Ex Ante Collusion and Design of Supervisory Institutions[1]

Dilip Mookherjee, Alberto Motta and Masatoshi Tsumagari[2]

This Version: February 28 2014

Abstract

A Principal seeks to design a mechanism for an agent (with privately informed cost) and a supervisor/intermediary (with a noisy signal of the agent's cost) that collude on *both* participation and reporting decisions. Under weak belief restrictions which generalize the assumption of passive beliefs, the problem reduces to selecting weakly collusion-proof (WCP) allocations that satisfy interim participation constraints. We characterize WCP allocations, and use this to show that it is valuable to employ the supervisor. Delegation is optimal, but only if supplemented by an appeal/dispute settlement mechanism mediated by the Principal, which serves as an outside option for coalitional bargaining.

KEYWORDS: mechanism design, intermediation, supervision, collusion, delegation

[1]We thank Yeon-Koo Che, Bart Lipman and participants in the Theory Workshop at Boston University for useful comments.

[2]Boston University, University of New South Wales and Keio University, respectively

1 Introduction

The potential for collusion is widely acknowledged to be a serious problem in contexts where a Principal relies on information provided by an intermediary or supervisor to make key production and compensation decisions for an agent. Examples of such contexts abound: an investor that relies on an investment bank or rating agency for information necessary to decide on financing an entrepreneur; a firm owner that relies on a manager for information needed to set production targets and compensation for workers or suppliers; an employer that relies on a referee for a recommendation for a job applicant; a government that relies on a regulator to advise on rates for a public utility, or on an assessor to set tax rates for taxpayers. In all these settings the supervisor is better informed about the agent's productivity or cost than the Principal, but is typically less informed than the agent in this regard. Eliciting the supervisor's information becomes problematic when he is willing to misreport information in exchange for a bribe. The question arises whether it is still possible or desirable for the Principal to usefully elicit the supervisor's information. If so, how should the incentives for the supervisor and agent be designed? When can such mechanisms completely overcome the collusion problem? Does the presence of collusion rationalize a decentralized arrangement where the Principal does not contract directly with the agent, and instead delegates this authority to the supervisor (as in franchising arrangements)? Or does such decentralization need to be supplemented by some centralized safeguards?

This problem has been studied in the literature, on the basis of an assumption that supervisor and agent collude only with respect to reporting decisions, after they have independently committed to participate in the mechanism. This is referred to as *interim* collusion. Here it is presumed that collusion over *both* participation and reporting decisions — referred to as *ex ante* collusion — is not possible. In many contexts it would be difficult for the Principal to preclude collusion over participation decisions, e.g. where the supervisor and agent have a prior relationship or know each other before they respond to the Principal's contract offer. Even in situations where the supervisor and agent do not know each other previously, and the Principal can get them to independently commit to participating, there may be ways to elicit more information from them at the participation stage. Motta (2009) argues in a class of models of interim collusion that the problem of collusion can be costlessly overcome by such mechanisms.

The purpose of this paper is to study the problem with *ex ante* collusion. We consider a setting where the agent produces a divisible good at a constant unit cost whose realization is known to him privately, and the Principal and the supervisor have a prior over this cost which is continuously distributed over some interval. The supervisor updates this prior by being able to costlessly observe a noisy signal of the agent's cost, where the signal can take a finite number of possible realizations. The agent also observes the realization of this signal, so the coalition of supervisor and agent is characterized by one-sided asymmetric information. The supervisor and agent can enter into a side-contract which coordinates on joint participation and cost/signal reports to the Principal, as well as a private side payment, following a private cost report made by the agent to the supervisor. The side contract is designed and offered by the supervisor to the agent, though we subsequently show that our results extend to contexts where they are designed instead by a third party that maximizes a weighted sum of their interim payoffs. The side contract and the internal communication and transfers within the coalition are unobserved by the Principal. Nevertheless, the presence of asymmetric information frictions within the coalition allows some room for the Principal to manipulate the side contract. All parties are risk neutral. This setting is contrasted to early formulations of the collusion problem (e.g., Tirole (1986), Laffont and Tirole (1993)) by the absence of any hard information, and transaction costs of collusion that are entirely endogenous.

In such environments the Principal can in general design a grand contract (GC) for both the supervisor and agent, in which noncooperative play forms the outside option for bargaining within the coalition over a side contract. We show restrictions on beliefs within the coalition consequent on rejection of offered side contracts are needed for collusion to be costly for the Principal. Specifically, there exists beliefs that vary with the particular side contract offered, in such a way that collusion is effectively deterred.[3] Since the Principal cannot really control beliefs of the supervisor and agent, such solutions to the problem are unlikely to be be relevant. We therefore restrict beliefs to

[3] This requires rejection of any non-null side contract offer to be associated with the belief that the supervisor will subsequently exit, whence the agent is offered a highly lucrative contract by the Principal. On the other hand, if no such side contract is offered, the agent believes the supervisor will participate, and that they will play a noncooperative equilibrium of the grand contract recommended by the Principal. With such beliefs, the agent is motivated to reject any non-null side contract. Anticipating this, no such side contract is offered.

be independent of the side contract offered, or whether or not it is offered, which generalizes the assumption of *passive beliefs* employed in a number of previous papers (e.g., Laffont and Martimort (1997, 2000), Faure-Grimaud, Laffont and Martimort (2003)). Our analysis generalizes the passive belief assumption by allowing for side-contracts to be rejected by some types of agents on the equilibrium path, to address the problem recently highlighted by Celik and Peters (2011) that restrictions that prevent equilibrium-path rejections may entail a loss of generality.[4] We propose a notion of Weak Perfect Bayesian Equilibrium (WPBE(w)) which incorporates such a restriction. We show this restriction corresponds to the notion of *weak collusion* considered in much of the existing literature on collusion, in which the outside options for bargaining within the supervisor-agent coalition do not vary with side-contract offers, and are determined by a truthful noncooperative equilibrium of the grand contract designed by the Principal. By designing the grand contract which forms a backdrop to side contract negotiations, the Principal is able to exercise some control over the nature of resulting collusion. We show that attention can be restricted to allocations that (besides satisfying interim participation constraints for the supervisor and agent) are *weak-collusion-proof (WCP)*: which leave no room for profitable collusion, if offered as a direct revelation mechanism in the grand contract. We shall refer to such allocations as *implementable with weak collusion*. WCP allocations can be characterized by a set of individual incentive constraints (for the agent) and coalitional incentive and participation constraints.

We use this characterization to prove the following results:

(a) Delegation to the supervisor (DS) is strictly dominated by not appointing any supervisor (NS).[5] Hence delegation cannot be rationalized as an optimal response of the Principal to weak ex ante collusion. This can be contrasted to the optimality of delegation with interim collusion, when there are two possible types of the agent and two possible signals of the supervisor (Faure-Grimaud, Laffont and Martimort (2003)).[6]

[4] They show examples of collusion among members of a cartel where a proposed cartel agreement is rejected with positive probability. Such rejections serve to communicate information within the cartel.

[5] If side contracts are designed by a third party that maximizes a weighted sum of the supervisor and agent's payoffs, the same result applies for delegation to the third party, as long as the third party assigns a positive welfare weight to the supervisor's payoff. When the supervisor is assigned a zero welfare weight, DS turns out to be equivalent to NS.

[6] With weak ex ante collusion in the two-type-two-signal case, it can be shown that the

(b) Centralized contracting with the supervisor and agent (CS) strictly dominates NS, so it is valuable for the Principal to employ the supervisor and induce full revelation of information, despite ex ante collusion.

(c) Sufficient conditions are provided for collusion to be costly for the Principal: the support of the conditional cost distribution is independent of the supervisor's signal, conditional distributions satisfy suitable monotonicity and monotone likelihood properties, and the Principal's gross benefit function exhibits sufficient curvature.

(d) Any allocation that is implementable with weak collusion can be implemented by a modified form of delegation, in which the Principal communicates and transacts with only the supervisor *on the equilibrium path.* This corresponds to a 'normal' hierarchical delegation arrangement where the Principal asks the supervisor to initially communicate and transact with the agent, and then submit a joint participation decision and cost-cum-signal report to the Principal on behalf of the coalition. These reports determine an output target and aggregate payment for the coalition made by the Principal to the supervisor, who subsequently relays the output target and makes a corresponding out-of-pocket payment to the agent. The mechanism leaves open the room for the agent to trigger a switch to a centralized mechanism (the grand contract) where both agent and supervisor make independent reports to the Principal. This may be thought of as an 'appeals' or 'dispute settlement' procedure mediated by the Principal, which is not activated in equilibrium but plays a key role by determining outside options for coalition partners when they negotiate a side-contract.

(e) Given outside options determined in this manner, the allocation of bargaining power within the coalition does not affect the set of implementable allocations with weak collusion. This implies that the rules determining the matching of supervisors and agents are irrelevant — whether agents select supervisors (so agents have more bargaining power), or supervisors select agents (whence supervisors have more bargaining power), or whether they are externally assigned to one another (where they have equal bargaining power). This is an implication of weak collusion, where outside options are independent of bargaining

Principal never benefits from appointing a supervisor.

power. In strong collusion (Dequiedt (2007), Che and Kim (2009)) where the side contract designer can commit to playing the subsequent grand contract in suitable ways, outside options depend on the allocation of bargaining power, which thereby ends up affecting the set of implementable allocations.

Related literature on mechanism design with collusion can be classified by the context (auctions, team production or supervision), the nature of collusion (ex ante or interim, weak or strong collusion), and whether type spaces are discrete or continuous. Auctions and team production involve multiple privately informed agents and no supervisor. For auctions, Dequiedt (2007) considers strong ex ante collusion with binary agent types and shows that efficient collusion is possible, implying that the second-best cannot be implemented. In contrast, Pavlov (2008) considers a model with continuous types where the second-best can be implemented with weak ex ante collusion, and Che and Kim (2009) find the same result with either weak or strong ex ante collusion with continuous types. Team production with binary types is studied by Laffont and Martimort (1997), who show the second best can be implemented with weak interim collusion; this analysis is extended to a public goods context in Laffont and Martimort (2000) where the role of correlation of types is explored. Che and Kim (2006) show how second-best allocations can be implemented in a team production context with continuous types in the presence of weak interim collusion. Quesada (2004) on the other hand shows strong ex ante collusion is costly in a team production model with binary types. Mookherjee and Tsumagari (2004) show delegation to one of the agents is worse than centralized contracting in the presence of weak ex ante collusion. The logic of this result is similar to that underlying our result that delegation to the supervisor is worse than not appointing a supervisor. Their paper also considers delegation to a supervisor who is perfectly informed about the costs of each agent, and show that its value relative to centralized contracting depends on complementarity or substitutability between inputs supplied by different agents. The current paper differs insofar as there is only one agent, and there is asymmetric information within the supervisor-agent coalition owing to the supervisor receiving a noisy signal of the agent's cost. This friction in coalitional bargaining plays a key role in the current paper.

In the context of collusion between a supervisor and agent, existing models (with the exception of Mookherjee-Tsumagari (2004)) have explored interim collusion only. Faure-Grimaud, Laffont and Martimort (2003) consider a model with binary types and signals (with full support for conditional distributions), a risk-averse supervisor where collusion is costly, where (unconditional) delegation turns out be an optimal response to collusion. Celik (2009) considers a model with three types and two signals (where the support of conditional distributions depends on the signal), and risk neutral supervisor and agent, in which unconditional delegation is dominated by no supervision, which in turn is dominated strictly by centralized contracting with supervision. Celik's results are similar to ours, but he considers interim rather than ex ante collusion. Our results can be viewed as finding that the results he derived in the context of interim collusion with a special information structure happen to obtain quite generally with ex ante collusion and continuous types. The need to examine ex ante rather than interim collusion is highlighted by Motta (2009) who shows that collusion can be rendered costless in models with discrete type and signal spaces and interim collusion, by using mechanisms where the Principal offers a menu of contracts to the agent which the latter must respond to before colluding with the supervisor.

The paper is organized as follows. Section 2 introduces the model, followed by Section 3 which defines and characterizes WCP allocations. It also provides a game theoretic foundation for this notion, but this part of the paper can be skipped by those more interested in the main results concerning optimal allocations. The main results are presented in Section 4 for the polar model, where optimal allocations are always interior and the supervisor has all the bargaining power within the coalition. Section 5 then considers a number of extensions, where side contracts are designed and offered by a third party maximizing a weighted sum of supervisor and agent's payoffs, where the supervisor may exhibit altruism towards the agent, and where the Principal's gross benefit function is linear (whereby optimal allocations are never interior). Finally, Section 6 concludes with a summary, some applications of our results, and directions for future work.

2 Model

2.1 Environment

We consider an organization composed of a principal (P), an agent (A) and a supervisor (S). P can hire A who delivers an output $q \geq 0$ at a personal cost of θq. P's return from q is $V(q)$ where $V(q)$ is twice continuously differentiable, increasing and strictly concave satisfying $\lim_{q\to 0} V^{'}(q) = +\infty$, $\lim_{q\to+\infty} V^{'}(q) = 0$ and $V(0) = 0$. These conditions imply that $q^*(\theta) \equiv \arg_q \max V(q) - \theta q$ is continuously differentiable, positive on $\theta \in [0, \infty)$ and strictly decreasing. In Section 5.4 we shall describe how the results are modified when V is linear and subject to a capacity constraint.

θ is a random variable whose realization is privately observed by A. It is common knowledge that everybody shares a common prior $F(\theta)$ over θ on the interval $\Theta \equiv [\underline{\theta}, \bar{\theta}] \subset \Re_+$. F has a density function $f(\theta)$ which is continuously differentiable and everywhere positive on its support. The 'virtual cost' $H(\theta) \equiv \theta + \frac{F(\theta)}{f(\theta)}$ is assumed to be strictly increasing in θ.

S plays no role in production, and costlessly acquires an informative signal η about A's cost θ. The set of possible realizations of η is Π, a finite set with $\#\Pi \geq 2$. It is common knowledge that the realization of η is observed by both S and A. $a(\eta \mid \theta) \in [0, 1]$ denotes the likelihood function of η conditional on θ, which is common knowledge among all agents. $a(\eta \mid \theta)$ is continuously differentiable and positive on $\Theta(\eta)$, where $\Theta(\eta)$ denotes the set of values of θ for which signal η can arise with positive probability. We assume $\Theta(\eta)$ is an interval, for every $\eta \in \Pi$. Define $\underline{\theta}(\eta) \equiv \inf \Theta(\eta)$ and $\bar{\theta}(\eta) \equiv \sup \Theta(\eta)$. We assume that for any $\eta \in \Pi$, $a(\eta \mid \theta)$ is not a constant function on Θ, and there are some portions of θ with positive measure such that $a(\eta \mid \theta) \neq a(\eta^{'} \mid \theta)$ for any $\eta, \eta^{'} \in \Pi$. In this sense each possible signal realization conveys information about the agent's cost. The information conveyed is partial, since Π is finite.

The conditional density function and the conditional distribution function are respectively denoted by $f(\theta \mid \eta) \equiv f(\theta)a(\eta \mid \theta)/p(\eta)$ (where $p(\eta) \equiv \int_{\underline{\theta}(\eta)}^{\bar{\theta}(\eta)} f(\tilde{\theta})a(\eta \mid \tilde{\theta})d\tilde{\theta}$) and $F(\theta \mid \eta) \equiv \int_{\underline{\theta}(\eta)}^{\theta} f(\tilde{\theta} \mid \eta)d\tilde{\theta}$. The 'virtual' cost conditional on the signal η is $h(\theta \mid \eta) \equiv \theta + \frac{F(\theta|\eta)}{f(\theta|\eta)}$. We do not impose any monotonicity assumption for $h(\theta \mid \eta)$. Let $\hat{h}(\theta \mid \eta)$ be constructed from $h(\theta \mid \eta)$ and $F(\theta \mid \eta)$ by the ironing procedure introduced by Myerson (1981).

All players are risk neutral. P's objective is to maximize the expected value of $V(q)$, less expected payment to A and S, represented by X_A and X_S respectively. S's objective is to maximize expected transfers $X_S - t$ where t is a transfer from S to A. A seeks to maximize expected transfers received, less expected production costs, $X_A + t - \theta q$. Both A and S have outside options equal to 0.

In this environment, a feasible (deterministic) allocation is represented by $(u_A, u_S, q) = \{(u_A(\theta, \eta), u_S(\theta, \eta), q(\theta, \eta)) \in \Re^2 \times \Re_+ \mid (\theta, \eta) \in K\}$ where

$K \equiv \{(\theta, \eta) \mid \eta \in \Pi, \theta \in \Theta(\eta)\}$, u_S, u_A denotes S and A's payoff respectively, and q represents the production level. P's payoff equals $u_P = V(q) - u_S - u_A - \theta q$. These payoffs relate to transfers and productions as follows: $u_A \equiv X_A + t - \theta q; u_S \equiv X_S - t; u_P \equiv V(q) - X_S - X_A$.

2.2 Mechanism in the Absence of Collusion

Consider as a benchmark the case where A and S do not collude, and P designs contracts for both. We call this organization NC (no collusion). Owing to risk-neutrality of all parties and concavity of V, P can restrict attention to a deterministic grand contract:

$$GC = (X_A(m_A, m_S), X_S(m_A, m_S), q(m_A, m_S); M_A, M_S)$$

where M_A (resp. M_S) is a message set for A (resp. S). This mechanism assigns a deterministic allocation, i.e. transfers X_S, X_A and output q, for any message $(m_A, m_S) \in M_A \times M_S$. M_A includes A's exit option $e_A \in M_A$, with the property that $m_A = e_A$ implies $X_A = q = 0$ for any $m_S \in M_S$. Similarly M_S includes S's exit option $e_S \in M_S$, where $m_S = e_S$ implies $X_S = 0$ for any $m_A \in M_A$. The set of all possible deterministic grand contracts is denoted by $\mathcal{GC}$.

A grand contract induces a Bayesian game of incomplete information between A and S. Let $p(\eta)$ denote a set of beliefs held by S regarding the distribution of θ, in states where signal η has been realized. The posterior beliefs of S based on Bayesian updating of prior beliefs on the basis of observation of η alone are denoted by $p_\emptyset(\eta)$. The two can differ when S receives additional information, e.g., based on whether A agreed to participate.

Definition 1 *A Bayesian equilibrium of the game played by A and S in state η relative to beliefs $p(\eta)$ is a set of functions $c \equiv (m_A(\theta, \eta); m_S(\eta))$ (where m_A maps K into M_A, while m_S maps Π into M_S) such that the following conditions are satisfied for all $\theta \in [\underline{\theta}(\eta), \bar{\theta}(\eta)]$:*

$$m_A(\theta, \eta) \in \arg\max_{m_A \in M_A}[X_A(m_A, m_S(\eta)) - \theta q(m_A, m_S(\eta))] \qquad (1)$$

$$m_S(\eta) \in \arg\max_{m_S \in M_S} E_{p(\eta)}[X_S(m_A(\theta, \eta), m_S)] \qquad (2)$$

where $E_{p(\eta)}$ denotes expectation taken with respect to beliefs $p(\eta)$. $C(p(\eta); \eta)$ denotes the set of Bayesian equilibria corresponding to the beliefs $p(\eta)$ in state η.

The timing of events in NC is as follows.

(NC1) A observes θ and η, S observes η.

(NC2) P offers the grand contract $GC \in \mathcal{GC}$, and for any $\eta \in \Pi$ recommends a Bayesian equilibrium $c(p_\emptyset(\eta); \eta)$ relative to posterior beliefs $p_\emptyset(\eta)$ based on Bayesian updating by S on the basis of observation of η alone.

(NC3) A and S play the recommended Bayesian equilibrium.

The order of the timing between (NC1) and (NC2) can be interchanged without altering any of the results. If P offers a null contract to S (defined by the property that M_S is the empty set and $X_S = 0$), this is an organization without a supervisor, which we will denote by NS. Such an organization obviously leaves no scope for collusion between A and S.

It is well-known that in (NC) the Principal can restrict attention to direct revelation games, where M_A, M_S reduce to reports of private information, besides participation decisions. Define the *second-best allocation* $(u_A^{SB}, u_S^{SB}, q^{SB})$ as follows:

$$u_A^{SB}(\theta, \eta) = \int_\theta^{\bar{\theta}(\eta)} q^{SB}(y, \eta) dy,$$

$$E[u_S^{SB}(\theta, \eta) \mid \eta] = 0$$

and

$$q^{SB}(\theta, \eta) \equiv q^*(\hat{h}(\theta \mid \eta)) = \arg\max_q [V(q) - \hat{h}(\theta \mid \eta) q]$$

where $\hat{h}(\theta \mid \eta)$ is constructed from $h(\theta \mid \eta)$ and $F(\theta \mid \eta)$ by the ironing procedure. It is well-known that this is the optimal allocation in (NC). It turns out that in (NC) it is possible for the second-best to be implemented as a unique Bayesian equilibrium.[7]

2.3 Mechanism with Ex Ante Collusion

Now we describe the game played with (ex ante) collusion. Collusion takes the form of communication and side-contracting between A and S, which

[7] A proof is available on request.

takes place before they respond to P's offer of the grand contract (including participation decisions). This is distinguished from (interim) collusion where they do not collude on their participation decisions, but collude on the reports they send to P and enter into side payments in the event of joint participation.

The game with ex ante collusion is different from the game without collusion following stage NC2. At that point, A and S can enter into a side-contract in which A sends a message to S following which they jointly decide on participation, reporting and side-payments. The side-contract is unobserved by P. As in existing literature, we assume the side-contract is costlessly enforceable. Moreover we assume S has all the bargaining power *vis-a-vis* A: S can make a take-it-or-leave-it offer of a side-contract. This assumption turns out to be inessential: Section 5.2 explains how the same results obtain with side contracts offered by an uninformed third party that assigns arbitrary welfare weights to the supervisor and agent.

After S offers the side contract, A retains the option of rejecting it; given that A's true cost is not known to S, this still enables A to earn some rents. This information friction within the coalition plays a key role in our analysis. If A rejects the side contract, they subsequently play the game associated with the grand contract non-cooperatively. If A were to reject the side-contract, some information regarding the realization of θ may be communicated to S. Hence their subsequent noncooperative play of the grand contract would be a Bayesian equilibrium relative to beliefs which could differ from $p_{\emptyset}(\eta)$.

The game in the presence of collusion replaces (NC3) above (while (NC1) and (NC2) are unchanged) by the following three-stage subgame (conditional on any $\eta \in \Pi$):

(i) S offers a side-contract SC which determines for any $\tilde{\theta} \in \Theta(\eta)$ to be privately reported by A to S, a probability distribution over joint messages $(m_A, m_S) \in M_A \times M_S$, and a side payment from S to A.[8] Formally, it is a pair of functions $\{\tilde{m}(\tilde{\theta}, \eta), t(\tilde{\theta}, \eta)\}$ where $\tilde{m}(\theta, \eta) : \Theta(\eta) \times \{\eta\} \longrightarrow \Delta(M_A \times M_S)$, the set of probability measures over $M_A \times M_S$, and $t : \Theta(\eta) \times \{\eta\} \longrightarrow \Re$. The case where S does not offer a side contract is represented by a null side-contract (NSC) with zero side payments ($t(\theta, \eta) \equiv 0$), and (deterministic) messages $(m_A(\theta, \eta); m_S(\eta))$ the same as those in the Bayesian equilibrium of the grand contract rec-

[8]The option of randomizing over possible messages is useful for technical reasons. Owing to quasilinearity of payoffs, there is no need to randomize over side transfers.

ommended by the Principal. We abuse terminology slightly and refer to the situation where no side contract is offered as one where NSC is offered.

(ii) A either accepts or rejects the SC offered, and the game continues as follows.

(iii) If A accepts the offered SC, he sends a private report $\theta' \in \Theta(\eta)$ to S, following which the SC is executed. If A rejects SC, S updates his beliefs to $p(SC; \eta)$ which is restricted to be $p_\emptyset(\eta)$ if NSC was offered in stage (i) above.[9] A and S then play a Bayesian equilibrium c of the grand contract relative to beliefs $p(SC; \eta)$.

The notion of Weak Perfect Bayesian Equilibrium (WPBE) of the game with collusion requires beliefs and continuation strategies to be specified corresponding to all information sets of the game.[10] Hence we need to specify how these are decided off the equilibrium path. On the equilibrium path, post-rejection beliefs are obtained upon applying Bayes rule if the SC offered in equilibrium is rejected with positive probability by A. But there is a lot of leeway in beliefs following rejection of off-equilibrium-path side contract offers. The specific notion of collusion depends critically on how such beliefs are formed, since these determine A's outside options in bargaining with S over a side contract. In the next section we shall impose a weak restriction on beliefs, and show that it corresponds to a notion of weak collusion proof (WCP) allocations.

It is worth mentioning that in the presence of ex ante collusion, the Principal can do no better than to confine attention to a static revelation mechanism when selecting a grand contract. We do not provide a formal proof, as the argument is a familiar one.

3 Weak Collusion Proof Allocations

3.1 Definition of WCP

We first provide a definition of weak collusion proofness. The notion can be explained quite simply: an allocation is weakly collusion proof if the

[9] This ensures that it is immaterial whether or not NSC was accepted or rejected, since in either case they play the grand contract non-cooperatively with prior beliefs.

[10] For definition of WPBE, see Mas-Colell, Whinston and Green (1995, p.285).

supervisor cannot benefit from offering a non-null side contract when the Principal selects a grand contract based on the associated direct revelation mechanism (i.e., when agent and supervisor make consistent reports about the state, the allocation corresponding to that state is chosen). This requires the null side contract to be the optimal side contract for S, when the outside option of A corresponds to his payoff resulting from the allocation.

Before proceeding to the formal definition, note that a *deterministic allocation* can be represented by payoff functions $(u_A(\theta,\eta), u_S(\theta,\eta))$ of the true state (θ,η) combined with the output function $q(\theta,\eta)$, as these determine the Principal's payoff function $u_P(\theta,\eta) \equiv V(q(\theta,\eta)) - u_S(\theta,\eta) - u_A(\theta,\eta) - \theta q(\theta,\eta)$, and the aggregate net transfers of S (equals $u_S(\theta,\eta)$) and A (equals $u_A(\theta,\eta) + \theta q(\theta,\eta)$). For technical convenience we consider randomized allocations, though it will turn out they will never actually need to be used on the equilibrium path. In a randomized allocation, $(u_A(\theta,\eta), u_S(\theta,\eta), q(\theta,\eta))$ denotes the expected payoffs of A, S and the expected output, conditional on the state (θ,η). For (conditional expected) allocation $(u_A(\theta,\eta), u_S(\theta,\eta), q(\theta,\eta))$, define functions $(\hat{X}(m), \hat{q}(m))$ on domain $m \in \hat{M} \equiv K \cup \{e\}$ (where $K \equiv \{(\theta,\eta) \mid \theta \in \Theta(\eta), \eta \in \Pi\}$) as follows:

$$(\hat{X}(\theta,\eta), \hat{q}(\theta,\eta)) = (u_A(\theta,\eta) + \theta q(\theta,\eta) + u_S(\theta,\eta), q(\theta,\eta))$$

$$(\hat{X}(e), \hat{q}(e)) = (0,0)$$

$(\hat{X}(\theta,\eta), \hat{q}(\theta,\eta))$ denote corresponding expected values of the sum of payments $X_S + X_A$ made by the principal, and the output delivered, in state θ,η. Also, let $\Delta(\hat{M})$ denote the set of the probability measures on $\hat{M}$, and use $\tilde{m} \in \Delta(\hat{M})$ to denote a randomized message submitted by the coalition to P. With a slight abuse of notation, we shall denote the corresponding conditional expected allocation by $(\hat{X}(\tilde{m}), \hat{q}(\tilde{m}))$, which is defined on $\Delta(\hat{M})$. $\tilde{m} = (\theta,\eta)$ or e will be used to denote the probability measure concentrated at (θ,η) or e respectively.

S's choice of an optimal (randomized) side-contract can be formally posed as follows. For any $\eta \in \Pi$, the side-contracting problem $P(\eta)$ is to select $(\tilde{m}(\theta \mid \eta), \tilde{u}_A(\theta,\eta))$ to maximize S's expected payoff

$$E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta,\eta) \mid \eta]$$

subject to $\tilde{m}(\theta \mid \eta) \in \Delta(\hat{M})$,

$$\tilde{u}_A(\theta,\eta) \geq \tilde{u}_A(\theta',\eta) + (\theta' - \theta)\hat{q}(\tilde{m}(\theta' \mid \eta))$$

for any $\theta, \theta' \in \Theta(\eta)$, and

$$\tilde{u}_A(\theta, \eta) \geq u_A(\theta, \eta)$$

for all $\theta \in \Theta(\eta)$. The first constraint states truthful revelation of the agent's true cost to S is consistent with the agent's incentives, and the second constraint requires A to attain a payoff at least as large as what he would expect to attain by playing the grand contract noncooperatively. Recall that the grand contract must satisfy $X_A(e_A, m_S) = q(e_A, m_S) = 0$ for any m_S and $X_S(m_A, e_S) = 0$ for any m_A. Hence S has the option of shutting down production altogether, implying that the expected payoff earned by S in the problem above is non-negative, conditional on any η.

Let the maximum payoff of S in the side contracting problem in state η be denoted by $V(\eta)$.

Definition 2 *The (conditional expected) allocation* $(u_A(\theta, \eta), u_S(\theta, \eta), q(\theta, \eta))$: $K \to \Re^2 \times \Re_+$ *is weakly collusion proof (WCP) if for every* $\eta \in \Pi$*:* $(\tilde{m}(\theta \mid \eta), \tilde{u}_A(\theta, \eta)) = ((\theta, \eta), u_A(\theta, \eta))$ *solves problem* $P(\eta)$ *in which S achieves a maximum payoff of* $V(\eta) = E[u_S(\theta, \eta) \mid \eta]$*.*

3.2 Characterization of WCP Allocations

We now characterize WCP allocations. This requires us to define a family of 'modified' virtual cost functions, representing the effective cost incurred by the coalition in delivering a unit of output to P, following selection of an optimal side-contract.

Definition 3 *For any* $\eta \in \Pi$*,* $Y(\eta)$ *is a collection of* **coalitional shadow cost (CSC)** *functions* $\pi(\cdot \mid \eta) : \Theta(\eta) \to \Re$ *which satisfy the following property. For any function in this collection, there exists a real-valued function* $\Lambda(\theta, \eta)$ *which is non-decreasing in* $\theta \in \Theta(\eta)$ *with* $\Lambda(\underline{\theta}(\eta) \mid \eta) = 0$ *and* $\Lambda(\bar{\theta}(\eta) \mid \eta) = 1$*, such that*

$$\pi(\theta|\eta) \equiv \theta + \frac{F(\theta \mid \eta) - \Lambda(\theta \mid \eta)}{f(\theta \mid \eta)} \tag{3}$$

Equation (3) modifies the usual expression for virtual cost $h(\theta|\eta) \equiv \theta + \frac{F(\theta|\eta)}{f(\theta|\eta)}$ by subtracting from it the non-negative term $\frac{\Lambda(\theta|\eta)}{f(\theta|\eta)}$. Intuitively, with collusion between S and A, it is as if P procures the good from a single entity, consisting of the coalition of S and A. If A's outside option payoff in the side-contracting problem were 0 instead of $u_A(\theta, \eta)$, S would incur a cost of $h(\theta|\eta)$ in arranging for delivery of one unit of the good. P's problem of procuring the good would then reduce to contracting with a single agent with an unknown cost of $h(\theta|\eta)$. This is worse for P compared with the situation where there is no supervisor at all — in the latter context, P would be contracting with A alone who incurs a cost of θ rather than $h(\theta|\eta)$. This is the well-known problem of double marginalization of rents (DMR), arising due to exercise of monopsony power by S in side-contracting with A. As elaborated later, this is why delegating the right to contract (with A) to S cannot result in any improvement for P compared to the situation where no S is employed.

To limit DMR, P contracts with both S and A, and provides A with an outside option (of $u_A(\theta, \eta)$) that effectively raises his bargaining power *vis-a-vis* S while negotiating the side contract. Meeting a larger outside option for A effectively induces S to deliver a higher output to P: this is what paying a higher rent to A necessitates. The extent of DMR is then curbed: the shadow cost for the coalition in delivering a unit of output to P is lowered. This lowering of the virtual cost is represented by the subtraction of the term $\frac{\Lambda(\theta|\eta)}{f(\theta|\eta)}$ from what it would have been $h(\theta|\eta)$ under delegated contracting. The derivative of $\Lambda(\theta \mid \eta)$ represents the Kuhn-Tucker multiplier on A's (type θ) participation constraint in S's problem of selecting an optimal side contract. Since the multiplier is non-negative, the $\Lambda(\theta \mid \eta)$ function is non-decreasing.

However, $\pi(\theta|\eta)$ is not the correct expression for the shadow cost of output for the coalition, if it is non-monotone in θ. In that case, it has to be replaced by its 'ironed' version (Myerson (1981)), using the distribution function $F(\theta|\eta)$. Let the corresponding ironed version of $\pi(\theta|\eta)$ be denoted by $z(\theta|\eta)$: we call this a *coalitional virtual cost function.*

Definition 4 *For any $\eta \in \Pi$, the set of* **coalitional virtual cost (CVC)** *functions is the set*

$$Z(\eta) \equiv \{z(\cdot \mid \eta) \quad \textit{ironed version of some} \quad \pi(\cdot \mid \eta) \in Y(\eta)\}.$$

of functions obtained by applying the ironing procedure to the set $Y(\eta)$ of CSC functions.[11] *Denote by $\Theta(\pi(\cdot \mid \eta), \eta)$ the corresponding pooling region of θ where $\pi(\cdot|\eta)$ is flattened by the ironing procedure.*

[11]The ironing procedure ensures these functions are continuous and non-decreasing.

As the next result shows, every WCP allocation satisfies coalitional participation and incentive constraints corresponding to some coalitional virtual cost function z. Combined with an individual incentive compatibility constraint for A, and a constraint that output must be constant over regions of constancy of z, these coalitional constraints characterize WCP allocations.

Proposition 1 *The allocation (u_A, u_S, q) is WCP if and only if the following conditions hold for every η. There exists a CVC function $z(\cdot|\eta) \in Z(\eta)$ such that*

(i) For every $(\theta, \eta), (\theta', \eta') \in K \equiv \{(\theta, \eta) \mid \theta \in \Theta(\eta), \eta \in \Pi\}$,

$$X(\theta, \eta) - z(\theta \mid \eta)q(\theta, \eta) \geq X(\theta', \eta') - z(\theta \mid \eta)q(\theta', \eta')$$

$$X(\theta, \eta) - z(\theta \mid \eta)q(\theta, \eta) \geq 0$$

where

$$X(\theta, \eta) \equiv u_A(\theta, \eta) + u_S(\theta, \eta) + \theta q(\theta, \eta)$$

(ii) For any $\theta, \theta' \in \Theta(\eta)$,

$$u_A(\theta, \eta) \geq u_A(\theta', \eta) + (\theta' - \theta)q(\theta', \eta)$$

(iii) $q(\theta, \eta)$ is constant on any interval of θ which is a subset of the corresponding pooling region of the CVC function z.

Conditions (i) and (ii) represent the coalitional incentive and participation constraints corresponding to contracting with a single agent with a unit cost of z. Condition (ii) is the individual incentive compatibility constraint for A. Condition (iii) states that the output must be constant over every interval in the pooling region.

In the rest of this section, we shall provide the game-theoretic foundation for focusing attention on WCP allocations. It can be skipped by those interested in our main results concerning the principal's mechanism design problem, where the outcome of such mechanisms are described by resulting WCP allocations that satisfy interim participation constraints.

3.3 Belief Restrictions Underlying WCP Allocations

We first explain the need for restricting off-equilibrium-path beliefs. In the absence of any such restrictions, we can use the notion of a Weak Perfect Bayesian equilibrium (WPBE) of the game with collusion. If the mechanism design problem is stated as selection of an allocation subject to the constraint that it can be implemented as the outcome of some WPBE, it is presumed that the Principal is free to select continuation beliefs and strategies for noncooperative play of the grand contract following off-equilibrium path rejections of offered side contracts by S to A.

Definition 5 *The organization with* **ultra-weak collusion (UWC)** *is the game in which stage (NC3) following any $\eta \in \Pi$ is replaced by the three-stage collusion game described above, in which the beliefs $p(SC;\eta)$ and Bayesian equilibrium $c(SC;\eta)$ played following off-equilibrium-path rejection of any SC can be specified by P (subject to the constraint that $c(SC;\eta) \in C(p(SC);\eta)$).*

Collusion can then be deterred completely, with appropriate selection of off-equilibrium-path continuations. A heuristic description of how the second-best payoff can be achieved by the Principal as a weak PBE is as follows. P selects a grand contract and recommends a noncooperative equilibrium of this contract in which (i) conditional of participation by S, noncooperative play results in the second-best allocation; (ii) S is paid nothing; and (iii) if S does not participate, P offers A a 'gilded' contract providing the latter a high payoff in all states. On the equilibrium path S always offers a null side contract. If A rejects any offer of a non-null side-contract, they mutually believe that subsequently S will not participate in the grand contract, and A will receive the gilded contract. This forms a weak PBE as rejection of any non-null side contract is sequentially rational for A given A's belief that S will exit following any rejection. And exit is sequentially rational for S given his belief that A will reject the side contract and they will subsequently play the grand contract noncooperatively where S will be paid nothing. Essential in this equilibrium is the way that continuation beliefs and strategies following the null side contract (where S does not exit) differ from those following any non-null side contract offer.

Proposition 2 *The second-best allocation can be implemented in (UWC).*

The assumption made often in the literature (e.g., Faure-Grimaud, Laffont and Martimort (2003)) of *passive beliefs* requires beliefs and continuation play following rejection of side-contract offers to not vary with the side-contract offered, and for the beliefs to be the same as the posterior beliefs resulting from Bayesian updating based on observation of η alone. Such a restriction rules out implementation of the second-best along the lines of Proposition 2.

Faure-Grimaud, Laffont and Martimort (2003), however, restrict attention to side contracts offered that are always accepted by A on the equilibrium path. Celik and Peters (2011) have shown in the context of a model of a two-firm cartel that this restriction may entail a loss of generality. In contrast to a standard principal-agent setting where agent outside options are exogenous, the consequences of rejection of a side-contract subsequently results in A and S playing a noncooperative game and are thus endogenous. Rejection of a side contract by some types of A can communicate information to S about A's type, affecting subsequent play and resulting payoffs in the noncooperative game. Celik and Peters show that there can be collusive allocations amongst cartel members which can only be supported by side-contract offers which are rejected with positive probability on the equilibrium path.

We thus allow for side contract offers that might be rejected by some types of A and accepted by others. This is combined with the following restriction on beliefs.

Definition 6 *A WPBE(w) is a Weak Perfect Bayesian Equilibrium (WPBE) satisfying the following restriction on beliefs (conditional on realization of any η): (a) there is a pair of beliefs $p(\eta)$ and Bayesian equilibrium $c(\eta) \in C(p(\eta); \eta)$ which results in the noncooperative play of the grand contract following rejection of* **any** *non-null side contract offered by S, where (b) $(p(\eta), c(\eta)) = (p_\emptyset(\eta), c_\emptyset(\eta))$ if S offers a null side-contract on the equilibrium path.*

Definition 7 *An organization with* **weak collusion (WC)** *is the game in which attention is restricted to WPBE(w) outcomes of the continuation game following choice of a grand contract by P in stage (NC2).*

Criterion (a) imposes the restriction that there is a common continuation belief and Bayesian equilibrium of the grand contract, following rejection of *any* non-null side-contract.[12] Criterion (b) additionally requires this continuation to be the same as the continuation that results when S offers a null side-contract on the equilibrium path.[13] In this case, the consequences of rejection are independent of the side contract offered, and are taken as given by the Principal.

One could argue that it would be reasonable to expand the scope of (b) and also require $(p(\eta), c(\eta)) = (p_\emptyset(\eta), c_\emptyset(\eta))$ whenever a non-null SC is offered and accepted by all types of A on the equilibrium path. Evidently, the definition of WPBE(w) is consistent with this stronger version of (b). However, it is not needed for the results that follow. The Faure-Grimaud, Laffont and Martimort (2003) assumption of passive beliefs (where rejection of any offered SC is followed by beliefs $(p_\emptyset(\eta), c_\emptyset(\eta))$) is therefore consistent with WPBE(w). Their approach can be rationalized by an underlying restriction to side contract offers that are either accepted by all types, or rejected by all types. So WPBE(w) may be viewed as a generalization of the assumption of passive beliefs, when one allows rejection of SCs by some types on the equilibrium path (in order to address the Celik-Peters criticism).

We now show that with this restriction on beliefs, the Celik-Peters criticism can be addressed: there is no loss of generality in confining attention to side-contract offers that are accepted by all types on the equilibrium path.

Lemma 1 *Given any grand contract, and any allocation resulting from a WPBE(w) in which S's side contract offer is rejected with positive probability on the equilibrium path, there exists another WPBE(w) resulting in the same allocation in which the side contract offered by S is accepted by all types of A on the equilibrium path.*

The argument resembles the standard one underlying the Revelation Principle: offering a new side-contract $\tilde{SC}$ which mimics the outcomes resulting from rejection of an original side-contract (SC), can result in acceptance by all types of A and the same resulting allocation. How can this be reconciled with the Celik-Peters (2013) demonstration of a collusive allocation for

[12]This is irrespective of whether or not this rejection occurs on the equilibrium path. If it does, whereby subsequent continuation beliefs are determined by Bayes Rule, (a) requires the same beliefs to ensue from rejection of some other non-null SC.

[13]Criterion (a) by itself is insufficient to allow collusion to have any bite, since the construction used in proving Proposition 2 satisfied (a). Hence part (b) is additionally required to avoid the conclusion of Proposition 2.

a two-firm cartel which is the outcome of a side-contract that is rejected with positive probability in equilibrium, which cannot be achieved by some other side contract that is not rejected on the equilibrium path? There are two main differences between our respective formulations of side-contracting. First, in our model S rather than some third-party offers the side-contract. In the latter case, a participation constraint for S has to be respected. In our model S offers the SC, so there is no need to incorporate a participation constraint for S. However this difference would disappear in the version of our model to be considered later, where side contracts are designed and offered by a third party. The second reason is the WPBE(w) restriction we have imposed. The construction of the example in Celik-Peters (2013) hinges on beliefs following rejection that vary with the side-contract in question, contrary to what WPBE(w) requires.[14]

The next step is to observe that the *collusion-proofness principle* — which states that P can do no better than to restrict attention to noncooperative equilibria of grand contracts that do not provide S with an incentive to offer a non-null side contract — holds for organizations with weak collusion. This simplifies the analysis of P's problem of designing a mechanism for such an organization.

Lemma 2 *An allocation (u_A, u_S, q) is a WPBE(w) outcome if and only if there exists a grand contract GC satisfying the following two properties:*

(i) In any state $\eta \in \Pi$: participation and truthful reporting by all types of S and A constitutes a Bayesian equilibrium relative to beliefs $p_\emptyset(\eta)$ obtained by updating on η alone, which results in state-η allocation: $(u_A(\cdot, \eta), u_S(\cdot, \eta), q(\cdot, \eta))$;

[14]To elaborate further, their example rests on the following feature. Rejection of the side contract analogous to our $\tilde{SC}$ (by the uninformed party) results in coalition members playing the grand contract noncooperatively with beliefs $p_\emptyset$, whereas rejection of the equilibrium side contract is followed by noncooperative play with a different set of beliefs (the same as on the equilibrium path p^*). If the two side contracts were associated with the same post-rejection continuation beliefs, the argument underlying Lemma 1 would apply, implying that the $\tilde{SC}$ contract would support the same allocation as the equilibrium side-contract. Their construction is based on the implicit assumption that the designer of the side-contract will disclose information regarding the type reported by the other party for some side contracts (e.g., the equilibrium side contract), and not others (e.g., $\tilde{SC}$) when a given party is the only one to reject the side contract. However, it is not clear whether or on what grounds this assumption can be justified.

(ii) there is a WPBE(w) of the resulting side-contracting game in which S offers no side-contract for any $\eta \in \Pi$.

The argument is straightforward. Lemma 1 ensures that without loss of generality attention can be focused on WPBE(w) in which the equilibrium side contract, if offered in any state η, is not rejected by any type of A. Then there is no room for further coordination by S and A which improves the expected payoff of S while meeting A's acceptance and incentive constraint. If the resulting allocation were offered directly in the grand contract, there would be no scope for S to benefit from any further side-contract.

Lemma 2 implies that implementable allocations in the game with weak collusion coincide with WCP allocations satisfying interim participation constraints for both A and S.

Proposition 3 *An allocation (u_A, u_S, q) is implementable in the weak collusion game, if and only if it is a WCP allocation satisfying interim participation constraints*

$$E[u_S(\theta, \eta)|\eta] \geq 0 \quad \textit{for all} \quad \eta \tag{4}$$

$$u_A(\theta, \eta) \geq 0 \quad \textit{for all} \quad (\theta, \eta) \tag{5}$$

4 Main Results

We are now in a position to present our main results. In this section we will compare the following organizational alternatives:

(a) *No Supervisor (NS):* where P does not employ S and contracts with A alone on the basis of his own prior information F over A's cost θ. This is a special case of the preceding model where $X_S \equiv 0, M_S \equiv \emptyset$ in the grand contract. It is well known that P attains an expected profit of $E[V(q^{NS}(\theta)) - H(\theta)q^{NS}(\theta)]$ where $q^{NS}(\theta)$ is defined by the property $V'(q^{NS}(\theta)) = H(\theta)$. We shall denote this profit by Π_{NS}.

(b) *Delegated Supervision (DS):* Here P contracts with S alone, and delegates to S the authority to contract with A and make production decisions. It is a special case of the preceding model where $X_A \equiv 0, M_A \equiv \emptyset$ in the grand contract. S enters into a side-contract with A, and then responds to P's contract offer with a message regarding the joint realization of θ and η, or some summary of the two variables. Here A has no outside option of rejecting the side contract and participating in the grand contract, which increases the bargaining power of S with A. We shall denote the resulting profit of P by Π_{DS}.

(c) *Centralized Supervision (CS):* This is the unrestricted version of the model considered so far, where P offers a grand contract involving both S and A. A now has an outside option of rejecting the side contract offered by S and participating in the grand contract noncooperatively. We shall denote the resulting profit of P by Π_{CS}.

We will also assess these relative to the benchmark of no collusion, which is associated with the second-best allocation defined previously. The associated profit will be denoted Π_{SB}. Since S has access to information about A's cost that is valuable in contracting with A, it is obvious that $\Pi_{NS} < \Pi_{SB}$, i.e., hiring S is valuable if there is no collusion. We now compare the three alternatives above against one another, and will subsequently assess them relative to the second-best.

Proposition 4 $\Pi_{DS} < \Pi_{NS}$: *delegated supervision is worse for the Principal compared to hiring no supervisor.*

The result of Faure-Grimaud, Laffont and Martimort (2003) therefore does not extend to the setting of our model with ex ante collusion, risk neutrality and continuous types. The intuitive reason is simple. Ex ante collusion implies that in contracting with P, the supervisor is subject to an ex post participation constraint: he can accept or reject the contract offered by P *after* he has learnt the realization of A's cost θ. This results in *double marginalization of rents (DMR)*: A earns rents owing to his private information regarding θ with respect to S, and then S earns rents owing to his private information regarding his costs of procuring from A (which depend on the realizations of θ and η). In DS, the Principal effectively contracts with a single agent whose unit cost equals $\hat{h}(\theta|\eta)$ which is the ironed version of $h(\theta|\eta) \equiv \theta + \frac{F(\theta|\eta)}{f(\theta|\eta)}$, who can decide whether to participate after observing the realization of his unit cost. Since $h(\theta|\eta) > \theta$ almost everywhere (which implies the same is true for its ironed version $\hat{h}(\theta|\eta)$), delegated supervision amounts to contracting with a single supplier whose cost is uniformly higher, compared to contracting with the agent alone in the absence of the supervisor.

While it is relatively easy to show that DS cannot dominate NS, the proof establishes the stronger result that DS is **strictly** dominated by NS.[15]

Proposition 5 $\Pi_{NS} < \Pi_{CS}$*: the Principal is strictly better off hiring S and contracting directly with both S and A, compared to hiring no supervisor.*

This states that P always benefits from hiring S despite the presence of ex ante collusion between S and A. Combining with the previous result, it follows that S is valuable only provided P does not delegate authority to S: it is essential that P contracts simultaneously with A as well, thus providing A an outside option which raises A's bargaining power within the coalition. This limits the DMR problem by countervailing S's tendency to behave monopsonistically with respect to A. By raising A's outside option, the coalitional virtual cost z is reduced, allowing an increase in output delivered, and raising P's expected payoff.

This helps explain how contracting directly with both S and A helps reduce the DMR problem inherent in DS which rendered it inferior to NS. However, it does not help explain why it manages to do so sufficiently that CS ends up being superior to NS. The explanation for this is more subtle, arising from P's ability to profitably utilize S's superior information concerning the agent's cost with a simple mechanism. This arises ultimately from the discrepancy between relative likelihoods of different cost states by P and S, which they use to weight different states in computing their respective payoffs.

It may help to outline the WCP allocation that can be used by P. Starting with the optimal allocation in NS (which corresponds to the special case of CS where $\Lambda(\theta \mid \eta)$ is chosen equal to $F(\theta \mid \eta)$, ensuring that the CSC and CVC functions both reduce to the identity function ($\pi(\theta|\eta) = z(\theta|\pi(\cdot|\eta), \eta) = \theta$)), P can construct a small variation in the CVC function z in some state η^*, raising it above θ for some interval Θ_H and lowering it for some other interval Θ_L, both of which have positive probability given η^*. The corresponding quantity procured $q(\theta, \eta^*)$ is set equal to $q^{NS}(z(\theta|\eta^*))$, the quantity procured in NS when the agent reported a cost of $z(\theta|\eta^*)$. This corresponds to raising

[15] The proof of strict domination is also straightforward in the case that $h(\theta|\eta)$ is continuous and nondecreasing in θ over a common support $[\underline{\theta}, \bar{\theta}]$ for every η. In that case an argument based on Proposition 1 in Mookherjee and Tsumagari (2004) can be applied. In the general case there are a number of additional technical complications, but the result still goes through.

the quantity procured from the coalition over Θ_L and lowering it over Θ_H. Payments to the coalition are set analogously at $X^{NS}(z(\theta|\eta^*))$, what the agent would have been paid in NS following such a cost report.[16] The agent is offered the associated rent: $u_A(\theta, \eta^*) = \int_{\theta}^{\bar{\theta}} q^{NS}(z(y|\eta^*))dy$. By construction, this allocation satisfies the agent's incentive and participation constraints, as well as the coalitional incentive constraint.[17]

Proposition 1 ensures such an allocation is WCP, provided S's interim participation constraint is satisfied. The variation over Θ_L lowers rents earned by S, and over Θ_H raises them. Since S does not earn any rents to start with (i.e, in NS), it is necessary to construct the variation such that S's expected rents in state η^* do not go down. The rate at which S's rents vary locally in state θ with the quantity procured equals $\frac{F(\theta|\eta^*)}{f(\theta|\eta^*)}$.[18] Intuitively this is the saving that can be pocketed by S when procuring one less unit of the good from A. Maintaining S's expected rent therefore implies a marginal rate of substitution between output variations over Θ_L and Θ_H that equals the ratio of the (average) conditional inverse hazard rates $\frac{F(\theta|\eta^*)}{f(\theta|\eta^*)}$ over these two intervals respectively.

On the other hand, P's benefit from a small expansion in output delivered in state θ equals $V'(q^{NS}(\theta)) - \theta$, where $q^{NS}(\theta)$ denotes the optimal allocation in NS.[19] This allocation satisfies $V'(q^{NS}(\theta)) = H(\theta) \equiv \theta + \frac{F(\theta)}{f(\theta)}$, the virtual cost of procurement without conditioning on information regarding η. Hence P's marginal benefit from output expansion in state θ equals the unconditional inverse hazard rate $\frac{F(\theta)}{f(\theta)}$. This implies that P's marginal rate of substitution between output variations over Θ_L and Θ_H equals the ratio of the (average) unconditional inverse hazard rates $\frac{F(\theta)}{f(\theta)}$ over these two

---

[16]Specifically, $X^{NS}(z(\theta|\eta)) = z(\theta|\eta)q^{NS}(z(\theta|\eta)) + \int_{z(\theta|\eta)}^{\bar{\theta}} q^{NS}(y)dy$.

[17]This requires checking that there exists a CSC function $\pi(\theta|\eta)$ corresponding to some distribution function $\Lambda(\cdot \mid \eta)$ on $[\underline{\theta}(\eta), \bar{\theta}(\eta)]$ such that $z(\theta \mid \eta)$ is the ironed version of $\pi(\theta \mid \eta)$. This is true, since we can select $\Lambda(\theta \mid \eta) = (\theta - z(\theta \mid \eta))f(\theta, \eta) + F(\theta, \eta)$, which is strictly increasing over Θ_L and Θ_H for a sufficiently small variation of z from the identity function. Then $\Lambda(\cdot \mid \eta)$ is a distribution function, which generates $\pi(\theta|\eta) = z(\theta \mid \eta)$ since $z(\theta \mid \eta)$ is a non-decreasing function.

[18]S's interim rent in state η equals the expected value conditional on η of $X^{NS}(z(\theta|\eta)) - u_A(z(\theta|\eta)) - \theta q^{NS}(z(\theta|\eta))$, i.e., equals $E[\{(z(\theta|\eta)) - h(\theta|\eta)\}q^{NS}((z(\theta|\eta)) - \int_{z(\theta|\eta)}^{\bar{\theta}} q^{NS}(z)dz|\eta]$.

[19]This follows from the fact that $\frac{\partial X^{NS}(z(\theta|\eta))}{\partial z} = z(\theta \mid \eta)q^{NS'}(z(\theta \mid \eta))$, implying that the marginal increase in payment evaluated at $z(\theta, \eta) = \theta$ equals θ times the marginal output change.

intervals. The informativeness of S's signals implies that P's marginal rate of substitution differs from S's in some state η^* over some pair of intervals Θ_L, Θ_H. Hence there exist variations of the type described above which raise P's expected payoff, while preserving the expected payoff of S.

The next question is whether it may be possible for P to attain the second-best payoff using a WCP mechanism. The following result provides a set of sufficient conditions when this is not possible.

Proposition 6 $\Pi_{CS} < \Pi_{SB}$: *P cannot attain the second-best payoff in CS if the following conditions hold:*

(i) The support of θ *does not vary with the signal:* $\Theta(\eta) = \Theta$ *for any* $\eta \in \Pi$*;*

(ii) there exists $\eta^* \in \Pi$ *such that* $f(\theta|\eta^*)$ *and* $\frac{f(\theta|\eta^*)}{f(\theta|\eta)}$ *are both strictly decreasing in* θ *for any* $\eta \neq \eta^*$*; and*

(iii) $V'''(q) \leq \frac{(V''(q))^2}{V'(q)}$ *for any* $q \geq 0$*.*

Condition (i) states that the support of θ does not vary with η, while (ii) is a form of a monotone likelihood property: there is a signal realization η^* which is 'better' news about θ than any other realization, in the sense of shifting weight in favor of low realizations of θ. It additionally requires that the conditional density $f(\theta|\eta^*)$ is strictly decreasing in θ, i.e., higher realizations of θ are less likely than low realizations when $\eta = \eta^*$. (ii) is satisfied for instance when θ has a uniform prior and there are just two possible signal values satisfying the standard monotone likelihood ratio property. Condition (iii) is satisfied if V is exponential ($V = 1 - \exp(-rq), r > 0$). It corresponds to the assumption of 'non-increasing absolute risk aversion' of the Principal's benefit function.

The proof develops necessary conditions for implementation of the second best given the distributional properties (i) and (ii). If the outputs must be second-best, they must be a monotone decreasing function of the (ironed) virtual cost $\hat{h}(\theta \mid \eta)$ in the second-best setting. If they also satisfy the coalitional incentive constraints, they must be monotone in CVC $z(\theta \mid \eta)$. These conditions imply the existence of a monotone transformation from $\hat{h}$ to z, and enable S's ex post rent to be expressed as a function of $\hat{h}$ alone. Condition (iii) is used to show that this rent function is strictly convex which in turn is used to show that the expected rents of S must be strictly higher in state η^* than any other state.

5 Extensions

5.1 Implementation via Modified Delegation

We now show that the optimal allocation can be implemented by a modified form of delegation, where P communicates and transacts only with S on the equilibrium path. In this arrangement, S is 'normally' expected to contract on behalf of the coalition $\{S, A\}$ with P, sending a joint participation decision and report of the state (θ, η) to P after having entered into a side contract with A. However A has the option of circumventing this 'normal' procedure and asking P to activate a grand contract in which A and S will send independent reports and participation decisions to P. The presence of this option ensures that A has sufficient bargaining power within the coalition; it does not have to be 'actually' used, i.e., on the equilibrium path. This mechanism can implement any implementable allocation as a WPBE(w) outcome.

The argument is quite simple, and outlined as follows (we omit a formal proof). Take any WCP allocation $(u_S(\theta, \eta), u_A(\theta, \eta), q(\theta, \eta))$ defined on K which satisfies interim participation constraints, and let aggregate payments to the coalition be $X(\theta, \eta) = u_A(\theta, \eta) + u_S(\theta, \eta) + \theta q(\theta, \eta)$. Let the associated grand contract be denoted as follows. The message spaces are $\tilde{M}_S, \tilde{M}_A$, where $\tilde{M}_S = \Pi \cup \{e_S\}$ and $\tilde{M}_A = K \cup \{e_A\}$. Both S and A report η, and A additionally reports θ. P cross-checks the two η reports, and conditional on these agreeing with one another, transfers are set in the obvious way corresponding to the allocation $(u_S(\theta, \eta), u_A(\theta, \eta), q(\theta, \eta))$, e.g., when neither party exits, both report η and A reports θ, $\tilde{X}_S(\theta, \eta) = u_S(\theta, \eta), \tilde{X}_A(\theta, \eta) = u_A(\theta, \eta) + \theta q(\theta, \eta), \tilde{q}(\theta, \eta) = q(\theta, \eta)$, otherwise these are all zero.

This 'original' grand contract can be augmented as follows. A is offered a message space $M_A = \tilde{M}_A \cup \{\emptyset\}$, while S is offered $M_S = \tilde{M}_S \cup K \cup \{e\}$. The interpretation is that if $m_A = \emptyset$, A decides not to communicate directly with P. And if $m_S \in K \cup \{e\}$, S decides to submit a joint report (θ, η) (or else communicates a joint shutdown decision e) to P on behalf of the coalition. The choice of $m_A = \emptyset, m_S \in K \cup \{e\}$ will correspond to the 'normal' delegation mode.

When the normal delegation mode is in operation, i.e., $m_A = \emptyset, m_S \in K \cup \{e\}$, P will communicate and transact with S alone. Hence transfers and output assignments in the augmented mechanism are defined as follows: (X_S, X_A, q) equals $(\tilde{X}_S, \tilde{X}_A, \tilde{q})$ on $\tilde{M}_S \times \tilde{M}_A$, $(0, X(m_S), q(m_S))$ if $m_A = \emptyset, m_S \in K \cup \{e\}$, and $(-T, -T, 0)$ otherwise where T is a large positive number. The last feature ensures that A and S will always coordinate on either the normal delegation mode, or the grand contract.

It is easy to check that this augmented mechanism has a WPBE(w) where both S and A opt for the normal delegation mode, S offers A a side contract with $m_S(\theta,\eta) = (\theta,\eta) \in K$ and $u^*_A(\theta,\eta) = u_A(\theta,\eta)$ for all (θ,η), which A accepts. To see this note first that if S and A play this augmented grand contract noncooperatively, A will never select $m_A = \emptyset$, since this results in a negative payoff for A no matter what S does. If $m_A = \emptyset, m_S \in K \cup \{e\}$, A is committed to producing a positive quantity while not getting paid anything, while $m_A = \emptyset, m_S \in \tilde{M}_S$ implies $X_A = -T, q = 0$. And given that A does not select $m_A = \emptyset$, neither will S select m_S in $K \cup \{e\}$, owing to the large penalty T for mis-coordination. Hence rejection of a side contract will result in noncooperative play of the original grand contract, with respect to prior beliefs (given the belief restrictions in WPBE(w)), which results in the desired payoffs.

Hence A has an outside option of earning $u_A(\theta,\eta)$ by rejecting any side contract offered by S. This (along with the fact that the allocation is WCP) implies that the side contract offered by S in equilibrium is optimal for S. The reason is that the outcome of any feasible side contract in the normal delegation mode was also attainable as the outcome of some feasible side contract in the original mechanism.

The fact that S could not profitably deviate from the equilibrium side-contract

It is possible to modify this mechanism slightly to ensure that this is the unique WPBE(w) outcome. It can be shown that there exists a (original) grand contract which implements the desired allocation uniquely as a noncooperative Bayesian equilibrium.[20] Transfers in the normal delegation mode of the augmented mechanism can be modifed as follows: (X_A, X_S, q) equals $(-\epsilon, X(m_S) + \delta, q(m_S))$ when $m_A = \emptyset, m_S \in K \cup \{e\}$, where $\delta > \epsilon > 0$. Noncooperative play will then necessarily result in the truthful Bayesian equilibrium of the original contract. And $\delta > \epsilon$ implies that the coalition benefits from the normal delegation mode. Hence S must offer A a side-contract where they agree to play the normal delegation mode. So unique implementation is possible at negligible cost to P, since ϵ and $\delta - \epsilon$ can be chosen to be arbitrarily small.

[20] A proof of this is available on request.

Proposition 7 *Any implementable allocation with weak collusion can be implemented as a WPBE(w) outcome of the modified delegation mechanism described above, where P communicates and transacts with S alone on the equilibrium path. It can be implemented as the unique WPBE(w) outcome at arbitrarily small cost to P.*

5.2 Side Contracts Designed by a Third Party, and Alternative Allocations of Bargaining Power

We now explain how the preceding results extend when the side contract is designed not by S, but instead by a third-party that manages the coalition and assigns arbitrary welfare weights to the payoffs of S and A respectively. Such a formulation has been used by a number of authors to model collusion, such as Laffont and Martimort (1997, 2000), Dequiedt (2006) and Celik and Peters (2011). An advantage of this approach is that it allows an evaluation of the effects of varying the allocation of bargaining power between colluding partners.

Our results extend to such a setting, under the following formulation of side contracts designed by a third party. We assume the third-party's objective is to maximize a weighted sum of S and A's interim payoffs. The third party designs the side contract after learning the realization of η.[21] Both S and A have the option to reject the side contract, in which case they play the grand contract noncooperatively. The WPBE(w) notion is extended as follows: if S rejects the offered side contract, they play the noncooperative equilibrium recommended by P based on prior beliefs. Consequences of rejection by A are the same as assumed previously.[22] This formulation is the natural extension of the assumption of passive beliefs underlying WPBE(w), whereby rejection of a side contract by either S or A will result in reversion to noncooperative play of the grand contract without any further updating of beliefs (except of course when rejection does occur on the equilibrium path with positive probability).

Letting $\alpha \in [0, 1]$ denote the welfare weight assigned by the third-party to

[21]This assumption can be dropped without affecting the results, since it can be shown the third-party can use cross-reporting of η by S and A to learn its true value.

[22]Implicit in this formulation is the assumption that a party that rejects a side contract is not told by the third party whether the other party rejected or accepted the side contract, or what reports the latter sent. Moreover, a party that accepts the side contract is not bound in any way, if the other party rejected it.

A's payoff, the side contract design problem reduces to selecting randomized message $\tilde{m}(\theta, \eta)$ and A's payoff $\tilde{u}_A(\theta, \eta)$ to (using the same notation for the formulation $P(\eta)$ of side contracts in Section 2):

$$\max E[(1-\alpha)[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta, \eta)] + \alpha\tilde{u}_A(\theta, \eta) \mid \eta]$$

subject to $\tilde{m}(\theta \mid \eta) \in \Delta(\hat{M})$,

$$\tilde{u}_A(\theta, \eta) \geq u_A(\theta, \eta)$$

$$\tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta', \eta) + (\theta' - \theta)\hat{q}(\tilde{m}(\theta' \mid \eta))$$

$$E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta, \eta) \mid \eta] \geq E[u_S(\theta, \eta) \mid \eta].$$

Besides modifying the objective function, this formulation adds a participation constraint for S. We refer to this as problem $TP(\eta; \alpha)$. It is straightforward to check that Lemma 2 extends: an allocation is a WPBE(w) outcome if and only if A and S's interim participation constraints are satisfied, and the null side contract is optimal in $TP(\eta; \alpha)$ for every η. Hence the definition of WCP can be extended to WCP(α) by requiring the null side contract to be optimal in $TP(\eta; \alpha)$ for every η.

We now claim that the set of implementable allocations unaffected by the allocation of bargaining power. This also implies that all our preceding results extend to side contracts designed by a third party.

Proposition 8 *The set of WCP(α) allocations is independent of* $\alpha \in [0, 1]$.

The reasoning is straightforward. Consider any $\alpha \in (0, 1)$. It is easy to check that a given allocation is WCP(α) if and only if there is no other allocation attainable by some non-null side contract which satisfies the incentive constraint for A, and which Pareto-dominates it (for A and S) with at least one of them strictly better off. The same characterization applies to any $\alpha' \in (0, 1)$, implying that the set of WCP(α) allocations is independent of $\alpha \in (0, 1)$. The transferability of utility then can be used to show that the set of WCP allocations for interior welfare weights are also the same at the boundary.[23]

[23]If an allocation is WCP(1) but not WCP(α) for some interior α, there must exist a non-null side contract SC^* which allows S to attain a strictly higher payoff, which leaves A's payoff unchanged. Then there exists another feasible non-null side-contract which gives A a slightly higher payoff in all states, which meets S's participation constraint. Hence it is possible to design a feasible side contract that raises A's expected payoff, so the original allocation could not have been WCP(1).

5.3 Altruistic Supervisors

Now consider a different variant, where S offers a side-contract to A, but S is altruistic towards A rather than just concerned with his own income. Suppose S's payoff is $u_S = X_S + t + \alpha[X_A - t - \theta q]$, where $\alpha \in [0, 1]$ is the weight he places on A's payoff. A on the other hand is concerned with only his own income: $u_A = X_A - t - \theta q$.

Our analysis extends as follows. It is easy to check that the expression for coalitional shadow cost is now modified to

$$\pi_\alpha(\theta|\eta) \equiv \theta + (1-\alpha)\frac{F(\theta \mid \eta) - \Lambda(\theta \mid \eta)}{f(\theta \mid \eta)}$$

instead of $\pi(\theta|\eta)$ in Definition 3. In DS, the corresponding expression for the cost of procuring one unit from S is modified from $h(\theta \mid \eta)$ to $h_\alpha(\theta \mid \eta) = \theta + (1-\alpha)\frac{F(\theta|\eta)}{f(\theta|\eta)}$. As long as $\alpha < 1$, this is strictly higher than θ, so DS will still continue to result in a lower profit than NS. The proof that CS dominates NS also goes through *in toto*.

It is interesting to examine the effect of changes in the degree of altruism on P's payoffs. An increase in α lowers S's shadow cost of output in DS $h_\alpha(\theta \mid \eta)$, which benefits P. This is intuitive: the DMR problem becomes less acute with a more altruistic supervisor. Note that with perfect altruism $\alpha = 1$, and the DMR problem disappears: DS then becomes equivalent to NS.

On the other hand, an increase in altruism cannot benefit P in CS. The set of WCP allocations can be shown to be non-increasing in α. Take any WCP allocation corresponding to α: the following argument shows that it is a WCP allocation corresponding to any $\alpha' < \alpha$. Let $z(\theta \mid \eta)$ be the CVC function that is associated with the allocation at α, i.e., it is the ironed version of $\pi_\alpha(\theta|\eta)$ corresponding to some distribution function $\Lambda_\alpha(\cdot|\eta)$ on $[\underline{\theta}(\eta), \bar{\theta}(\eta)]$. We can then select

$$\Lambda_{\alpha'}(\theta \mid \eta) = \frac{\alpha - \alpha'}{1-\alpha'}F(\theta \mid \eta) + \frac{1-\alpha}{1-\alpha'}\Lambda_\alpha(\theta \mid \eta)$$

when the altruism parameter is α', which is a distribution function since $\alpha > \alpha'$. This ensures that the same CSC and CVC function is available when the altruism parameter is α', since by construction $\pi_\alpha(\theta|\eta) = \pi_{\alpha'}(\theta|\eta)$. Hence the allocation satisfies the sufficient condition for WCP when the altruism parameter is α'.

Finally, if $\alpha = 1$, the CSC function π_α coincides with the identity function θ, the cost of the agent in NS. We thus obtain

Proposition 9 *In CS, P's optimal payoff is non-increasing in α. In DS, P's optimal payoff is increasing in α. When $\alpha = 1$, P's optimal payoffs in DS, NS and CS coincide.*

5.4 Linear Benefits

So far we have assumed that V is strictly concave, satisfying Inada conditions so as to guarantee interior allocations. We now briefly describe how preceding results extend when V is linear upto some capacity limit, and optimal allocations are typically non-interior. This case is relevant for an indivisible project where q refers to the probability of the project being carried out, or to financial contexts where q denotes a revenue stream accruing to the Principal.

We now consider the implications of assuming that $V(q) = Vq$ with $V \in (\underline{\theta}, \bar{\theta})$ and $q \in [0, 1]$. For simplicity we focus on the case of a binary signal $\eta \in \{\eta_1, \eta_2\}$ with the monotone likelihood ratio property ($a(\eta_2 \mid \theta)/a(\eta_1 \mid \theta)$ is increasing in θ) and full support of θ for each signal: $\Theta(\eta_1) = \Theta(\eta_2) = \Theta$. The monotone likelihood ratio property implies that the distribution of θ conditional on η_2 first order stochastically dominates that on η_1: $F(\theta \mid \eta_1) > F(\theta \mid \eta_2)$ for $\theta \in (\underline{\theta}, \bar{\theta})$. It also implies the following ranking among the virtual costs:

$$h(\theta \mid \eta_2) < H(\theta) < h(\theta \mid \eta_1)$$

for any $\theta \in (\underline{\theta}, \bar{\theta}]$. The same ranking is preserved after applying the ironing transformation as well:

$$\hat{h}(\theta \mid \eta_2) < H(\theta) < \hat{h}(\theta \mid \eta_1)$$

for any $\theta \in (\underline{\theta}, \bar{\theta}]$.

With linear V, the second best output schedules $(q^{SB}(\theta, \eta_1), q^{SB}(\theta, \eta_2))$ are characterized by the thresholds $(\theta_1^{SB}, \theta_2^{SB})$ such that the output is 1 for smaller θ than the threshold and 0 for larger θ than it where $\theta_i^{SB} \equiv \sup\{\theta \mid V \geq \hat{h}(\theta \mid \eta_i)\}$.[24] The ranking among virtual costs implies $\theta_1^{SB} < \theta^{NS} < \theta_2^{SB}$.

[24]More precisely $q^{SB}(\theta, \eta_i)$ can be the arbitrary non-increasing function on $\{\theta \mid \hat{h}(\theta \mid \eta_i) = V\}$ as far as it takes constant value on connected open interval of θ satisfying

The Principal's expected payoff in NS is $(V - \theta^{NS})F(\theta^{NS})$ and in SB, it is

$$p(\eta_1)(V - \theta_1^{SB})F(\theta_1^{SB} \mid \eta_1) + p(\eta_2)(V - \theta_2^{SB})F(\theta_2^{SB} \mid \eta_2).$$

The result that delegated contracting DS is inferior to NS continues to go through without any modification, as DS is associated with an increase in the unit cost of delivering output by S, compared to the cost of output delivered by A in NS. The argument used earlier to show that CS dominates NS however cannot be applied, since output allocations in NS are not interior. It turns out that a different variation can be constructed, involving adjustment in the threshold θ^{NS} in NS to thresholds θ_i in state $\eta_i, i = 1, 2$ where $\theta_1 < \theta^{NS} < \theta_2$, such that P is better off. The proof rests on showing that such thresholds can be selected in the neighborhood of θ^{NS} such that P's profit rises in both states η_1, η_2 owing to outputs moving closer to the corresponding second-best outputs $\theta_1^{SB}, \theta_2^{SB}$. These thresholds nevertheless have to be selected carefully to ensure that the resulting allocation is WCP.

Proposition 10 *Suppose $V(q) = Vq$ with $V \in (\underline{\theta}, \bar{\theta})$, with $q \in [0, 1]$, and $\Pi \equiv \{\eta_1, \eta_2\}$ such that $a(\eta_2 \mid \theta)/a(\eta_1 \mid \theta)$ is strictly increasing and $\Theta(\eta_1) = \Theta(\eta_2) = \Theta$. Then $\Pi_{CS} > \Pi_{NS}$: the Principal benefits from hiring the supervisor.*

With regard to the attainability of second-best payoffs, examples can be constructed where this is and is not possible.[25]

6 Concluding Comments

We have a considered a model of weak ex ante collusion between a supervisor and agent, where collusion arises with regard to both participation and reporting decisions, and outside option payoffs in coalitional bargaining are determined by noncooperative equilibria of a grand contract designed by the Principal. We showed in such settings that the Principal can still benefit from employing the supervisor. This requires the Principal to design a grand contract involving both the supervisor and the agent, rather

$\{\theta \mid \int_0^{F(\theta|\eta_i)} h(F^{-1}(\phi \mid \eta_i) \mid \eta_i) d\phi > \int_0^{F(\theta|\eta_i)} \hat{h}(F^{-1}(\phi \mid \eta_i) \mid \eta_i) d\phi\}$ (which is the pooling region where $h(\theta \mid \eta)$ is flattered). However in our analysis, without loss of generality, our attention can be restricted to output schedule with one threshold.

[25] Details are available on request.

than delegating authority over contracting with the agent to the supervisor in an unconditional manner. It is essential for the Principal to give both parties suitable outside option payoffs by designing such a grand contract judiciously. The presence of such a centralized safeguard as an option then allows optimal outcomes to be implemented by delegating authority to the supervisor. These results are consistent with the widespread prevalence of delegation to information intermediaries, managers and regulators, and highlight the importance of centralized oversight mechanisms which are necessary supplements to mitigate 'abuse of power' by the concerned intermediaries. While the commonsense justification for such mechanism is typically based on considerations of fair treatment of agents, our analysis shows how such mechanisms are essential to prevent inefficient output contractions and loss of profits of the Principal owing to monopsonistic behavior by intermediaries to whom authority is delegated.

In many contexts of hierarchical supervision or management, examples of such oversight mechanisms are commonly observed. Taxpayers are usually able to appeal assessments made by auditors or assessors appointed by the government, in which taxpayers and assessors argue their respective cases with an appellate tribunal or judge. Firms establish internal adjudication procedures where employees can appeal decisions of their hierarchical superiors. The presence of such 'rights' is a privilege of firm's employees not shared by external suppliers: this may be an important distinguishing feature of intra-firm relationships.

Additional questions arise regarding how supervisors and agents ought to be matched. Should agents have the right to select their supervisors, or should they be assigned by the Principal? Our analysis of consequences of altruism of the auditor towards the agent confirms the common-sense notion that the Principal should ensure absence of any overt conflict-of-interest or likelihood of favoritism of the auditor towards the agent. This implies the need for the Principal to check for possible external, social or personal relationships between supervisors and agents.

Subject to such constraints, can agents be allowed to select their supervisor, as commonly observed in the context of firms that are allowed to select their financial auditor or quality-rating agency? This raises the bargaining power of firms vis-a-vis their supervisor, compared to a system where supervisors are externally appointed. Our model showed that optimal mechanisms are unaffected by varying the allocation of bargaining power between the supervisor and agent, once their outside option payoffs have been set by the ‘fallback’ noncooperative outcomes of the grand contract designed by the Principal.

This result, however, is based critically on the assumption of weak collusion, where outside options are unaffected by the allocation of bargaining power. With the alternative notion of strong collusion, this is no longer true — when one party can make a take-it-or-leave-it offer of the side contract to the other party, and the former can commit to threats of how that party will behave subsequently in the grand contract should the side contract offer be rejected. Such a solution concept may be reasonable in settings where one of the parties is a long-lived player and develops a reputation for ‘toughness’, while the other party is a short-term player. How our results will be modified in the case of strong collusion remains an interesting task for future research.

References

Celik, G. (2009), "Mechanism Design with Collusive Supervision", *Journal of Economic Theory*, 144(1), 69-95.

Celik, G. and M. Peters (2011), "Equilibrium Rejection of a Mechanism", *Games and Economic Behavior*, 73(2), 375-387.

Che, Y. K. and J. Kim (2006), "Robustly Collusion-Proof Implementation", *Econometrica*, 74(4), 1063-1107.

Che, Y. K. and J. Kim (2009), "Optimal Collusion-Proof Auctions", *Journal of Economic Theory*, 144(2), 565-603.

Dequiedt, V. (2006), "Ratification and Veto Constraints in Mechanism Design", *Working Paper*, CERDI.

Dequiedt, V. (2007), "Efficient Collusion in Optimal Auctions", *Journal of Economic Theory*, 136(1), 302-323.

Faure-Grimaud, A., J. J. Laffont and D. Martimort (2003), "Collusion, Delegation and Supervision with Soft Information", *The Review of Economic Studies*, 70(2), 253-279.

Laffont, J. J. and D. Martimort (1997), "Collusion under Asymmetric Information", *Econometrica*, 65(4), 875-911.

Laffont, J. J. and D. Martimort (2000), "Mechanism Design with Collusion and Correlation", *Econometrica*, 68(2), 309-342.

Laffont, J. J. and J. Tirole (1993), "A Theory of Incentives in Procurement and Regulation", MIT Press, Cambridge, MA.

Mookherjee, D. and M. Tsumagari (2004), "The Organization of Supplier Networks: Effects of Delegation and Intermediation", *Econometrica*, 72(4), 1179-1219.

Motta, A. (2009), "Collusion and Selective Supervision", working paper.

Pavlov, G. (2008), "Auction Design in the presence of Collusion", *Theoretical Economics*, 3(3), 383-429.

Quesada, L. (2004), "Collusion as an Informed Principal Problem", University of Wisconsin.

Tirole, J. (1986), "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations", *Journal of Law, Economics, and Organization*, 2, 181-214.

Appendix: Proofs

Proof of Proposition 1: Consider the necessity part. Suppose the allocation (u_A, u_S, q) is WCP. Then the null side contract is optimal for S for every η, so must be feasible in $P(\eta)$. This implies $(u_A(\theta,\eta), q(\theta,\eta))$ satisfies A's incentive compatibility condition. Now consider the problem $P(\eta)$. The incentive constraint

$$\tilde{u}_A(\theta,\eta) \geq \tilde{u}_A(\theta^{'},\eta) + (\theta^{'} - \theta)\hat{q}(\tilde{m}(\theta^{'} \mid \eta))$$

is equivalent to

$$\tilde{u}_A(\theta,\eta) = \tilde{u}_A(\bar{\theta}(\eta),\eta) + \int_{\theta}^{\bar{\theta}(\eta)} \hat{q}(\tilde{m}(y \mid \eta))dy$$

and $\hat{q}(\tilde{m}(\theta \mid \eta))$ is non-increasing in θ. Then the problem can be rewritten as

$$\max E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta,\eta) \mid \eta]$$

subject to $\tilde{m}(\theta \mid \eta) \in \Delta(\hat{M})$ where $\hat{M} \equiv K \cup \{e\}$,

$$\tilde{u}_A(\theta,\eta) = \tilde{u}_A(\bar{\theta}(\eta),\eta) + \int_{\theta}^{\bar{\theta}(\eta)} \hat{q}(\tilde{m}(y \mid \eta))dy \geq u_A(\theta,\eta)$$

and $\hat{q}(\tilde{m}(\theta \mid \eta))$ non-increasing in θ. Since randomized side contracts can be chosen, the objective function is concave and the feasible set is convex. So the solution maximizes (subject to the constraint $\hat{q}(\tilde{m}(\theta \mid \eta))$ is non-increasing in θ) the following Lagrangian expression corresponding to some non-decreasing function $\tilde{\Lambda}(\theta \mid \eta)$:

$$\begin{aligned} \mathcal{L} \equiv\ & E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta,\eta)|\eta] \\ +\ & \int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\tilde{u}_A(\theta,\eta) - u_A(\theta,\eta)]d\tilde{\Lambda}(\theta \mid \eta) \end{aligned}$$

where $\hat{X}(\tilde{m}), \hat{q}(\tilde{m})$ denote expected values of $\hat{X}(m), \hat{q}(m)$ taken with respect to probability measure $\tilde{m}$ over $m \in \hat{M}$. Note that without loss of generality, $\tilde{u}_A(\theta,\eta)$ is a deterministic function.

A's incentive constraint implies $\tilde{u}_A(\theta,\eta)$ is continuous on $\Theta(\eta)$. Hence integration by parts yields:

$$\begin{aligned}
&\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} \tilde{u}_A(\theta,\eta)d\tilde{\Lambda}(\theta \mid \eta) = \tilde{\Lambda}(\bar{\theta}(\eta) \mid \eta)\tilde{u}_A(\bar{\theta}(\eta),\eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta)\tilde{u}_A(\underline{\theta}(\eta),\eta) \\
+ &\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} \tilde{\Lambda}(\theta \mid \eta)\hat{q}(\tilde{m}(\theta \mid \eta))d\theta \\
= &[\tilde{\Lambda}(\bar{\theta}(\eta) \mid \eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta)]\tilde{u}_A(\bar{\theta}(\eta),\eta) \\
+ &\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\tilde{\Lambda}(\theta \mid \eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta)]\hat{q}(\tilde{m}(\theta \mid \eta))d\theta.
\end{aligned}$$

The second equality comes from

$$\tilde{u}_A(\underline{\theta}(\eta),\eta) = \tilde{u}_A(\bar{\theta}(\eta),\eta) + \int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} \hat{q}(\tilde{m}(y \mid \eta))dy.$$

Next consider the effect of raising uniformly A's outside option function from $u_A(\theta,\eta)$ to $u_A(\theta,\eta)+\Delta$ where Δ is an arbitrary positive scalar. It is evident that the solution is unchanged, except that $\tilde{u}_A(\theta,\eta)$ is raised uniformly by Δ. Hence the maximized payoff of S must fall by Δ, implying that

$$\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} \Delta d\tilde{\Lambda}(\theta \mid \eta) = [\tilde{\Lambda}(\bar{\theta}(\eta) \mid \eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta)]\Delta = \Delta,$$

and so $\tilde{\Lambda}(\bar{\theta}(\eta) \mid \eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta) = 1$ in the optimal solution. Now define $\Lambda(\theta \mid \eta) \equiv \tilde{\Lambda}(\theta \mid \eta) - \tilde{\Lambda}(\underline{\theta}(\eta) \mid \eta)$. Then $\Lambda(\theta \mid \eta)$ is non-decreasing in θ with $\Lambda(\underline{\theta}(\eta) \mid \eta) = 0$ and $\Lambda(\bar{\theta}(\eta) \mid \eta) = 1$.

This implies

$$\begin{aligned}
\mathcal{L} \equiv &\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\hat{X}(\tilde{m}(\theta \mid \eta)) - \pi(\theta \mid \eta)\hat{q}(\tilde{m}(\theta \mid \eta))]dF(\theta \mid \eta) \\
- &\int_{(\underline{\theta}(\eta),\bar{\theta}(\eta)]} u_A(\theta,\eta)d\Lambda(\theta \mid \eta) \qquad (6)
\end{aligned}$$

where $\pi(\theta \mid \eta) \equiv \theta + \frac{F(\theta|\eta)-\Lambda(\theta|\eta)}{f(\theta|\eta)}$. This has to be maximized subject to the constraint that $\hat{q}(\tilde{m}(\theta \mid \eta))$ is non-increasing in θ. This reduces to the unconstrained maximization of the corresponding expression where the CSC function $\pi(\cdot \mid \eta)$ is replaced by the corresponding CVC function $z(\cdot \mid \eta)$ using the ironing procedure relative to the cdf $F(\theta \mid \eta)$.

If $\tilde{m}^*(\theta \mid \eta)$ is optimal in problem $P(\eta)$, there exists $\pi(\cdot \mid \eta) \in Y(\eta)$ so that the optimal side contract $\tilde{m} = \tilde{m}^*(\theta \mid \eta)$ maximizes

$$\hat{X}(\tilde{m}(\theta \mid \eta)) - z(\theta \mid \eta)\hat{q}(\tilde{m}(\theta \mid \eta))$$

where $z(\theta \mid \eta) \equiv z(\theta \mid \pi(\cdot \mid \eta), \eta)$. Moreover $\hat{q}(\tilde{m}^*(\theta \mid \eta))$ must be non-increasing in θ and flat on any interval of θ which is a subset of $\Theta(\pi(\cdot \mid \eta), \eta)$.

If the optimal side contract is degenerate and concentrated at (θ, η), it must be the case that

$$\hat{X}(\theta, \eta) - z(\theta \mid \eta)\hat{q}(\theta, \eta) \geq \hat{X}(\tilde{m}^{'}) - z(\theta \mid \eta)\hat{q}(\tilde{m}^{'})$$

for any $\tilde{m}^{'} \in \Delta(\hat{M})$. This implies

$$\hat{X}(\theta, \eta) - z(\theta \mid \eta)q(\theta, \eta) \geq \hat{X}(\theta^{'}, \eta^{'}) - z(\theta \mid \eta)q(\theta^{'}, \eta^{'})$$

$$\hat{X}(\theta, \eta) - z(\theta \mid \eta)q(\theta, \eta) \geq 0$$

for any $(\theta, \eta), (\theta^{'}, \eta^{'})$, implying (i) in the proposition. Obviously $q(\theta, \eta)$ must be non-increasing in θ and must be flat on any interval of θ which is a subset of $\Theta(\pi(\cdot \mid \eta), \eta)$ (implying (iii) in the proposition).

Now consider the sufficiency part. Consider any state η. Suppose there is a CSC function $\pi(\cdot \mid \eta) \in Y(\eta)$ which is ironed to yield the CVC function $z(\cdot|\eta)$ such that $(u_S(\theta, \eta), u_A(\theta, \eta), q(\theta, \eta))$ satisfies all the conditions in the proposition. Define $(\hat{X}(m), \hat{q}(m))$ on $\hat{M} \equiv K \cup \{e\}$ such that

$$(\hat{X}(\theta, \eta), \hat{q}(\theta, \eta)) = (u_S(\theta, \eta) + u_A(\theta, \eta) + \theta q(\theta, \eta), q(\theta, \eta))$$

and

$$(\hat{X}(e), \hat{q}(e)) = (0, 0).$$

and extend this to $(\hat{X}(\tilde{m}), \hat{q}(\tilde{m}))$ on $\Delta(\hat{M})$ in the obvious manner. Consider the problem $P(\eta)$ as selection of $\tilde{m}(\theta|\eta), \tilde{u}_A(\theta, \eta)$ to maximize

$$E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta, \eta) \mid \eta]$$

subject to

$$\tilde{u}_A(\theta, \eta) \geq u_A(\theta, \eta)$$

for any $\theta \in \Theta(\eta)$,

$$\tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta^{'}, \eta) + (\theta^{'} - \theta)\hat{q}(\tilde{m}(\theta^{'} \mid \eta))$$

for any $\theta, \theta' \in \Theta(\eta)$. For $\tilde{u}_A(\theta,\eta)$ which satisfies constraints of the problem, we have

$$\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\tilde{u}_A(\theta,\eta) - u_A(\theta,\eta)] d\Lambda(\theta \mid \eta) \geq 0.$$

Then

$$\begin{aligned} & E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta,\eta) \mid \eta] \\ \leq \ & E[\hat{X}(\tilde{m}(\theta \mid \eta)) - \theta\hat{q}(\tilde{m}(\theta \mid \eta)) - \tilde{u}_A(\theta,\eta) \mid \eta] \\ + \ & \int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\tilde{u}_A(\theta,\eta) - u_A(\theta,\eta)] d\Lambda(\theta \mid \eta). \end{aligned}$$

Now consider the problem of maximizing the right hand side of this inequality, subject to the constraint that $\hat{q}(\tilde{m}(\theta \mid \eta))$ is non-increasing in θ. Using the same steps in the proof of the necessity part, this can be expressed as a problem of selecting $\tilde{m}(\theta|\eta)$ to maximize the Lagrangean (6) subject to the constraint that $\hat{q}(\tilde{m}(\theta \mid \eta))$ is non-increasing in θ. Conditions (i)-(iii) imply that the right-hand-side is maximized at $\tilde{m}(\theta \mid \eta) = (\theta,\eta)$ and $\tilde{u}_A(\theta,\eta) = u_A(\theta,\eta)$. Since

$$\int_{[\underline{\theta}(\eta),\bar{\theta}(\eta)]} [\tilde{u}_A(\theta,\eta) - u_A(\theta,\eta)] d\Lambda(\theta \mid \eta) = 0$$

when $\tilde{u}_A(\theta,\eta) = u_A(\theta,\eta)$, this shows that the left hand side of the above inequality is also maximized at $\tilde{m}(\theta \mid \eta) = (\theta,\eta)$ and $\tilde{u}_A(\theta,\eta) = u_A(\theta,\eta)$. Hence $(\tilde{m}(\theta \mid \eta), \tilde{u}_A(\theta,\eta)) = ((\theta,\eta), u_A(\theta,\eta))$ solves $P(\eta)$. ■

Proof of Proposition 2: Let P offer the following grand contract which is a revelation mechanism satisfying

$$(X_A(m_A,m_S), X_S(m_A,m_S), q(m_S,m_A); M_S, M_A)$$

where $M_S = \Pi \cup \{e_S\}$ and $M_A = \Theta \cup \{e_A\}$.

(i) $X_S(m_A,m_S) = 0$ for any (m_A,m_S).

(ii) $q(\theta,\eta) = q^{SB}(\hat{h}(\theta \mid \eta))$ and $X_A(\theta,\eta) = \theta q^{SB}(\hat{h}(\theta \mid \eta)) + \int_\theta^{\bar{\theta}(\eta)} q^{SB}(\hat{h}(y \mid \eta))dy$, if $(\theta,\eta) \in K$, otherwise both are set equal to zero.

(iii) $X_A(e_A,m_S) = q(e_A,m_S) = 0$ for any m_S.

(iv) $(X_A(\theta, e_S), q(\theta, e_S)) = (\hat{X}_A(\theta), \hat{q}(\theta))$ satisfies the following properties: (a) $\hat{X}_A(\theta) - \theta\hat{q}(\theta) \geq \hat{X}_A(\theta') - \theta\hat{q}(\theta')$, (b) $\hat{X}_A(\theta) - \theta\hat{q}(\theta) \geq 0$ and (c) there exists θ' such that $\hat{q}(\theta') = q(\theta, \eta)$ and $\hat{X}_A(\theta') > X_A(\theta, \eta)$ for any $(\theta, \eta) \in \Theta \times \Pi$.[26]

For any η, P recommends S and A play the Bayesian equilibrium in which they both participate and report truthfully. Following any SC offered by S, the consequences of rejection of this SC by A are specified by P as follows. Beliefs held are $p_\emptyset(\eta)$. Continuation strategies are: S exits, A plays the revelation mechanism according to the 'gilded' contract specified in (iv), participates always and reports truthfully.

If A rejects any side-contract offer, S will subsequently earn nothing owing to property (i). S cannot benefit from deciding to participate in the grand contract following rejection of any SC. So it is sequentially rational for S to exit following rejection of an offered SC.

We claim that it is optimal for S to not offer any non-null side contract. By construction of the 'gilded' contract, it dominates any offered non-null side-contract from A's point of view in state (θ, η), unless the side-payment $t(\theta, \eta)$ specified in the latter is positive. Hence S will lose money from a non-null side-contract in any state in which it is accepted, and cannot benefit by offering it.

Consequently, there is a Weak Perfect Bayesian equilibrium of this game in which S never offers any side contract. This implies that S and A play the recommended Bayesian equilibrium, and the second-best allocation is implemented. ■

Proof of Lemma 1: Suppose on the equilibrium path S offers a side contract SC^* in some state $\eta \in \Pi$ which is rejected by a set $T_r \subseteq \Theta(\eta)$ of types of A with positive measure conditional on η. Let the continuation beliefs following rejection of SC^* be denoted p^*, and the Bayesian equilibrium of the grand contract thereafter is denoted $c^* \in C(p^*)$ (here we are suppressing η in the notation for expositional convenience).

Now suppose S offers an alternative side contract $\tilde{SC}$, which agrees with SC^* if A reports $\theta \in \Theta(\eta) \backslash T_r$ to S, i.e., results in the same coordinated report to P and the same side-payment as stipulated by SC^*. If instead A

[26] For instance, we can choose $(\hat{X}_A(\theta), \hat{q}(\theta))$ such that (i) $\hat{q}(\theta)$ is continuous and strictly decreasing in θ with $\hat{q}(\underline{\theta}) = \max_{(\theta,\eta)\in\Theta\times\Pi} q(\theta, \eta)$ and $\hat{q}(\bar{\theta}) = \min_{(\theta,\eta)\in\Theta\times\Pi} q(\theta, \eta)$, and (ii) $\hat{X}_A(\theta) = \theta\hat{q}(\theta) + \int_\theta^{\bar{\theta}} \hat{q}(y)dy + R$ for sufficiently large $R > 0$

reports $\theta \in T_r$, S proposes the same joint report (θ, η) they would have made independently in c^*, with no side-payment. If $\tilde{SC}$ is rejected by A, they play according to (p^*, c^*) in the grand contract. This ensures consistency with criteria (a) and (b) in the definition of WPBE(w).

If all types of A accept $\tilde{SC}$ and report truthfully, it results in the same allocation as in SC^*. Rejecting it results in the same continuation play of the grand contract that resulted from rejecting SC^*. Conditional on accepting $\tilde{SC}$, no type θ of A can benefit from deviating from truthful-reporting. Otherwise, if $\theta \in \Theta(\eta) \backslash T_r$ benefitted from deviating, this would imply they would have had a profitable deviation from their equilibrium response to SC^*. If $\theta \in T_r$ benefits by deviating, this type would have benefitted earlier also, either by accepting SC^*, or rejecting it and then deviating to the strategy played by some other type of A while playing the Bayesian equilibrium of the grand contract.

Owing to restriction (a) of Definition 6, rejection of any other side-contract offer SC' will also result in the same continuation outcomes in the grand contract. Hence the consequences of S deviating to some other side contract offer remain unchanged. The consequences of not offering a side contract have not changed. So it is optimal for S to offer $\tilde{SC}$. ∎

Proof of Lemma 2: sketched in the text.

Proof of Proposition 3:

Necessity follows straightforwardly from Lemmas 1 and 2. To show sufficiency, consider a WCP allocation satisfying participation constraints. Let P offer the following revelation mechanism in the grand contract: $X_S = X_A = q = 0$ if $m_A = e_A$ or $m_S = e_S$. If $m_A \neq e_A$ and $m_S \neq e_S$, and A reports (θ, η_A) while S reports η_S, $q((\theta, \eta_A), \eta_S) = q(\theta, \eta_S), X_S((\theta, \eta_A), \eta_S) = u_S(\theta, \eta_A), X_A((\theta, \eta_A), \eta_S) = \theta q(\theta, \eta_S) + u_A(\theta, \eta_S) - T(\eta_S, \eta_A)$ where T equals zero if $\eta_A = \eta_S$ and $(\theta, \eta_A) \in K$, and a large negative number otherwise. We first show property (i) of Lemma 2 holds. Consider any η. Conditional on both S and A participating, it is optimal for S to report $\eta_S = \eta$ since S's payoff does not depend on η_S. Given that S is reporting truthfully, it is optimal for A to report $\eta_A = \eta$. WCP implies that the null side contract is feasible in the side contracting problem for every η, hence it is optimal for A to report θ truthfully, given that η is being reported truthfully. Given that both S and A report truthfully conditional on participation, the interim participation constraints imply it is optimal for them to always participate.

Let this equilibrium be denoted c^*. We claim that there is a WPBE(w) in which S always offers a null side contract, whose outcome is c^*. The WPBE(w) restriction implies c^* must be the consequence of rejection by A of any offered non-null side contract. Hence $u_A(\theta, \eta)$ is the outside option of A which S takes as given while selecting a side contract. Since the allocation resulting from c^* is WCP, S cannot benefit from offering any non-null side contract. ■

Proof of Proposition 4:

At the first step, note that the optimal side contract problem for S in DS involves an outside option for A which is identically zero. This reduces to a standard problem of contracting with a single agent with adverse selection and an outside option of zero, where the principal has a prior distribution $F(\theta|\eta)$ over the agent's cost θ in state η. The CSC function equals $h(\theta|\eta)$, and the CVC function $z(\theta|\eta)$ reduces to $\hat{h}(\theta|\eta)$ obtained by applying the ironing rule to $h(\theta|\eta)$ and distribution $F(\theta|\eta)$.

Given this, P's contract with S in DS is effectively a contracting problem for P with a single supplier whose unit supply cost is $\hat{h}(\theta|\eta)$. P's prior over this supplier's cost is given by distribution function

$$G(h) \equiv \Pr((\theta, \eta) \mid \hat{h}(\theta \mid \eta) \leq h)$$

for $h \geq \underline{\theta}$ and $G(h) = 0$ for $h < \underline{\theta}$. Let $G(h \mid \eta)$ denote the cumulative distribution function of $h = \hat{h}(\theta \mid \eta)$ conditional on η:

$$G(h \mid \eta) \equiv \Pr(\theta \mid \hat{h}(\theta \mid \eta) \leq h, \eta)$$

for $h \geq \hat{h}(\underline{\theta}(\eta) \mid \eta)(= \underline{\theta}(\eta))$ and $G(h \mid \eta) = 0$ for $h < \underline{\theta}(\eta)$. Then $G(h) = \Sigma_{\eta \in \Pi} p(\eta) G(h \mid \eta)$. Since $\hat{h}(\theta \mid \eta)$ is continuous on $\Theta(\eta)$, $G(h \mid \eta)$ is strictly increasing in h on $[\underline{\theta}(\eta), \hat{h}(\bar{\theta}(\eta) \mid \eta)]$. However, $G(h \mid \eta)$ may fail to be left-continuous.

Hence P's problem in DS reduces to

$$\max E_h[V(q(h)) - X(h)]$$

subject to

$$X(h) - hq(h) \geq X(h^{'}) - hq(h^{'})$$

for any $h, h^{'} \in [\underline{\theta}, \bar{h}]$ and

$$X(h) - hq(h) \geq 0$$

for any $h \in [\underline{\theta}, \bar{h}]$ where the distribution function of h is $G(h)$ and $\bar{h} \equiv \max_{\eta \in \Pi} \hat{h}(\bar{\theta}(\eta) \mid \eta)$. The corresponding problem in NS is

$$\max E_\theta[V(q(\theta)) - X(\theta)]$$

subject to

$$X(\theta) - \theta q(\theta) \geq X(\theta^{'}) - \theta q(\theta^{'})$$

for any $\theta, \theta^{'} \in \Theta$ and

$$X(\theta) - \theta q(\theta) \geq 0$$

for any $\theta \in \Theta$. The two problems differ only in the underlying cost distributions of P: $G(h)$ in the case of DS and $F(\theta)$ in the case of NS. Since $\theta < \hat{h}(\theta \mid \eta)$ for $\theta > \underline{\theta}(\eta)$,

$$G(h \mid \eta) \equiv \Pr(\theta \mid \hat{h}(\theta \mid \eta) \leq h, \eta) < \Pr(\theta \mid \theta \leq h, \eta) = F(h \mid \eta)$$

for $h \in (\underline{\theta}(\eta), \hat{h}(\bar{\theta}(\eta) \mid \eta))$, implying

$$G(h) = \Sigma_{\eta \in \Pi} p(\eta) G(h \mid \eta) < \Sigma_{\eta \in \Pi} p(\eta) F(h \mid \eta) = F(h)$$

for any $h \in (\underline{\theta}, \bar{h})$. Therefore the distribution of h in DS (strictly) dominates that of θ in NS in the first order stochastic sense.

It remains to show that this implies that P must earn a lower profit in DS. We prove the following general statement. Consider two contracting problems with a single supplier which differ only in regard to the cost distributions G_1 and G_2, where $G_1(h) < G_2(h)$ for any $h \in (\underline{h}, \bar{h})$. Let the maximized profit of P with distribution G be denoted $W(G)$. We will show $W(G_1) < W(G_2)$.

Let $q_1(h)$ denote the optimal solution of the problem based on $G_1(h)$.

(i) First we show that $V^{'}(q_1(h)) < h$ does not hold for any h. Suppose otherwise that there exists some interval over which $V^{'}(q_1(h)) < h$. Then we can replace the portion of $q_1(h)$ with $V^{'}(q_1(h)) < h$ by $q^*(h)$ with $V^{'}(q^*(h)) = h$, without violating the constraint that $q(h)$ is non-increasing. It raises the value of the objective function, since $V(q_1(h)) - hq_1(h) < V(q_1^*(h)) - hq_1^*(h)$ for h where $q_1(h)$ is replaced by $q^*(h)$, and $\int_h^{\bar{h}} q(y)dy$ decreases with this replacement. This is a contradiction.

(ii) Next we show that for any $h^{'} \in [\underline{h}, \bar{h})$, there exists a subinterval of $[h^{'}, \bar{h})$ over which $V^{'}(q_1(h)) > h$. Otherwise, there exists $h^{'} \in [\underline{h}, \bar{h})$ such that $q_1(h) = q^*(h)$ almost everywhere on $[h^{'}, \bar{h})$. Then for any $h \in [h^{'}, \bar{h})$,

$$V(q^*(h)) - hq^*(h) - \int_h^{\bar{h}} q^*(y)dy = V(q^*(\bar{h})) - \bar{h}q^*(\bar{h}),$$

since $V(q^*(h)) - hq^*(h) = \int_h^{\bar{h}} q^*(y)dy + V(q^*(\bar{h})) - \bar{h}q^*(\bar{h})$ (which follows from the Envelope Theorem: $d[V(q^*(h)) - hq^*(h)]/dh = -q^*(h)$). Then

$$\begin{aligned} & W(G_1) = (1 - G_1(h^{'}))[V(q^*(\bar{h})) - \bar{h}q^*(\bar{h})] \\ + \ & G_1(h^{'})E[V(q_1(h)) - hq_1(h) - \int_h^{h^{'}} q_1(y)dy \mid h \leq h^{'}] - G_1(h^{'}) \int_{h^{'}}^{\bar{h}} q^*(y)dy. \end{aligned}$$

Now consider output schedule $q(h)$ such that $q(h) = q_1(h)$ for $h \leq h^{'}$ and $q(h) = q^*(\bar{h})$ for $h > h^{'}$. It is evident that $q(h)$ is non-increasing in h and generates a higher value of the objective function, since $\int_{h^{'}}^{\bar{h}} q^*(y)dy > \int_{h^{'}}^{\bar{h}} q^*(\bar{h})dy$. This is a contradiction.

(iii) We show there does not exist q such that $q_1(h) = q$ almost everywhere. Otherwise, $q_1(h) = q$ almost everywhere for some q. Then

$$V(q) - hq - \int_h^{\bar{h}} qdy = V(q) - \bar{h}q,$$

which is not larger than $V(q^*(\bar{h})) - \bar{h}q^*(\bar{h})$ which equals $\max_{\tilde{q}}[V(\tilde{q}) - \bar{h}\tilde{q}]$. We can show that the value of the objective function is increased by choosing the following output schedule $\tilde{q}(h)$:

$$\tilde{q}(h) = \begin{cases} q^*(\bar{h}) & h \in [h^*, \bar{h}] \\ q^*(\bar{h}) + \epsilon & h \in [\underline{h}, h^*] \end{cases}$$

where h^* is any element of $(\underline{h}, \bar{h})$, and $\epsilon > 0$ is chosen so that $V(q^*(\bar{h}) + \epsilon) - V(q^*(\bar{h})) > \epsilon h^*$. This is possible since $\lim_{\epsilon \to 0} \frac{V(q^*(\bar{h})+\epsilon) - V(q^*(\bar{h}))}{\epsilon} = V^{'}(q^*(\bar{h})) = \bar{h}$, implying existence of $\epsilon > 0$ such that $V(q^*(\bar{h}) + \epsilon) - V(q^*(\bar{h})) > \epsilon h^*$ for any $h^* < \bar{h}$.

Then we obtain a contradiction, since

$$\begin{aligned} & V(q^*(\bar{h})) - \bar{h}q^*(\bar{h}) \\ < \ & (1 - G_1(h^*))[V(q^*(\bar{h})) - \bar{h}q^*(\bar{h})] + G_1(h^*)[V(q^*(\bar{h}) + \epsilon) - \bar{h}q^*(\bar{h}) - \epsilon h^*] \\ = \ & \int_{\underline{h}}^{\bar{h}} [V(\tilde{q}(h)) - h\tilde{q}(h) - \int_h^{\bar{h}} \tilde{q}(y)dy]dG_1(h). \end{aligned}$$

(iv) Define

$$\Phi(h) \equiv V(q_1(h)) - hq_1(h) - \int_h^{\bar{h}} q_1(y)dy.$$

We claim that $\Phi(h)$ is left-continuous and bounded. First we show that $q_1(h)$ is left-continuous. Otherwise, there exists $h^{'} \in (\underline{h}, \bar{h})$ such that $q_1(h^{'}-) > q_1(h^{'})$. Now consider $\tilde{q}_1(h)$ (which is left-continuous at $h^{'}$) such that $\tilde{q}_1(h^{'}) = q_1(h^{'}-)$ and $\tilde{q}_1(h) = q_1(h)$ for any $h \neq h^{'}$. Defining $\tilde{\Phi}(h) \equiv V(\tilde{q}_1(h)) - h\tilde{q}_1(h) - \int_h^{\bar{h}} \tilde{q}_1(y)dy$, observe that $\tilde{\Phi}(h) = \Phi(h)$ for $h \neq h^{'}$ and $\tilde{\Phi}(h) > \Phi(h)$ when $h = h^{'}$. Then

$$\begin{aligned} \int_{[\underline{h},\bar{h}]} \tilde{\Phi}(h)dG(h) &= \int_{[\underline{h},\bar{h}]\backslash h'} \tilde{\Phi}(h)dG(h) + \tilde{\Phi}(h^{'})[G(h^{'}+) - G(h^{'}-)] \\ \geq & \int_{[\underline{h},\bar{h}]\backslash h'} \tilde{\Phi}(h)dG(h) + \Phi(h^{'})[G(h^{'}+) - G(h^{'}-)] = \int_{[\underline{h},\bar{h}]} \Phi(h)dG(h) \end{aligned}$$

with strict inequality if $G(h)$ is discontinuous at $h = h^{'}$. This is a contradiction. This implies in turn that $\Phi(h)$ is also left-continuous. Moreover, $\Phi(h)$ is bounded, since

$$\Phi(h) \leq \Phi(\underline{h}) \leq V(q_1(\underline{h})) - \underline{h}q_1(\underline{h}) \leq V(q^*(\underline{h})) - \underline{h}q^*(\underline{h}) < \infty$$

because of $\underline{h} > 0$, and

$$\Phi(h) \geq \Phi(\bar{h}) = V(q_1(\bar{h})) - \bar{h}q_1(\bar{h}) \geq 0$$

because of $V^{'}(q) > V^{'}(q_1(\bar{h})) \geq \bar{h}$ for $q < q_1(\bar{h})$ and $V(0) = 0$.

(v) We claim that $\Phi(h)$ is non-increasing in h and is not constant on $(\underline{h}, \bar{h})$. To show the former, note that for any h, we have

$$\begin{aligned} & \lim_{\epsilon \to 0+} \frac{\Phi(h+\epsilon) - \Phi(h)}{\epsilon} \\ = & \lim_{\epsilon \to 0+} (1/\epsilon)[V(q_1(h+\epsilon)) - (h+\epsilon)q_1(h+\epsilon) - \int_{h+\epsilon}^{\bar{h}} q_1(y)dy \\ - & [V(q_1(h)) - hq_1(h) - \int_h^{\bar{h}} q_1(y)dy]] \\ = & [V^{'}(\hat{q}(h)) - h] \lim_{\epsilon \to 0+} \frac{q_1(h+\epsilon) - q_1(h)}{\epsilon} \\ - & q_1(h+) + \lim_{\epsilon \to 0+} (1/\epsilon) \int_h^{h+\epsilon} q_1(y)dy \\ = & [V^{'}(\hat{q}(h)) - h] \lim_{\epsilon \to 0+} \frac{q_1(h+\epsilon) - q_1(h)}{\epsilon} \end{aligned}$$

for some $\hat{q}(h) \in [q_1(h+), q_1(h)]$. This is non-positive since $V^{'}(\hat{q}(h)) \leq V^{'}(q_1(h+)) \leq h$ and $\lim_{\epsilon \to 0+} \frac{q_1(h+\epsilon)-q_1(h)}{\epsilon} \leq 0$. Because of left-continuity of $\Phi(h)$, it implies that $\Phi(h)$ is non-increasing in h.

Next we show that $\Phi(h)$ is not constant on $(\underline{h}, \bar{h})$. First we consider the case that there exists $h \in (\underline{h}, \bar{h})$ such that $q_1(h+) < q_1(h-)$. Then

$$\begin{aligned} & \Phi(h+) \\ = \;& V(q_1(h+)) - hq_1(h+) - \int_h^{\bar{h}} q_1(y)dy] \\ < \;& V(q_1(h-)) - hq_1(h-) - \int_h^{\bar{h}} q_1(y)dy = \Phi(h-) \end{aligned}$$

The inequality follows from $V^{'}(q_1(h+)) > V^{'}(q_1(h-)) \geq V^{'}(q^*(h)) = h$. Therefore $\Phi(h)$ decreases discontinuously at h, implying that $\Phi(h)$ is not constant on $(\underline{h}, \bar{h})$. Second we consider the case that $q(h)$ is continuous on $(\underline{h}, \bar{h})$. Then from (ii) and (iii) above, there exists an interval (h^-, h^+) with the positive measure such that $q_1(h)$ is strictly decreasing and $V^{'}(q_1(h)) > h$ on (h^-, h^+). $\Phi(h)$ is continuous and almost everywhere differentiable (because of monotonicity of $q_1(h)$). At any point of differentiability,

$$\Phi^{'}(h) = [V^{'}(q_1(h)) - h]q_1^{'}(h).$$

This is negative almost everywhere on (h^-, h^+). Hence $\Phi(h)$ is strictly decreasing in h on (h^-, h^+).

(vi) Now consider the contracting problem with cost distribution $G_2(h)$. Since $q_1(h)$ is non-increasing in h, it is feasible for P to select this output schedule when the cost distribution is G_2. Hence $W(G_2) \geq \int_{\underline{h}}^{\bar{h}} \Phi(h)dG_2(h)$. Therefore if $\int_{\underline{h}}^{\bar{h}} \Phi(h)dG_2(h) > \int_{\underline{h}}^{\bar{h}} \Phi(h)dG_1(h) = W(G_1)$, it follows that $W(G_2) > W(G_1)$. Since $G_1(h)$ is right-continuous and $\Phi(h)$ is left-continuous and bounded, we can integrate by parts:

$$\int_{\underline{h}}^{\bar{h}} \Phi(h)dG_1(h) + \int_{\underline{h}}^{\bar{h}} G_1(h)d\Phi(h) = \Phi(\bar{h})G_1(\bar{h}) - \Phi(\underline{h})G_1(\underline{h}) = \Phi(\bar{h}).$$

Similarly for $G_2(h)$,

$$\int_{\underline{h}}^{\bar{h}} \Phi(h)dG_2(h) + \int_{\underline{h}}^{\bar{h}} G_2(h)d\Phi(h) = \Phi(\bar{h})G_2(\bar{h}) - \Phi(\underline{h})G_2(\underline{h}) = \Phi(\bar{h}).$$

Hence

$$\int_{\underline{h}}^{\bar{h}} \Phi(h)dG_2(h) - \int_{\underline{h}}^{\bar{h}} \Phi(h)dG_1(h) = \int_{\underline{h}}^{\bar{h}} [G_1(h) - G_2(h)]d\Phi(h).$$

By (iv) and $G_2(h) > G_1(h)$ for $h \in (\underline{h}, \bar{h})$, this is positive. ■

Proof of Proposition 4:

Step 1: For any $\eta \in \Pi$ and any closed interval $[\theta^{'}, \theta^{''}] \subset \Theta(\eta)$ such that $\underline{\theta}(\eta) < \theta^{'} < \theta^{''} < \bar{\theta}(\eta)$, there exists $\delta > 0$ such that $z(\cdot) \in Z(\eta)$ for any $z(\cdot)$ satisfying the following properties:

(i) $z(\theta)$ is increasing and differentiable with $|z(\theta) - \theta| < \delta$ and $|z^{'}(\theta) - 1| < \delta$ for any $\theta \in \Theta(\eta)$

(ii) $z(\theta) = \theta$ for any $\theta \notin [\theta^{'}, \theta^{''}]$.

Proof of Step 1

For arbitrary $\eta \in \Pi$ and arbitrary closed interval $[\theta^{'}, \theta^{''}] \subset \Theta(\eta)$ such that $\underline{\theta}(\eta) < \theta^{'} < \theta^{''} < \bar{\theta}(\eta)$, we choose ϵ_1 and ϵ_2 such that

$$\epsilon_1 \equiv \min_{\theta \in [\theta^{'}, \theta^{''}]} f(\theta \mid \eta)$$

and

$$\epsilon_2 \equiv \max_{\theta \in [\theta^{'}, \theta^{''}]} |f^{'}(\theta \mid \eta)|.$$

From our assumptions that $f(\theta \mid \eta)$ is continuously differentiable and positive on $\Theta(\eta)$, $\epsilon_1 > 0$, and ϵ_2 is positive and bounded above. We choose $\delta > 0$ such that

$$\delta \in (0, \frac{\epsilon_1}{\epsilon_1 + \epsilon_2}).$$

For this δ, it is obvious that there exists $z(\theta)$ which satisfies conditions (i) and (ii) of the statement. Define

$$\Lambda(\theta \mid \eta) \equiv (\theta - z(\theta))f(\theta \mid \eta) + F(\theta \mid \eta).$$

Since $z(\theta)$ is differentiable on $\Theta(\eta)$, $\Lambda(\theta \mid \eta)$ is also so. It is equal to $\Lambda(\theta \mid \eta) = F(\theta \mid \eta)$ on $\theta \notin [\theta', \theta'']$. For $\theta \in [\theta', \theta'']$,

$$\begin{aligned}
\frac{\partial \Lambda(\theta \mid \eta)}{\partial \theta} &= (2 - z'(\theta))f(\theta \mid \eta) + (\theta - z(\theta))f'(\theta \mid \eta) > (1-\delta)f(\theta \mid \eta) - \delta|f'(\theta \mid \eta)| \\
&\geq (1-\delta)\epsilon_1 - \delta\epsilon_2.
\end{aligned}$$

This is positive by the definition of $(\epsilon_1, \epsilon_2, \delta)$. Then $\Lambda(\theta \mid \eta)$ is increasing in θ on $\Theta(\eta)$ with $\Lambda(\underline{\theta}(\eta) \mid \eta) = 0$ and $\Lambda(\bar{\theta}(\eta) \mid \eta) = 1$. Since $z(\theta)$ is increasing in θ by the definition, it is preserved even by ironing rule. Therefore $z(\cdot) \in Z(\eta)$.
■

Step 2: There exist $\eta \in \Pi$ and an interval of θ with positive measure such that $\frac{F(\theta|\eta)}{f(\theta|\eta)} / \frac{F(\theta)}{f(\theta)}$ is increasing in θ.

The proof of Step 2

Define

$$A(\theta \mid \eta) \equiv \frac{F(\theta \mid \eta)}{f(\theta \mid \eta)} / \frac{F(\theta)}{f(\theta)} \equiv \frac{\int_{\underline{\theta}(\eta)}^{\theta} f(y) a(\eta|y) dy}{a(\eta|\theta)F(\theta)}.$$

If the result is false, $A(\theta \mid \eta)$ is non-increasing in $\theta \in (\underline{\theta}(\eta), \bar{\theta}(\eta))$ for all η. Then

$$\begin{aligned}
\partial A(\theta \mid \eta)/\partial\theta = \frac{1}{F(\theta)^2 a(\eta \mid \theta)^2}[F(\theta)a(\eta \mid \theta)^2 f(\theta) \\
- \int_{\underline{\theta}(\eta)}^{\theta} f(y)a(\eta \mid y)dy\{F(\theta)\partial a(\eta \mid \theta)/\partial\theta + f(\theta)a(\eta \mid \theta)\}] \leq 0
\end{aligned}$$

holds for $\theta \in (\underline{\theta}(\eta), \bar{\theta}(\eta))$. Equivalently

$$\partial a(\eta \mid \theta)/\partial\theta \geq \frac{f(\theta)}{F(\theta)}[1/A(\theta \mid \eta) - 1]a(\eta \mid \theta).$$

Define $\Pi(\theta) \equiv \{\eta \in \Pi \mid \theta \in (\underline{\theta}(\eta), \bar{\theta}(\eta))\}$. By $\Sigma_{\eta \in \Pi(\theta)} a(\eta \mid \theta) = 1$, $\Sigma_{\eta \in \Pi(\theta)} \partial a(\eta \mid \theta)/\partial\theta = 0$. This implies that

$$0 = \Sigma_{\eta \in \Pi(\theta)} \partial a(\eta \mid \theta)/\partial\theta \geq \frac{f(\theta)}{F(\theta)}[\Sigma_{\eta \in \Pi(\theta)} a(\eta \mid \theta)/A(\theta \mid \eta) - 1],$$

or $\Sigma_{\eta\in\Pi(\theta)}a(\eta \mid \theta)/A(\theta \mid \eta) \leq 1$ holds any for $\theta \in (\underline{\theta}, \bar{\theta})$. Since $1/A$ is convex in A and $\Sigma_{\eta\in\Pi(\theta)}a(\eta \mid \theta)A(\theta \mid \eta) = 1$,

$$\Sigma_{\eta\in\Pi(\theta)}a(\eta \mid \theta)/A(\theta \mid \eta) \geq 1/[\Sigma_{\eta\in\Pi(\theta)}a(\eta \mid \theta)A(\theta \mid \eta)] = 1$$

with strict inequality if there exists $\eta \in \Pi(\theta)$ such that $A(\theta \mid \eta) \neq 1$. This means that $A(\theta \mid \eta) = 1$ must hold for any $\eta \in \Pi(\theta)$ and any $\theta \in \Theta$. Then $h(\theta \mid \eta) = H(\theta)$ for any $(\theta, \eta) \in K$. This is a contradiction, since η is informative about θ. ■

Step 3:

From Step 2, we can choose $\eta^* \in \Pi$ and a closed interval $[\theta^{'}, \theta^{''}] \subset \Theta(\eta^*)$ such that $\underline{\theta}(\eta^*) < \theta^{'} < \theta^{''} < \bar{\theta}(\eta^*)$ and $A(\theta \mid \eta^*) \equiv \frac{F(\theta|\eta^*)}{f(\theta|\eta^*)}/\frac{F(\theta)}{f(\theta)}$ is increasing in θ on $[\theta^{'}, \theta^{''}]$. According to the procedure in Step 1, we select $\delta > 0$ for η^* and $[\theta^{'}, \theta^{''}]$. Then we also choose $\lambda > 0$, closed intervals $\Theta^L \subset [\theta^{'}, \theta^{''}]$ and $\Theta^H \subset [\theta^{'}, \theta^{''}]$,

$$\lambda < \frac{F(\theta)}{f(\theta)}/\frac{F(\theta \mid \eta^*)}{f(\theta \mid \eta^*)} \text{ for } \theta \in \Theta^L \equiv [\underline{\theta}^L, \bar{\theta}^L] \subset [\theta^{'}, \theta^{''}]$$

$$\lambda > \frac{F(\theta)}{f(\theta)}/\frac{F(\theta \mid \eta^*)}{f(\theta \mid \eta^*)} \text{ for } \theta \in \Theta^H \equiv [\underline{\theta}^H, \bar{\theta}^H] \subset [\theta^{'}, \theta^{''}]$$

with $\bar{\theta}^L < \underline{\theta}^H$. These conditions are equivalent to

$$H(\theta) - (1 - \lambda)\theta - \lambda h(\theta \mid \eta^*) > 0 \text{ for } \theta \in \Theta^L$$

and

$$H(\theta) - (1 - \lambda)\theta - \lambda h(\theta \mid \eta^*) < 0 \text{ for } \theta \in \Theta^H.$$

Step 4: Construction of $z(\theta \mid \eta)$

Now let us construct $z(\theta \mid \eta)$ which satisfies the following conditions.

(A) For $\eta \neq \eta^*$, $z(\theta \mid \eta) = \theta$ for any $\theta \in \Theta(\eta)$.

(B) For η^*, $z(\theta \mid \eta^*)$ satisfies

 (i) $z(\theta \mid \eta^*)$ is increasing and differentiable with $|z(\theta \mid \eta^*) - \theta| < \delta$ and $|z^{'}(\theta \mid \eta^*) - 1| < \delta$ for any $\theta \in \Theta(\eta^*)$

(ii) $z(\theta \mid \eta^*) = \theta$ for any $\theta \notin \Theta^H \cup \Theta^L$

(iii) For $\theta \in \Theta^L$, $z(\theta \mid \eta^*)$ satisfies (a) $z(\theta \mid \eta^*) \leq \theta$ with strict inequality for some subinterval of Θ^L of positive measure, and (b) $H(z) - (1-\lambda)z - \lambda h(\theta \mid \eta^*) > 0$ for any $z \in [z(\theta \mid \eta^*), \theta]$.

(iv) For $\theta \in \Theta^H$, $z(\theta \mid \eta^*)$ satisfies (a) $z(\theta \mid \eta^*) \geq \theta$ with strict inequality for some some subinterval of Θ^H of positive measure, (b) $z(\theta \mid \eta^*) < h(\theta \mid \eta^*)$ and (c) $H(z) - (1-\lambda)z - \lambda h(\theta \mid \eta^*) < 0$ for any $z \in [\theta, z(\theta \mid \eta^*)]$.

(v) $E[(z(\theta \mid \eta^*) - h(\theta \mid \eta^*))q^{NS}(z(\theta \mid \eta^*)) + \int_{z(\theta|\eta^*)}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*] = 0.$

We now argue there exists $z^*(\theta \mid \eta^*)$ which satisfies (B(i)-(v)). Step 3 guarantees that we can select $z(\theta \mid \eta^*)$ which satisfies (B(i)-(iv)). Since

$$(z - h(\theta \mid \eta^*))q^{NS}(z) + \int_z^{\bar{\theta}(\eta^*)} q^{NS}(y)dy$$

is increasing in z for $z < h(\theta \mid \eta^*)$, and

$$E[(\theta - h(\theta \mid \eta^*))q^{NS}(\theta) + \int_\theta^{\bar{\theta}(\eta^*)} q^{NS}(y)dy \mid \eta^*] = 0,$$

the choice of $z(\theta \mid \eta^*) \leq \theta$ on Θ_L (or $z(\theta \mid \eta^*) \geq \theta$ on Θ_H) reduces (or raises)

$$E[(z(\theta \mid \eta^*) - h(\theta \mid \eta^*))q^{NS}(z(\theta \mid \eta^*)) + \int_{z(\theta|\eta^*)}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*]$$

away from zero. For any pair of parameters α_H, α_L lying in $[0,1]$, define a function $z_{\alpha_L,\alpha_H}(\theta|\eta^*)$ which equals $(1-\alpha_L)z(\theta|\eta^*) + \alpha_L\theta$ on Θ_L, equals $(1-\alpha_H)z(\theta|\eta^*) + \alpha_H\theta$ on Θ_H and equals θ elsewhere. It is easily checked that any such function also satisfies conditions (B(i)–(iv)). Define

$$Q(\alpha_L, \alpha_U) \equiv E[(z_{\alpha_L,\alpha_H}(\theta \mid \eta^*) - h(\theta \mid \eta^*))q^{NS}(z_{\alpha_L,\alpha_H}(\theta \mid \eta^*)) + \int_{z_{\alpha_L,\alpha_H}(\theta|\eta^*)}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*].$$

Then Q is continuously differentiable, strictly increasing in α_L and strictly decreasing in α_H. By (B(v)), $Q(1,1) = 0$. The Implicit Function Theorem ensures existence of α_L^*, α_H^* both smaller than 1 such that $Q(\alpha_L^*, \alpha_H^*) = 0$. Hence the function $z_{\alpha_L^*,\alpha_H^*}(\theta|\eta^*)$ satisfies (B(i)-(v)).

Step 5

By Step 1, $z(\cdot \mid \eta)$ constructed in Step 4 is in $Z(\eta)$ for any $\eta \in \Pi$. Consider the following allocation (u_A, u_S, q):

$$q(\theta, \eta) = q^{NS}(z(\theta \mid \eta))$$

$$u_A(\theta, \eta) = \int_{\theta}^{\bar{\theta}} q^{NS}(z(y \mid \eta))dy$$

$$u_S(\theta, \eta) = X^{NS}(z(\theta \mid \eta)) - \theta q^{NS}(z(\theta \mid \eta)) - \int_{\theta}^{\bar{\theta}(\eta)} q^{NS}(z(y \mid \eta))dy - \int_{\bar{\theta}(\eta)}^{\bar{\theta}} q^{NS}(y)dy.$$

where

$$X^{NS}(z(\theta \mid \eta)) \equiv z(\theta \mid \eta)q^{NS}(z(\theta \mid \eta)) + \int_{z(\theta|\eta)}^{\bar{\theta}} q^{NS}(z)dz.$$

The construction of $z(\theta \mid \eta)$ implies that $z(\bar{\theta}(\eta) \mid \eta) \leq \bar{\theta}$ for any $\eta \in \Pi$. Hence

$$X^{NS}(z(\theta \mid \eta)) - z(\theta \mid \eta)q^{NS}(z(\theta \mid \eta)) \geq 0$$

for any $(\theta, \eta) \in K$ and

$$E[u_S(\theta, \eta) \mid \eta] = 0$$

from (A) and (B(v)). Then (u_A, u_S, q) is a WCP allocation satisfying interim PCs. Now we show that this allocation generates a higher payoff to P than the optimal allocation in NS. P's resulting expected payoff conditional on η^* (maintaining the expected payoff conditional on $\eta \neq \eta^*$ unchanged) is:

$$E[V(q^{NS}(z(\theta \mid \eta^*))) - z(\theta \mid \eta^*)q^{NS}(z(\theta \mid \eta^*)) - \int_{z(\theta|\eta^*)}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*].$$

With $E[u_S(\theta,\eta^*) \mid \eta^*] = 0$, this is equal to

$$
\begin{aligned}
& E[V(q^{NS}(z(\theta \mid \eta^*))) - z(\theta \mid \eta^*) q^{NS}(z(\theta \mid \eta^*)) - \int_{z(\theta|\eta^*)}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*] \\
+ \quad & \lambda E[(z(\theta \mid \eta^*) - h(\theta \mid \eta^*)) q^{NS}(z(\theta \mid \eta^*)) + \int_{z(\theta|\eta^*)}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*] \\
= \quad & E[V(q^{NS}(z(\theta \mid \eta^*))) - [(1-\lambda) z(\theta \mid \eta^*) + \lambda h(\theta \mid \eta^*)] q^{NS}(z(\theta \mid \eta^*)) \\
- \quad & (1-\lambda) \int_{z(\theta|\eta^*)}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*] \mid \eta^*] \\
- \quad & \lambda \int_{\bar{\theta}(\eta^*)}^{\bar{\theta}} q^{NS}(z)dz
\end{aligned}
$$

On the other hand,

$$
E[(\theta - h(\theta \mid \eta^*)) q^{NS}(\theta) + \int_{\theta}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*] = 0.
$$

P's expected payoff conditional on η^* in the optimal allocation in NS is:

$$
\begin{aligned}
& E[V(q^{NS}(\theta)) - \theta q^{NS}(\theta) - \int_{\theta}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*] \\
= \quad & E[V(q^{NS}(\theta)) - \theta q^{NS}(\theta) - \int_{\theta}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*] \\
+ \quad & \lambda E[(\theta - h(\theta \mid \eta^*)) q^{NS}(\theta) + \int_{\theta}^{\bar{\theta}(\eta^*)} q^{NS}(z)dz \mid \eta^*] \\
= \quad & E[V(q^{NS}(\theta)) - [(1-\lambda)\theta + \lambda h(\theta \mid \eta)] q^{NS}(\theta) - (1-\lambda) \int_{\theta}^{\bar{\theta}} q^{NS}(z)dz \mid \eta^*] \\
- \quad & \lambda \int_{\bar{\theta}(\eta^*)}^{\bar{\theta}} q^{NS}(z)dz
\end{aligned}
$$

The difference between two payoffs is

$$
\begin{aligned}
& E[V(q^{NS}(z(\theta \mid \eta^*))) - [(1-\lambda)z(\theta \mid \eta^*) + \lambda h(\theta \mid \eta^*)]q^{NS}(z(\theta \mid \eta^*)) \\
- & (1-\lambda)\int_{z(\theta|\eta^*)}^{\bar\theta} q^{NS}(z)dz \mid \eta^*] \\
- & E[V(q^{NS}(\theta)) - [(1-\lambda)\theta + \lambda h(\theta \mid \eta^*)]q^{NS}(\theta) - (1-\lambda)\int_{\theta}^{\bar\theta} q^{NS}(z)dz \mid \eta^*] \\
= & E[\int_{\theta}^{z(\theta|\eta^*)} [V'(q^{NS}(z)) - \{(1-\lambda)z + \lambda h(\theta \mid \eta^*)\}]q^{NS'}(z)dz \mid \eta^*] \\
= & E[\int_{\theta}^{z(\theta|\eta^*)} [H(z) - \{(1-\lambda)z + \lambda h(\theta \mid \eta^*)\}]q^{NS'}(z)dz \mid \eta^*].
\end{aligned}
$$

The second equality uses $V'(q^{NS}(z)) = H(z)$. From the construction of $z(\theta \mid \eta^*)$ in Step 4 and $q^{NS'}(z) < 0$, this is positive. We have thus found an implementable allocation generating a higher payoff to P in CS compared to the optimal allocation in NS. ■

Proof of Proposition 6:

Since $f(\theta \mid \eta^*)$ is decreasing in θ, $h(\theta \mid \eta^*)$ is increasing in θ, implying $h(\theta \mid \eta^*) = \hat h(\theta \mid \eta^*)$. Since $\frac{f(\theta|\eta^*)}{f(\theta|\eta)}$ is strictly decreasing in θ for any $\eta \neq \eta^*$, $\frac{f(\theta'|\eta^*)}{f(\theta|\eta^*)} > \frac{f(\theta'|\eta)}{f(\theta|\eta)}$ for $\theta > \theta'$. $\Theta(\eta) = \Theta(\eta^*) = \Theta$ then implies

$$\frac{F(\theta \mid \eta^*)}{f(\theta \mid \eta^*)} = \int_{\underline\theta}^{\theta} \frac{f(\theta' \mid \eta^*)}{f(\theta \mid \eta^*)} d\theta' > \int_{\underline\theta}^{\theta} \frac{f(\theta' \mid \eta)}{f(\theta \mid \eta)} d\theta' = \frac{F(\theta \mid \eta)}{f(\theta \mid \eta)}.$$

Hence $h(\theta \mid \eta^*) > h(\theta \mid \eta)$ for $\theta \in (\underline\theta, \bar\theta]$ and $h(\underline\theta \mid \eta^*) = h(\underline\theta \mid \eta) = \underline\theta$. The ironing procedure then ensures that $\hat h(\theta \mid \eta^*) > \hat h(\theta \mid \eta)$ for any $\theta > \underline\theta$ and any $\eta \neq \eta^*$. Thus $\hat h(\bar\theta|\eta^*) > \hat h(\bar\theta|\eta)$ while $\hat h(\underline\theta|\eta^*) = \hat h(\underline\theta|\eta) = \underline\theta$ for $\eta \neq \eta^*$, i.e., the range of $\hat h$ conditional on η^* includes the range of $\hat h$ conditional on η. Since $h(\theta \mid \eta^*) = \hat h(\theta \mid \eta^*)$ is strictly increasing and continuously differentiable, $q^*(\hat h(\theta \mid \eta^*))$ is also continuously differentiable and strictly decreasing in θ.

Suppose the result is false, and the second best allocation $(u_A^{SB}(\theta,\eta), u_S^{SB}(\theta,\eta), q^{SB}(\theta,\eta))$ is implementable with weak collusion. Then Proposition 1 implies existence of $\pi(\cdot \mid \eta) \in Y(\eta)$ such that

$$q^{SB}(\theta, \eta) = q^*(\hat h(\theta \mid \eta))$$

$$X^{SB}(\theta,\eta) - z(\theta \mid \eta)q^{SB}(\theta,\eta) \geq 0$$
$$X^{SB}(\theta,\eta) - z(\theta \mid \eta)q^{SB}(\theta,\eta) \geq X^{SB}(\theta^{'},\eta^{'}) - z(\theta \mid \eta)q^{SB}(\theta^{'},\eta^{'})$$

where $z(\theta \mid \eta) \equiv z(\theta, \pi(\theta \mid \eta), \eta)$ and

$$X^{SB}(\theta,\eta) \equiv u_A^{SB}(\theta,\eta) + u_S^{SB}(\theta,\eta) + \theta q^{SB}(\theta,\eta).$$

Step 1: $z(\theta \mid \eta) \in [z(\underline{\theta} \mid \eta^*), z(\bar{\theta} \mid \eta^*)]$ holds for any (θ,η).

The proof is as follows. Since $\hat{h}(\theta \mid \eta) < \hat{h}(\theta \mid \eta^*)$ for any $\theta > \underline{\theta}$ and $\eta \neq \eta^*$,

$$q^{SB}(\theta,\eta^*) = q^*(\hat{h}(\theta \mid \eta^*)) < q^*(\hat{h}(\theta \mid \eta)) = q^{SB}(\theta,\eta).$$

Then $z(\theta \mid \eta^*) \geq z(\theta \mid \eta)$ follows from the coalition incentive constraints.

If on the other hand $z(\underline{\theta}|\eta) < z(\underline{\theta}|\eta^*)$, there exists a non-degenerate interval T of θ for which $z(\theta|\eta) \in (z(\underline{\theta}|\eta), z(\underline{\theta}|\eta^*))$. The second-best output in either state $(\underline{\theta},\eta)$ or $(\underline{\theta},\eta^*)$ is the first-best level $q^*(\underline{\theta})$ corresponding to cost $\underline{\theta}$. The coalitional incentive constraints imply output must be constant over T given η, so must equal the first-best $q^*(\underline{\theta})$ corresponding to cost $\underline{\theta}$. But $\hat{h}(\theta,\eta) \geq \theta$ for every $\theta \in T$, implying $q^{SB}(\theta,\eta) = q^*(\hat{h}(\theta,\eta)) \leq q^*(\theta) < q^*(\underline{\theta})$, and we obtain a contradiction.

In what follows, we denote $[z(\underline{\theta} \mid \eta^*), z(\bar{\theta} \mid \eta^*)]$ by $[\underline{z}, \bar{z}]$.

Step 2:

Now we claim that there exists $\phi(\cdot) : [\underline{h}, \bar{h}] \to [\underline{z}, \bar{z}]$ which satisfies

(i) $z(\theta \mid \eta) = \phi(\hat{h}(\theta \mid \eta))$.

(ii) $\phi(h)$ is continuous, and non-decreasing in h.

(iii) $h - \phi(h)$ is non-negative and increasing in h.

First we show that for any (θ,η) and $(\theta^{'},\eta^{'})$ such that $\hat{h}(\theta \mid \eta) = \hat{h}(\theta^{'} \mid \eta^{'})$, $z(\theta \mid \eta) = z(\theta^{'} \mid \eta^{'})$. Otherwise, there exists $(\theta^{'},\eta^{'})$ and $(\theta^{''},\eta^{''})$ such that $\hat{h}(\theta^{'} \mid \eta^{'}) = \hat{h}(\theta^{''} \mid \eta^{''})$ and $z(\theta^{'} \mid \eta^{'}) \neq z(\theta^{''} \mid \eta^{''})$. Suppose $z(\theta^{'} \mid \eta^{'}) < z(\theta^{''} \mid \eta^{''})$ without loss of generality. By Step 1 and the continuity of $z(\theta \mid \eta^*)$, there exists θ_1 and θ_2 $(\theta_1 < \theta_2)$ such that

$$z(\theta_1 \mid \eta^*) = z(\theta^{'} \mid \eta^{'}) < z(\theta^{''} \mid \eta^{''}) = z(\theta_2 \mid \eta^*).$$

Since $z(\theta \mid \eta^*)$ is continuous in θ and non-decreasing in θ,

$$z(\theta^{'} \mid \eta^{'}) \leq z(\theta \mid \eta^*) \leq z(\theta^{''} \mid \eta^{''})$$

for any $\theta \in [\theta_1, \theta_2]$. The coalitional incentive constraints imply

$$q^{SB}(\theta^{'}, \eta^{'}) \geq q^{SB}(\theta, \eta^*) \geq q^{SB}(\theta^{''}, \eta^{''})$$

for any $\theta \in [\theta_1, \theta_2]$. On the other hand $\hat{h}(\theta^{'} \mid \eta^{'}) = \hat{h}(\theta^{''} \mid \eta^{''})$ implies $q^{SB}(\theta^{'}, \eta^{'}) = q^{SB}(\theta^{''}, \eta^{''})$. Therefore $q^{SB}(\theta, \eta^*) = q^{SB}(\theta^{'}, \eta^{'}) = q^{SB}(\theta^{''}, \eta^{''})$ for any $\theta \in [\theta_1, \theta_2]$. This contradicts the property that $q^{SB}(\theta, \eta^*)$ must be strictly decreasing in θ.

Hence there exists a function $\phi(\cdot) : [\underline{h}, \bar{h}] \to [\underline{z}, \bar{z}]$ such that $z(\theta \mid \eta) = \phi(\hat{h}(\theta \mid \eta))$. Since $z(\theta \mid \eta^*)$ and $\hat{h}(\theta \mid \eta^*)$ are continuous in θ, $\phi(h)$ must be continuous.

Second we show that $\phi(h)$ is non-decreasing in h. For any (θ, η) and $(\theta^{'}, \eta^{'})$ such that $\hat{h}(\theta \mid \eta) < \hat{h}(\theta^{'} \mid \eta^{'})$,

$$q^{SB}(\theta, \eta) = q^*(\hat{h}(\theta \mid \eta)) > q^*(\hat{h}(\theta^{'} \mid \eta^{'})) = q^{SB}(\theta^{'}, \eta^{'}).$$

The coalitional incentive constraints then imply $z(\theta \mid \eta) \leq z(\theta^{'} \mid \eta^{'})$.

Third we show $h - \phi(h)$ is non-negative and increasing in h. Since $q^{SB}(\theta, \eta^*) = q^*(\hat{h}(\theta \mid \eta^*))$ is strictly decreasing in θ, the pooling region $\Theta(\pi(\cdot \mid \eta^*), \eta^*)$ must be empty. Hence it must be the case that

$$z(\theta \mid \eta^*) = \theta + \frac{F(\theta \mid \eta^*) - \Lambda(\theta \mid \eta^*)}{f(\theta \mid \eta^*)},$$

implying

$$\frac{\Lambda(\theta \mid \eta^*)}{f(\theta \mid \eta^*)} = \hat{h}(\theta \mid \eta^*) - \phi(\hat{h}(\theta \mid \eta^*)).$$

The LHS is non-negative and increasing in θ, since $f(\theta \mid \eta^*)$ is decreasing in θ and $\Lambda(\theta \mid \eta^*)$ is non-negative and non-decreasing in θ. So $h - \phi(h)$ must be non-negative and increasing in $h \in [\underline{h}, \bar{h}]$.

Step 3:

Define $R(z) \equiv \max_{(\tilde{\theta}, \tilde{\eta}) \in K}[X^{SB}(\tilde{\theta}, \tilde{\eta}) - zq^{SB}(\tilde{\theta}, \tilde{\eta})]$ for any $z \in [\underline{z}, \bar{z}]$. Then

$$R(z(\theta \mid \eta)) = X^{SB}(\theta, \eta) - z(\theta \mid \eta)q^{SB}(\theta, \eta)$$

and by the Envelope Theorem, $R^{'}(z(\theta \mid \eta)) = -q^{SB}(\theta,\eta) = -q^{*}(\hat{h}(\theta \mid \eta))$. It also implies $R^{'}(\phi(h)) = -q^{*}(h)$. Then S's interim payoff is

$$\begin{aligned} & E[X^{SB}(\theta,\eta) - h(\theta \mid \eta)q^{SB}(\theta,\eta) \mid \eta] \\ = \; & E[X^{SB}(\theta,\eta) - z(\theta \mid \eta)q^{SB}(\theta,\eta) + (z(\theta \mid \eta) - h(\theta \mid \eta))q^{SB}(\theta,\eta) \mid \eta] \\ = \; & E[R(\phi(\hat{h}(\theta \mid \eta))) + (\phi(\hat{h}(\theta \mid \eta)) - \hat{h}(\theta \mid \eta))q^{*}(\hat{h}(\theta \mid \eta)) \mid \eta] \end{aligned}$$

with the last equality using the property of the ironing rule.

Next define

$$L(h) \equiv R(\phi(h)) + (\phi(h) - h)q^{*}(h).$$

$L(h)$ is continuous and differentiable almost everywhere, since the monotonicity implies the differentiability of $\phi(h)$ almost everywhere. If the second best allocation is implementable with weak collusion, $E[L(\hat{h}(\theta \mid \eta)) \mid \eta] = 0$ holds for any η. The first derivative of $L(h)$ is

$$L^{'}(h) = (\phi(h) - h)q^{*'}(h) - q^{*}(h).$$

Since $q^{*}(h)$ is continuously differentiable, $L^{'}(h)$ is continuous and also differentiable almost everywhere and

$$L^{''}(h) = (\phi^{'}(h) - 1)q^{*'}(h) + (\phi(h) - h)q^{*''}(h) - q^{*'}(h).$$

By using $V^{'}(q^{*}(h)) = h$, we can show that $V^{'''}(q) \leq 0$ implies $q^{*''}(h) \leq 0$, and $0 < V^{'''}(q) \leq \frac{(V^{''}(q))^2}{V^{'}(q)}$ implies $q^{*''}(h) > 0$ and $hq^{*''}(h) + q^{*'}(h) < 0$. By $\phi^{'}(h) - 1 < 0$ and $\phi(h) - h \leq 0$, it follows that $L^{''}(h) > 0$.

The strict convexity of L then implies $L(h) > L(h^{'}) - (h^{'} - h)L^{'}(h^{'})$ for any $h \neq h^{'}$. Hence

$$\begin{aligned} & E[L(\hat{h}(\theta \mid \eta^{*})) \mid \eta^{*}] = E[L(h(\theta \mid \eta^{*})) \mid \eta^{*}] \\ > \; & E[L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - h(\theta \mid \eta^{*})]L^{'}(\hat{h}(\theta \mid \eta)) \mid \eta^{*}] \\ = \; & E[L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - \theta]L^{'}(\hat{h}(\theta \mid \eta)) + \int_{\theta}^{\bar{\theta}} L^{'}(\hat{h}(y \mid \eta))dy \mid \eta^{*}] \end{aligned}$$

for any $\eta \neq \eta^{*}$. $L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - \theta]L^{'}(\hat{h}(\theta \mid \eta)) + \int_{\theta}^{\bar{\theta}} L^{'}(\hat{h}(y \mid \eta))dy$ is non-increasing in θ, since

$$-[\hat{h}(\theta \mid \eta) - \theta]L^{''}(\hat{h}(\theta \mid \eta)) < 0$$

and is strictly decreasing in θ over some interval (since the ironing rule ensures $\hat{h}(\theta \mid \eta)$ is continuous with $\hat{h}(\underline{\theta} \mid \eta) = \underline{\theta}$ and $\hat{h}(\bar{\theta} \mid \eta) > \bar{\theta}$). Then property (ii) implies $F(\theta \mid \eta^*) > F(\theta \mid \eta)$ for $\theta \in (\underline{\theta}, \bar{\theta})$ and for any $\eta \neq \eta^*$. A first order stochastic dominance argument then ensures

$$
\begin{aligned}
& E[L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - \theta]L^{'}(\hat{h}(\theta \mid \eta)) + \int_{\theta}^{\bar{\theta}} L^{'}(\hat{h}(y \mid \eta))dy \mid \eta^*] \\
> \; & E[L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - \theta]L^{'}(\hat{h}(\theta \mid \eta)) + \int_{\theta}^{\bar{\theta}} L^{'}(\hat{h}(y \mid \eta))dy \mid \eta] \\
= \; & E[L(\hat{h}(\theta \mid \eta)) - [\hat{h}(\theta \mid \eta) - h(\theta \mid \eta)]L^{'}(\hat{h}(\theta \mid \eta)) \mid \eta] \\
= \; & E[L(\hat{h}(\theta \mid \eta)) \mid \eta].
\end{aligned}
$$

where the last equality utilizes a property of the ironing transformation. Therefore S must earn a positive rent in state η^*, as $E[L(h(\theta \mid \eta^*)) \mid \eta^*] > E[L(\hat{h}(\theta \mid \eta)) \mid \eta] \geq 0$. This is a contradiction. ■

Proof of Proposition 10

Step 1 and 2 are technical steps needed to prepare for the proof of the statements in Step 3 and 4.

Step 1: There exists $\theta^{'}$, $\theta^{''}$ and $\epsilon > 0$ such that

(i) $\theta_1^{SB} < \theta^{'} < \theta^{NS} < \theta^{''} < \theta_2^{SB}$

(ii)For any $\eta \in \{\eta_1, \eta_2\}$ and any $\epsilon^{'} \in (0, \epsilon]$, there exists $z(\cdot \mid \eta) \in Z(\eta)$ such that $z(\theta \mid \eta) = \theta - \epsilon^{'}$ on $[\theta^{'}, \theta^{''}]$.

Proof of Step 1

The proof is based on the construction. Evidently $\underline{\theta} < \theta_1^{SB} < \theta_2^{SB} < \bar{\theta}$ by $V \in (\underline{\theta}, \bar{\theta})$. Then choose $\epsilon > 0$ which satisfies

(i) $\epsilon < \frac{\min_{[\theta_1^{SB}, \theta_2^{SB}]} |f(\theta|\eta)|}{\max_{[\theta_1^{SB}, \theta_2^{SB}]} |f_\theta(\theta|\eta)|}$ for $\eta \in \{\eta_1, \eta_2\}$

(ii) $\epsilon < \theta^{NS} - \theta_1^{SB}$

(iii) $\epsilon < \frac{F(\theta_2^{SB}|\eta) - F(\theta^{NS}|\eta)}{f(\theta^{NS}|\eta)}$ for $\eta \in \{\eta_1, \eta_2\}$.

The continuous differentiability of $f(\theta \mid \eta)$ guarantees the existence of $\epsilon > 0$ satisfying (i)-(iii). (i) implies that for any $\eta \in \{\eta_1, \eta_2\}$ and for any $\epsilon^{'} \leq \epsilon$, $F(\theta \mid \eta) + \epsilon^{'} f(\theta \mid \eta)$ is increasing in θ on $[\theta_1^{SB}, \theta_2^{SB}]$. Now we select $\theta^{'}$ and $\theta^{''}$ as follows:

$$\theta^{'} \equiv \theta_1^{SB} + \epsilon$$

$$\theta^{''} \equiv \min\{\theta_1^{''}, \theta_2^{''}\}$$

for $\theta_1^{''}$ and $\theta_2^{''}$ satisfying

$$F(\theta_1^{''} \mid \eta_1) + \epsilon f(\theta_1^{''} \mid \eta_1) = F(\theta_1^{SB} \mid \eta_1)$$

$$F(\theta_2^{''} \mid \eta_2) + \epsilon f(\theta_2^{''} \mid \eta_2) = F(\theta_2^{SB} \mid \eta_2).$$

(ii) implies $\theta_1^{SB} < \theta^{'} < \theta^{NS}$. By (i) and (iii), $\theta^{NS} < \theta^{''} < \theta_2^{SB}$. Now for $\eta_i \in \{\eta_1, \eta_2\}$ and some $\epsilon^{'} \in (0, \epsilon]$, define $\Lambda(\theta \mid \eta_i)$ as follows

- $\Lambda(\theta \mid \eta_i) \equiv F(\theta \mid \eta_i)$ for $\theta \in [\underline{\theta}, \theta_1^{SB}) \cup (\theta_2^{SB}, \bar{\theta}]$
- $\Lambda(\theta \mid \eta_i) \equiv \min\{F(\theta \mid \eta_i) + \epsilon^{'} f(\theta \mid \eta_i), F(\theta_i^{SB} \mid \eta_i)\}$ for $\theta \in [\theta_1^{SB}, \theta_2^{SB}]$

$\Lambda(\theta \mid \eta_i)$ is non-decreasing in θ and $\Lambda(\underline{\theta} \mid \eta_i) = 0$ and $\Lambda(\bar{\theta} \mid \eta_i) = 1$. Then $\pi(\theta) \equiv \theta + \frac{F(\theta|\eta_i) - \Lambda(\theta|\eta_i)}{f(\theta|\eta_i)} = \theta - \epsilon^{'}$ for $\theta \in [\theta^{'}, \theta^{''}]$, since

$$F(\theta^{''} \mid \eta_i) + \epsilon^{'} f(\theta^{''} \mid \eta_i) \leq F(\theta_i^{SB} \mid \eta_i).$$

Moreover $\pi(\theta) < \pi(\theta^{'})$ for any $\theta < \theta^{'}$, since $\pi(\theta_1^{SB}) = \theta_1^{SB} \leq \pi(\theta^{'}) = \theta^{'} - \epsilon^{'}$ and $\pi(\theta) > \pi(\theta^{''})$ for any $\theta > \theta^{''}$. The ironing procedure implies that $z(\theta \mid \eta_i) = \pi(\theta) = \theta - \epsilon^{'}$ on $[\theta^{'}, \theta^{''}]$.

Step 2: There exists θ_L, θ_U such that

(i) $\theta_1^{SB} < \theta_L < \theta^{NS} < \theta_U < \theta_2^{SB}$

(ii) $(V - \theta)F(\theta \mid \eta_1) > (V - \theta^{NS})F(\theta^{NS} \mid \eta_1)$ *for any* $\theta \in (\theta_L, \theta^{NS})$

(iii) $(V - \theta)F(\theta \mid \eta_2) > (V - \theta^{NS})F(\theta^{NS} \mid \eta_2)$ *for any* $\theta \in (\theta^{NS}, \theta_U)$

Proof of Step 2

By the definition of θ_1^{SB}, it is true that

$$(V - \theta_1^{SB})F(\theta_1^{SB} \mid \eta_1) > (V - \theta^{NS})F(\theta^{NS} \mid \eta_1).$$

However it does not mean that $(V - \theta)F(\theta \mid \eta_1) > (V - \theta^{NS})F(\theta^{NS} \mid \eta_1)$ for any $\theta \in [\theta_1^{SB}, \theta^{NS})$, unless $h(\theta \mid \eta_1)$ is non-decreasing in θ. However this inequality holds for θ sufficiently close to θ^{NS}.

$$\begin{aligned} & \frac{\partial[(V-\theta)F(\theta \mid \eta_1)]}{\partial \theta}|_{\theta=\theta^{NS}} = f(\theta^{NS} \mid \eta_1)(V - h(\theta^{NS} \mid \eta_1)) \\ = \quad & f(\theta^{NS} \mid \eta_1)(H(\theta^{NS}) - h(\theta^{NS} \mid \eta_1)) < 0 \end{aligned}$$

The continuity of $(V - \theta)F(\theta \mid \eta_1)$ implies that there exists $\theta_L \in [\theta_1^{SB}, \theta^{NS})$ satisfying (ii). A similar argument ensures the existence of θ_U satisfying (i) and (iii).

Step 3: There exists $(\theta_1, \theta_2, \hat{\theta})$, $z(\cdot \mid \eta_1) \in Z(\eta_1)$ *and* $z(\cdot \mid \eta_2) \in Z(\eta_2)$ *such that*

(i) $\max\{\hat{\theta}, \theta_L\} < \theta_1 < \theta^{NS} < \theta_2 < \theta_U$

(ii) $\hat{\theta} = z(\theta_1 \mid \eta_1) = z(\theta_2 \mid \eta_2)$

(iii) For each $\eta_i \in \{\eta_1, \eta_2\}$, $z(\theta \mid \eta_i)$ *is increasing in* θ *at* θ_i

(iv) $(\theta_1 - \hat{\theta})F(\theta_1 \mid \eta_1) = (\theta_2 - \hat{\theta})F(\theta_2 \mid \eta_2)$

Proof of Step 3

For θ', θ'' and ϵ selected in Step 1 and θ_L, θ_U in Step 2, choose θ_1 and θ_2 such that

(a) $\max\{\theta', \theta_L\} < \theta_1 < \theta^{NS} < \theta_2 < \min\{\theta'', \theta_U\}$

(b) $0 < (\theta_2 - \theta_1)\frac{F(\theta_1|\eta_1)}{F(\theta_1|\eta_1) - F(\theta_2|\eta_2)} < \epsilon$.

Since $F(\theta^{NS} \mid \eta_1) > F(\theta^{NS} \mid \eta_2)$, such θ_1 and θ_2 always exist. Define ϵ_1 and ϵ_2 as follows:

$$\epsilon_1 \equiv [\theta_2 - \theta_1]\frac{F(\theta_2 \mid \eta_2)}{F(\theta_1 \mid \eta_1) - F(\theta_2 \mid \eta_2)}$$

$$\epsilon_2 \equiv [\theta_2 - \theta_1]\frac{F(\theta_1 \mid \eta_1)}{F(\theta_1 \mid \eta_1) - F(\theta_2 \mid \eta_2)}.$$

Since (b) implies $F(\theta_1 \mid \eta_1) > F(\theta_2 \mid \eta_2)$, $0 < \epsilon_1 < \epsilon_2 < \epsilon$. By Step 1, we can choose $z(\cdot \mid \eta_i) \in Z(\eta_i)$ such that $z(\theta \mid \eta_i) = \theta - \epsilon_i$ on $[\theta', \theta'']$ for $i \in \{1, 2\}$. It is evident that $z(\cdot \mid \eta_i)$ satisfies (iii). By the definition of ϵ_1 and ϵ_2, it is evident that $z(\theta_1 \mid \eta_1) = z(\theta_2 \mid \eta_2)$. With the definition of $\hat{\theta}$ such that $\hat{\theta} \equiv z(\theta_1 \mid \eta_1) = z(\theta_2 \mid \eta_2)$, $\hat{\theta} < \theta_1$ which satisfies (i), and (iv) is also automatically satisfied.

Step 4:

Based on $(\theta_1, \theta_2, \hat{\theta})$ and $z(\cdot \mid \eta_i)$ in Step 3, consider the allocation

$$(u_A(\theta, \eta), u_S(\theta, \eta), q(\theta, \eta))$$

as follows:

- $u_A(\theta, \eta_i) = \max\{\theta_i - \theta, 0\}$
- $u_S(\theta, \eta_i) = \hat{\theta} - \theta_i + K$ for $\theta \leq \theta_i$ and K for $\theta > \theta_i$ where

$$K \equiv (\theta_1 - \hat{\theta})F(\theta_1 \mid \eta_1) = (\theta_2 - \hat{\theta})F(\theta_2 \mid \eta_2) > 0$$

- $q(\theta, \eta_i) = \tilde{q}(z(\theta \mid \eta_i))$ for $\tilde{q}(z)$ satisfying $\tilde{q}(z) = 1$ for $z \leq \hat{\theta}$ and 0 for $z > \hat{\theta}$.

We claim this is implementable in the weak collusion. First consider the agent's participation and incentive constraint. It is evident that $u_A(\theta, \eta_i) \geq 0$ for any (θ, η_i). We can also show

$$u_A(\theta, \eta_i) \geq u_A(\theta', \eta_i) + (\theta' - \theta)q(\theta', \eta)$$

for any θ, θ' and any $\eta_i \in \{\eta_1, \eta_2\}$, since

$$\max\{\theta_i - \theta, 0\} \geq \max\{\theta_i - \theta', 0\} + (\theta' - \theta) = \theta_i - \theta$$

for any θ' such that $z(\theta', \eta_i) \leq \hat{\theta}$ or equivalently $\theta' \leq \theta_i$, and

$$\max\{\theta_i - \theta, 0\} \geq \max\{\theta_i - \theta', 0\} = 0$$

for any θ' such that $z(\theta' \mid \eta_i) > \hat{\theta}$ or $\theta' > \theta_i$.

Next we can check the participation constraint of S:

$$E[u_S(\theta, \eta_i) \mid \eta_i] = F(\theta_i \mid \eta_i)[\hat{\theta} - \theta_i + K] + [1 - F(\theta_i \mid \eta_i)]K = 0.$$

Finally, let us check WCP of the allocation. Define $\tilde{X}(z)$ as follows:

$$\tilde{X}(z) = z\tilde{q}(z) + \int_{z}^{\bar{z}} \tilde{q}(y)dy + K$$

where $\bar{z} \equiv \max\{z(\bar{\theta} \mid \eta_1), z(\bar{\theta} \mid \eta_2)\}$. It satisfies $\tilde{X}(z) - z\tilde{q}(z) \geq 0$ and $\tilde{X}(z) - z\tilde{q}(z) \geq \tilde{X}(z^{'}) - z\tilde{q}(z^{'})$ for any $z, z^{'}$ on possible range of z. Since

$$\tilde{X}(z(\theta \mid \eta_i)) = \hat{\theta} + K$$

for $\theta \leq \theta_i$, and

$$\tilde{X}(z(\theta \mid \eta_i)) = z(\theta \mid \eta_i)\tilde{q}(z(\theta \mid \eta_i)) + \int_{z(\theta|\eta_i)}^{\bar{z}} \tilde{q}(y)dy + K = K$$

for $\theta > \theta_i$, it implies that

$$\tilde{X}(z(\theta \mid \eta_i)) = u_A(\theta, \eta_i) + u_S(\theta, \eta_i) + \theta q(\theta, \eta_i).$$

The allocation satisfies all conditions of the allocation which is implementable in the weak collusion. In this allocation, the principal's payoff is

$$\begin{aligned} & \Sigma_\eta \Pr(\eta) E[V(\tilde{q}(z(\theta \mid \eta))) - h(\theta \mid \eta)\tilde{q}(z(\theta \mid \eta)) \mid \eta] \\ = \;& p(\eta_1)(V - \theta_1)F(\theta_1 \mid \eta_1) + p(\eta_2)(V - \theta_2)F(\theta_2 \mid \eta_2) \end{aligned}$$

which is strictly larger than the optimal payoff in NS: $(V - \theta^{NS})F(\theta^{NS})$ by Step 2, $\theta_1 \in (\theta_L, \theta^{NS})$ and $\theta_2 \in (\theta^{NS}, \theta_U)$. ■