\underline{c}\,|y_i)}, \ \forall c_i \in (\widehat{c},\overline{c}]. \end{aligned}$$

In either inequality, the RHS approaches $q_i^{sb}(\widehat{c})$ for c_i close to $\widehat{c}$. Moreover, at the solution to (Γ), $q_i^{sb}(\widehat{c}) = q_i^{fb}(\widehat{c})$ and $q_i^I(c_i) < q_i^{fb}(c_i) < q_i^V(c_i)$ $\forall c_i \in [\underline{c},\overline{c}]$. Hence, output pooling arises around $\widehat{c}$. The pooling interval is unique for the same reasons as in Lewis and Sappington [6] (compare pp. 309-310 in their paper) and the proof is here omitted. Since the unique pooling interval includes $\widehat{c}$ and since $q_i^{sb}(\widehat{c}) = q_i^{fb}(\widehat{c})$, we have $q_i^{sb}(c_i) = q_i^{fb}(\widehat{c})$ $\forall c_i \in \left[c_i^-, c_i^+\right]$, where c_i^- is defined by $q_i^I(c_i^-) = q_i^{fb}(\widehat{c})$ and c_i^+ by $q_i^V(c_i^+) = q_i^{fb}(\widehat{c})$.