Abstract

We study regulatory mechanism design with strong collusion between a privately informed agent and a less well-informed supervisor, incorporating ‘extortion’ which permits redistribution of rents within the coalition. We show that the Principal does not benefit from hiring the supervisor if the latter has less ex ante bargaining power vis-a-vis the agent. We provide sufficient conditions for the supervisor to be valuable if she has greater bargaining power. These results suggest the importance of anti-collusion strategies that augment bargaining power of supervisors vis-a-vis agents.

KEYWORDS: mechanism design, supervision, collusion, bargaining power
1 Introduction

The design of mechanisms to limit the harmful effects of collusion between supervisors and agents in adverse selection settings has been studied by many authors following Tirole (1986) and Laffont and Tirole (1993), with many applications to design of procurement, regulation and internal organization of firms. Subsequent literature has explored the implications of enlarging the severity of the collusion problem, such as soft information which provides wider scope for manipulation of reports (Faure-Grimaud, Laffont and Martimort (2003), Celik (2009)), and collusion over both reporting and participation decisions (Mookherjee, Motta and Tsumagari (2020)).

All these papers however assume that collusion is ‘weak’ in the sense that the supervisor and agent play non-cooperatively if either vetoes the offered side-contract. This ensures collusion occurs only to realize joint gains for the colluding parties. The allocation of bargaining power within the coalition then does not matter for the Principal’s capacity to control collusion. This implication of weak collusion follows from the Collusion-Proof Principle which applies quite generally in this class of models (Tirole (1992)).\footnote{For details of the argument, see Faure Grimaud et al. (2003) in a setting where the supervisor and agent collude only over reporting, or Mookherjee et al. (2020) where they collude both on participation and reporting decisions. The Principle implies that attention can be confined to revelation mechanisms which do not allow scope for any deviating (feasible) side contract for the coalition to increase the welfare-weighted sum of their payoffs. Given transferability of utility (owing to presence of side payments and absence of binding budget constraints for either colluding party), it follows that there cannot exist a deviating feasible side contract which makes both of them strictly better off. This property is independent of the relative welfare weights, i.e., if it holds for one set of weights it also holds for any other.} It is important to note that the irrelevance of bargaining power (among parties engaging in weak collusion) applies much more generally than conditions usually needed for the Coase Theorem to apply. For instance, bargaining takes place with asymmetric information between the supervisor and agent, either at the ex ante or interim stage (as defined by Holmstrom and Myerson (1983)). These models
are therefore incapable of explaining the observed success of anti-corruption policies preventing agents from selecting their own auditor in many countries such as India (Duflo et al. (2013)) and Italy (Vanutelli (2020)) which lowered bargaining power of the agent vis-a-vis the supervisor.\footnote{See Section 2 for more details of these policies and resulting outcomes.}

In this paper we argue that a strengthening of the nature of collusion to allow for ‘extortion’ can explain why relative bargaining power of the supervisor could matter. Extortion permits the colluding partner with greater bargaining power to extract rents from the other party, by threatening to send reports to the Principal which would hurt the latter if (s)he did not agree to the offered side-contract. Colluding parties can commit to such threats at the time of bargaining. This alters the sub-game following refusal of an offered side-contract by one party, from a simultaneous move noncooperative game, to one where the report of the accepting party is stipulated by the side-contract (with the other party choosing a best response). Following earlier work by Dequiedt (2007) and Che and Kim (2009) of an analogous notion in an auction design setting, we refer to this as ‘strong collusion’. While we do not endogenize the source of commitment power, we have in mind settings where S and A interact with one another in many other transactions with many principals, either at the same time or at later dates, and reputational concerns provide the required enforcement.\footnote{Consider a setting with many customers, each of whom wants to contract with an agent to commission a required production task (such as remodeling a kitchen), along with a supervisor who is better informed than customers about the cost of remodeling their respective kitchens. There are many ex ante identical potential agents and brokers/supervisors in the industry which constitutes a close-knit community, in which all members know one another, side-contracting between agents and supervisors happen, and outcomes of these side contracts are observed by the rest of the community. If they are far-sighted, any agent and supervisor appointed by a given customer have an incentive to develop and maintain a reputation for following through on promises and threats they make during side contracting.} It should be added that this is consistent with the rest of the mechanism design literature which examines the consequence of alternative assumptions.
regarding commitment among players, rather than providing an explicit microfounda-
tion for these assumptions. The existing literature which has studied weak collusion
has also been based on a particular assumption regarding lack of commitment among
colluding players. Here we are interested in understanding the consequence of the
opposite assumption where they can commit to reporting threats when they fail to
agree on a side-contract.

In a law review article, Ayres (1997) refers to ‘bribery’ and ‘extortion’ as the
‘twin faces of corruption’. In the language of mechanism design theory, bribery corre-
sponds to weak collusion, while the combination of bribery and extortion corresponds
to strong collusion. The importance of extortion has been stressed by numerous au-
thors in descriptive accounts of corruption, from medieval England (Cam (1930)),
to more contemporary accounts of corruption in Burma (Furnivall (1956)) or other
developing countries (Klitgaard (1988)). Ayres (1997) and Andrianova and Melissas
(2008) discuss extortion from a legal standpoint. Some papers studying tax evasion,
regulations or intra-firm organization in a moral hazard setting (Mookherjee (1997),
Hindricks et al. (1999), Khalil et al. (2010)) have shown that anti-corruption policy
design is significantly altered by the presence of extortion. By contrast our focus is
on an adverse selection context, where the modeling issues as well as detailed results
are quite different. Section 2 provides a more detailed discussion of related literature.

Section 3 describes the model of strong collusion when the good to be delivered
is perfectly divisible, and allocation of bargaining power between the supervisor (S)
and agent (A) is given and known by the Principal (P). In many contexts P may
have no control over the allocation of bargaining power, and takes this as given. In
others P may be able to influence bargaining power, e.g., if the game is embedded in
a setting where P can influence the process by which S and A are appointed which
affects their relative bargaining power. Consider for instance a setting where A is
allowed to appoint her own supervisor, from a set of potential supervisors. This
would give A the opportunity to exercise monopsony power over S in the choice
of side contract, resulting in a higher welfare weight on A’s payoff. An alternative institutional rule (as in the policy reforms in India and Italy studied by Duflo et al. (2013) and Vanutelli (2020)) is one where the supervisor is selected by P instead, and assigned to a given agent. This would alter the side contract negotiation to a bilateral monopoly where bargaining power is more equally divided between S and A. There may also be situations (e.g., in construction, or procurement of a particular service) where there are many potential agents available to carry out the project for P, in which P appoints a supervisor S and delegates the choice of A to S. This would confer monopoly power to S over A and tilt bargaining power in favor of the former.

Given a certain allocation of bargaining power within the coalition, Section 3 verifies that the Collusion Proofness Principle continues to apply in a strong collusion setting, once message spaces are augmented to include some non-type messages. This result is of some independent interest, as it contrasts with the moral hazard setting studied by Khalil et al. (2010). Other differences in our results from the moral hazard setting are described in Section 2. The result is used to characterize the class of feasible allocations in terms of a set of coalition incentive compatibility constraints.

This characterization is used in Section 4 which provides two main results for alternative ranges of bargaining power allocation. First, if A has a greater bargaining power than S,7 appointing a supervisor is worthless for P (Proposition 1). The intuition for this result is that when A has greater bargaining power, extortion allows A to extract all of S’s rents, since A knows everything that S knows. Hence A becomes the residual claimant on the coalition’s total surplus, and is able to push S down to her true outside option. This eliminates all bargaining frictions within the coalition, and the presence of S becomes redundant (by having no choice other than to ‘rubber-stamp’ whatever A wants to report).

Our next result (Proposition 2) shows that appointing a supervisor is valuable (for

---

7By this we mean A is assigned a higher welfare weight in the selection from the set of feasible side-contract payoff pairs.
generic information structures) when S has greater bargaining power, her signal has
two possible realizations and satisfies a monotone likelihood ratio property. In this
case, there are bargaining frictions within the coalition, since A is better informed
than S about her own cost. This prevents full rent extraction by S. Hence P is
able to control the costs of collusion to some extent, by compounding asymmetric
information frictions between S and A — a strategy of ‘throwing sands into the
wheels of corruption’.

The preceding results suggest but do not settle the question whether P’s welfare
is monotonically decreasing in A’s bargaining power over the range where A has
lower bargaining power than S. Section 5 studies a simpler setting where the good
to be procured by P is indivisible. Here we provide sufficient conditions for the
monotonicity property to hold. We also show that P’s welfare is always strictly lower
in strong collusion compared with weak collusion. Besides, this section provides
an illustration how our preceding results can be used to analyze the corresponding
optimal contracting problem with strong collusion.

The Appendix provides details of the proof of Proposition 2. The Supplementary
(online) Appendix includes proofs of required technical results such as the Collusion-
Proofness Principle, an associated characterization of strong collusion proof alloca-
tions, details of omitted steps in the proof of Propositions 2 and 3.8

2 Related Literature

In the mechanism design literature, implications of a similar notion of strong collusion
have been studied in the context of auction design featuring collusion among bidders,
collusion’ to refer to a different concept, which requires every Bayesian equilibrium of
the game induced by P’s contract to be weakly collusion proof. We are not aware of

8The Supplementary Appendix is located at http://people.bu.edu/dilipm/wkpap/index.html.
any previous study of the implications of strong collusion (representing a combination of bribery and extortion) in settings of supervision with a privately informed agent. In such a setting, our analysis shows how standard characterizations of feasible and optimal allocations in the existing literature on weak collusion (Faure-Grimaud et al. (2003), Celik (2009), Mookherjee et al. (2020)) need to be modified, besides providing results concerning the costs imposed by extortion and how they depend on intra-coalition bargaining power allocation. Most importantly, manipulating the bargaining power of the supervisor becomes an important tool of the Principal in curbing the costs of corruption, by increasing asymmetric information frictions within the supervisor-agent coalition.\(^9\)

The practical relevance of these results is highlighted by recent policy experience. Duflo et al. (2013) study a controlled experiment in India during 2009-10 where a treatment group of firms were no longer allowed to appoint their own pollution auditors, but had randomly assigned auditors instead. They found significant increases in pollution reports by the assigned auditors, and corresponding decline in actual pollution levels (verified from special backchecks conducted by the research team). Vanutelli (2020) studies a related policy reform in Italy introduced in 2011, where auditors of municipal budgets were randomly assigned instead of being appointed by local mayors. This resulted in increased property tax collections, larger budget surpluses and debt repayments. These effects were significantly larger in places with higher ‘risk of corruption’ (measured by prior investigations of corruption-related crimes).\(^10\)

\(^9\)This is broadly similar though different in details from the strategy proposed by Ortner and Chassang (2018) where P deliberately creates asymmetric information between agent and monitor by randomizing the latter’s incentive contract and not letting the agent observe the monitor’s contract with P. This particular tool is presumed unavailable in the settings we examine, e.g., the incentive contract for both parties are required to be in the public domain, as is commonly the case for public sector procurement or regulatory settings.

\(^10\)Using suitable extensions of our model, an earlier version of this paper illustrated how random assignment of a supervisor would raise bargaining power of the supervisor relative to a context where
The consequences of extortion for agent incentives in a supervision setting with moral hazard have been studied by a class of models (Mookherjee (1997), Hindricks et al. (1999), Khalil et al. (2010)). The most closely related paper in this group is Khalil et al. (2010), which also finds that extortion imposes a larger cost to the Principal’s welfare when the agent has larger bargaining power. However, many features and results of their model differ from ours. The information structure is different (the supervisor is either perfectly informed or perfectly uninformed; the supervisor’s information is hard unless the agent agrees to collude), and the Collusion-Proofness Principle does not hold in their setting. Most of their attention is subsequently devoted to the question whether bribery (i.e., resulting in mutual gain to both supervisor and agent) or extortion is the ‘greater evil’. This question is not meaningful in our setting owing to the Collusion-Proof Principle which applies in our setting and implies the optimality of eliminating both bribery and extortion. Moreover, in contrast to our results, hiring the supervisor is still valuable in their model even if the agent has all the bargaining power within the coalition (when the supervisor’s signal is informative enough). Extortion is costless in their setting if the supervisor has all the bargaining power; this may not be the case in our model (as verified in the indivisible good setting of Section 5).

3 Divisible Output Model

3.1 Technology, Preferences and Information

An appointed agent A delivers an output $q$ to the Principal P at a personal cost of $\theta q$. P’s return from $q$ is $V(q)$, a twice continuously differentiable, increasing and strictly concave function satisfying the Inada condition ($\lim_{q \to 0} V'(q) = +\infty$ and the agent appoints the supervisor. The idea is simple: if there are many competing potential agents and supervisors on each side of the market, giving a chosen agent or supervisor the power to appoint the other party would tilt relative bargaining power in its favor.
lim_{q \to +\infty} V'(q) = 0) and V(0) = 0. The realization of \( \theta \) is privately observed by A. \( \Theta \), which denotes the support of \( \theta \), constitutes an interval \([\tilde{\theta}, \bar{\theta}] \subset (0, \infty)\). It is common knowledge that everybody shares a common distribution function \( F(\theta) \) over \( \Theta \). It has a density function \( f(\theta) \) which is continuously differentiable and everywhere positive on \( \Theta \). Moreover \( H(\theta) \equiv \theta + F(\theta) f(\theta) \) is strictly increasing in \( \theta \).

An appointed supervisor S costlessly acquires an informative signal \( \eta \in \Pi \equiv \{\eta_1, \eta_2, \ldots, \eta_m\} \) about A’s cost \( \theta \) with \( m \geq 2 \). The realization of S’s signal is observed by A. \( a(\eta | \theta) \in [0, 1] \), which denotes the likelihood function of \( \eta \) conditional on \( \theta \), is continuously differentiable and positive-valued on \( \Theta \). We assume that for any \( \eta \in \Pi \), \( a(\eta | \theta) \) is not a constant function on \( \Theta \), and there are some subsets of \( \theta \) with positive measure satisfying \( a(\eta | \theta) \neq a(\eta' | \theta) \) for every \( \eta, \eta' \in \Pi \). In this sense each possible signal realization conveys information about the agent’s cost. The information conveyed is partial, since \( \Pi \) is finite. The cdf over \( \theta \) conditional on \( \eta \) is denoted \( F(\theta | \eta) \). Conditional on \( \eta \), the density function and distribution function are respectively denoted by \( f(\theta | \eta) \equiv f(\theta) a(\eta | \theta)/p(\eta) \) and \( F(\theta | \eta) \equiv \int_{\tilde{\theta}}^{\theta} f(\theta | \eta)d\theta \), where \( p(\eta) \equiv \int_{\tilde{\theta}}^{\bar{\theta}} f(\theta) a(\eta | \theta)d\theta \). Let \( K \equiv \Theta \times \Pi \) denote the set of possible states.

All players are risk neutral. S’s payoff is \( u_S = X_S + t_S \) where \( X_S \) is the transfer from P to S, and \( t_S \) is a transfer received by S within the coalition. A’s payoff is \( u_A = X_A + t_A - \theta q \) where \( X_A \) is the transfer from P to A and \( t_A \) is a transfer received by A within the coalition. P’s objective is a weighted average of profit \( (V(q) - X_A - X_S) \) and welfare of A and S \( (u_A + u_S) \), with a lower relative weight on the latter. With \( k \in (\frac{1}{2}, 1] \) which denotes the weight on profit, and \( 1 - k \) on welfare of A and S, P’s payoff reduces to \( k[V(q) - (X_S + X_A)] + (1 - k)[X_S + X_A - \theta q] \).\(^{13}\)

\(^{11}\)If S incurs a fixed cost \( c \) to acquire the signal, transfers received by S must be replaced by transfers net of this fixed cost while measuring S’s payoff. Increases in \( c \) will of course lower the value of appointing the supervisor, but it is easy to see how the results will be modified.

\(^{12}\)This assumes that the support of \( \theta \) given \( \eta \) is \( \Theta \) for all \( \eta \), in which sense \( \theta \) has full support. We adopt this assumption purely to simplify the exposition; our results extend to the case of non-full support (details available on request).

\(^{13}\)We exclude \( k = \frac{1}{2} \) because in that case the first-best can be achieved (as in Baron and Myerson
Hence the model applies both to the organization of private firms whose owners seek to maximize profit \((k = 1)\), as well as regulation or taxation contexts where \(P\) is a social planner pursuing a welfare objective that includes payoffs of \(A\) and \(S\) as well as \(P\)’s surplus, but assigns a higher weight to the latter.\(^{14}\)

In this economic environment, a (deterministic) allocation is denoted by

\[
\{(u_A(\theta, \eta), u_S(\theta, \eta), q(\theta, \eta)) \mid (\theta, \eta) \in K\}.
\]

### 3.2 Mechanism, Collusion Game and Equilibrium Concept

\(P\) designs a grand contract (GC) played by an appointed pair of \(A\) and \(S\), describing production decisions and transfers made by \(P\) in response to message sent by \(S\) and \(A\). We focus on deterministic mechanisms:

\[
GC = (X_A(m_A, m_S), X_S(m_A, m_S), q(m_A, m_S); M_A, M_S)
\]

where \(M_A\) (resp. \(M_S\)) denotes a message set for \(A\) (resp. \(S\)).\(^{15}\) In order to avoid technical complications, we assume that \(M_A\) and \(M_S\) are compact subsets of finite dimensional Euclidean spaces, and the domain of \((X_A, X_S, q)\) is also compact. These assumptions enable us to apply the minimax theorem (Nikaido (1954)) and guarantee the existence of an optimal side contract. Message spaces include exit options for \(A\) and \(S\) respectively \((e_A \in M_A, e_S \in M_S)\), where \(X_A = q = 0\) whenever \(m_A = e_A\), and \(X_S = 0\) whenever \(m_S = e_S\). The set of grand contracts satisfying these restrictions is denoted by \(\mathcal{GC}\). As a special case, \(P\) has the option to not hire \(S\), which we denote by \(No\ Supervision\ (NS)\), where \(M_S\) is null and \(X_S \equiv 0\).

Collusion between \(S\) and \(A\) takes the form of a side contract (SC) which is unobserved by \(P\). As explained in the Introduction, we treat the allocation of bargaining

\(^{14}\)The latter would be the case e.g., if \(P\) represents the interests of consumers, who need to be taxed to finance transfers to \(A\) and \(S\), and these taxes involve deadweight losses.

\(^{15}\)We ignore stochastic mechanisms which randomize the allocation conditional on messages, since they lower \(P\)’s welfare (owing to strict concavity of \(V\)) without affecting \(S\) or \(A\)’s payoffs.
power as a parameter, and represent it by relative welfare weights on the ex ante payoffs of S and A at the time that the side contract is chosen by the coalition in response to the contract GC offered by P. Formally, the side contract is selected at the ex ante stage by a fictional (uninformed) third party acting as a mediator, who maximizes $\alpha u_A + (1 - \alpha) u_S$ ($\alpha \in [0, 1]$), where $u_A$ and $u_S$ respectively denote ex ante payoffs of A and S, in response to choices of $\alpha$ and GC made by P. The third party does not play any budget breaking role, hence transfers within the coalition must balance: $t_A + t_S \leq 0$. No side payments can be exchanged at the ex ante stage; they can only be exchanged at the ex post stage after payments from P have been received. The side contract cannot be renegotiated at the interim or ex post stage, though either party can decide to withdraw from the agreement at these stages. The side contract allows A and S to exchange messages privately among one another, which determine a side payment and joint set of messages they respectively send to P. Since message spaces include exit as well as type reports, collusion takes the ex ante form studied in Mookherjee et al. (2020).

The stages of the game are as follows. Following the choice of GC by P, at stage 1 (the ex ante stage), the third party offers a side contract to S and A. A null side contract (NSC) could also be offered.

Next at stage 2 (the interim stage) S observes $\eta$ and A observes $(\theta, \eta)$. If a NSC was offered, they play the GC noncooperatively based on their prior beliefs, just as in a game without any collusion. If a non-null side contract was offered, S and A independently decide whether to accept it. Specifically, the game proceeds as follows. $i = A, S$ selects a message $d_i \in D_i$ ($i = A, S$) where $D_i$ is $i$’s message set specified in the side-contract. $D_i$ includes $i$’s exit option $\hat{e}_i$ from the side-contract. If $d_A \neq \hat{e}_A$ and $d_S \neq \hat{e}_S$, their reports to P are selected according to randomized message choices $\mu(d_A, d_S) \in \Delta(M)$ where $M \equiv M_A \times M_S$, and side payments to A and S are determined according to functions $t_A(d_A, d_S)$ and $t_S(d_A, d_S)$ respectively. If $d_A = \hat{e}_A$

$\Delta(M)$ denotes the set of probability measures on $M$. 

16
and $d_S = \hat{e}_S$, A and S play GC non-cooperatively.

What happens when one accepts and the other does not? Then SC specifies a reporting strategy of the party that accepted it, which can be interpreted as a threat that party commits to. The party that rejected it then plays a best response to this threat. Hence if $d_i \neq \hat{e}_i$ and $d_j = \hat{e}_j$ ($i, j = A, S$), $i$’s message to P is selected according to $\mu_i(d_i) \in \Delta(M_i)$, and the side payment to $i$ is $t_i(d_i).$\(^\text{17}\) On the other hand, $j$ plays GC without any constraint imposed by the side contract, and without any side transfer.\(^\text{18}\)

We focus on Perfect Bayesian Equilibrium (PBE) of this strong collusion game induced by the grand contract GC and bargaining weight parameter $\alpha.$\(^\text{19}\) However, there may be multiple PBE in a given game. We assume collusion permits parties to coordinate the choice of a PBE, hence the third party can specify a selected PBE to maximize the welfare-weighted sum of ex ante payoffs of S and A in the event of multiple PBE. The resulting equilibrium concept is denoted by PBE(sc). In case there are two PBE(sc) where the third party receives the same payoff, we assume that P can select the more desirable one.

Feasible allocations in strong collusion can now be defined:

**Definition 1** An allocation $(u_A, u_S, q)$ is achievable in strong collusion with $\alpha$ if it is realized in PBE(sc) under $\alpha$ for some $GC \in GC$.

### 3.3 Strong Collusion-Proof Allocations

We now explain the notion of a strong-collusion-proof allocation. To aid the exposition and highlight the contrast with the corresponding notion of weak-collusion-proof allocations which is familiar from past literature, we start with an informal discussion.

\(^\text{17}\)Owing to the budget balance condition, $t_i(d_i) \leq 0.$

\(^\text{18}\)When collusion is weak instead, the side contract ceases to apply for the subsequent messages for either player when one of them exits — S and A play GC noncooperatively.

\(^\text{19}\)For definition of PBE, see Fudenberg and Tirole (1991).
of the main difference between the two contexts. As explained in the Introduction, in a strong collusion setting side contracts allow the player with a greater bargaining power to commit to reporting threats that induce ‘extortion’ of rents of the other player. The player with weaker bargaining power could be pushed below the payoff she would earn in the absence of any collusion. The extent to which these redistributions can occur depends on the reporting options available. P may then seek to provide the weaker player with additional reporting options that provide it with protection against extortion. For instance, if A is more powerful than S within the coalition, P might want to protect S against extortion by A by providing an ‘escape route’ or auxiliary message option. If chosen it will guarantee S a minimum level of payoff, thereby limiting the extent to which A can extort from S. On the other hand, this option may enlarge the scope for collusion between S and A, or the extent to which S can extort from A in the reverse situation where S is more powerful. Hence P needs to weigh these advantages and disadvantages while augmenting reporting options.

It turns out that it suffices for P to restrict attention to mechanisms which augments the set of possible \( \eta \) reports of S’s signal with a single auxiliary option, denoted \( \eta_0 \), which provides S with a constant payoff (denoted \( \omega \)) irrespective of whatever \( \theta \) happens to be reported. \( \omega \) equals the minmax payoff of S, representing the worst punishment that A can impose on S. \( \omega \) must obviously be non-negative since S always has the option to not participate in the mechanism. The option to report this auxiliary message serves to protect S against extortion by A. Conversely, A is protected from extortion by a cross-checking rule where inconsistent reports of \( \eta \) result in punishments imposed only on S.

Analogous to the weak collusion context, a version of the Collusion Proof Principle can then be shown to hold — given any arbitrary mechanism chosen by P and any allocation resulting from a PBE(sc) in that mechanism, there exists an augmented revelation mechanism of the kind explained above (with a single auxiliary message
\(\eta_0\), which is ‘strong collusion proof’. The latter criterion requires that there does not exist a deviating side contract which raises a welfare weighted sum of payoffs of S and A subject to the following feasibility constraints: (a) a truthful reporting constraint for A within the coalition, and (b) a set of participation constraints for S and A with respective outside option payoffs that correspond to the ‘worst’ punishment that can be imposed by the other party if the player in question refuses to participate in the deviating side contract.

These coalitional participation constraints differ from those in weak collusion, as we now explain. It is helpful to recall the nature of these constraints in a weak collusion setting (see e.g., Mookherjee et al. (2020)). In such a context, there is no need to augment the message spaces, so we can focus on a standard revelation mechanism. Let the revelation mechanism result in output \(q(\theta, \eta)\), payoffs \(u_A(\theta, \eta), u_S(\theta, \eta)\) for S and A and corresponding aggregate payment \(X(\theta, \eta) \equiv u_A(\theta, \eta) + u_S(\theta, \eta) + \theta q(\theta, \eta)\) when \((\theta, \eta) \in K\) is reported and there is no inconsistency between S and A’s reports, and \(X = q = 0\) if \(e = (e_A, e_S)\) is reported. Recall also the nature of individual incentive compatibility constraints arising in the absence of any collusion: pertaining to

\((IC_A)\): truthful reporting of \(\theta\) by A:

\[
u_A(\theta, \eta) \geq u_A(\theta', \eta) + (\theta' - \theta)q(\theta', \eta) \quad (1)
\]

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

\((PC_A)\): interim participation constraint for A

\[
u_A(\theta, \eta) \geq 0 \quad (2)
\]

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

\((PC_S)\): interim participation constraint for S

\[
E[u_S(\theta, \eta) | \eta] \geq 0 \quad (3)
\]
for any $\eta \in \Pi$. We say that $(u_A, u_S, q)$ satisfies individual incentive compatibility (IIC) if and only if it satisfies all three conditions $IC_A$, $PC_A$ and $PC_S$.

Now consider coalitional incentive constraints under weak collusion. Since $S$ and $A$ can collude over participation decisions as well as reports, they could deviate to a side contract where they jointly report according to strategy $\mu(\theta, \eta) \in \Delta(K \cup \{e\})$ and enter into side-payments resulting in payoffs $\tilde{u}_A(\theta, \eta), \tilde{u}_S(\theta, \eta)$ instead. The feasibility requirements on such a deviating side contract are the following.

First, we have the requirement of budget balance within the coalition:

$$\tilde{u}_A(\theta, \eta) + \tilde{u}_S(\theta, \eta) \leq X(\mu(\theta, \eta)) - \theta q(\mu(\theta, \eta)) \tag{4}$$

Second, as $A$ is privately informed about the realization of $\theta$, it must satisfy individual incentive constraints requiring $A$ to report $\theta$ truthfully within the coalition:

$$\tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta', \eta) + (\theta' - \theta) q(\mu(\theta', \eta)) \text{ for any } \theta' \in \Theta \tag{5}$$

If the deviating side contract offer is rejected by either $A$ or $S$, they revert to noncooperative play of the revelation mechanism in which they report truthfully and end up with payoffs $u_A(\theta, \eta), u_S(\theta, \eta)$. These serve as their respective outside options. Since either party can decide to opt out of the side contract at the interim stage, the relevant participation constraints under weak collusion are:

$$E[\tilde{u}_S(\cdot, \eta) | \eta] \geq E[u_S(\cdot, \eta) | \eta] \tag{6}$$

$$\tilde{u}_A(\theta, \eta) \geq u_A(\theta, \eta). \tag{7}$$

With strong collusion, these constraints need to be modified as follows. First, the message space is augmented, and the definition of the mechanism has to be extended to this augmented message space. Define $\bar{K} \equiv \Theta \times \bar{\Pi}$ where $\bar{\Pi} \equiv \Pi \cup \{\eta_0\}$. The augmented message space is then $\bar{K} \cup \{e\}$. The allocation --- denoted by $(\bar{u}_A^e, u_S^e, q^e)$ --- is now defined over $\bar{K} \cup \{e\}$. Let $X^e(\theta, \eta) \equiv u_A^e(\theta, \eta) + u_S^e(\theta, \eta) + \theta q^e(\theta, \eta)$ denote the
corresponding aggregate transfer from \( P \) to the coalition. Moreover, \((X^e(e), q^e(e)) = (0, 0)\). Say that \((u_A^e, u_S^e, q^e)\) satisfies \(IC_A, PC_A, PC_S\) when this function satisfies \((1, 2, 3)\) respectively over the entire domain \( \bar{K} \cup \{e\} \).

In any deviating side contract, the reporting strategy is now denoted \( \mu(\theta, \eta) \in \Delta(\bar{K} \cup \{e\}) \), payoffs are \( \tilde{u}_A(\theta, \eta), \tilde{u}_S(\theta, \eta) \). The budget balance and individual incentive constraints \((4, 5)\) are then extended in a similar manner:

\[
\tilde{u}_A(\theta, \eta) + \tilde{u}_S(\theta, \eta) \leq X^e(\mu(\theta, \eta)) - \theta q^e(\mu(\theta, \eta)) \quad (8)
\]

\[
\tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta', \eta) + (\theta' - \theta) q^e(\mu(\theta', \eta)) \quad \text{for any} \quad \theta' \in \Theta. \quad (9)
\]

The main difference from weak collusion is in the nature of the participation constraints. In strong collusion, if the offer is rejected by one of the two parties, the party who accepted the contract commits to a reporting strategy designed to punish the rejecting party. Suppose that \( A \) was the one to reject the side contract in some state of the world \( \theta, \eta \), while \( S \) accepted it. Then \( S \) could threaten to send a false report \( \eta' \) about the realization of her own signal. Let this threat be denoted by the mixed strategy \( P(\eta'|\eta) \). \( A \) will play a best response to this. Since the mechanism satisfies individual incentive constraints, it will be optimal for \( A \) to report \( \theta \) truthfully, no matter what \( S \) reports. Hence \( A \) will end up with a payoff of \( u_A^e(\theta, \eta') \) if \( S \) happens to report \( \eta' \). Given \( S \)'s threat to report according to \( P(\eta'|\eta) \), the outside option payoff for \( A \) at the interim stage will therefore be \( \sum_{\eta' \in \bar{\Pi}} P(\eta' | \eta) u_A^e(\theta, \eta') \). The side contract participation constraint for \( A \) reduces to:

\[
\tilde{u}_A(\theta, \eta) \geq \sum_{\eta' \in \bar{\Pi}} P(\eta' | \eta) u_A^e(\theta, \eta'). \quad (10)
\]

This is milder than the corresponding participation constraint \((7)\) under weak collusion, since it reduces to the latter only for truthful reporting strategies where \( P(\eta|\eta) = 1 \) for all \( \eta \in \Pi \). By allowing the side contract to stipulate threats of \( S \) reporting her signal untruthfully, the agent’s participation constraint can be relaxed by the side contract designer. This reduces \( P \)'s capacity to manipulate the outcome of bargaining over the side contract.
Next consider the situation where S may reject the side contract, while A accepts it. Then A can threaten to punish S by misreporting in a way that lowers S’s payoff. The worst punishment that can be imposed on S is represented by S’s minmax payoff \( \omega \).\(^{20}\) Hence S’s participation constraint can be written as

\[
E[\tilde{u}_S(\cdot, \eta) | \eta] \geq \omega
\]

instead of (6) in weak collusion. By construction \( \omega \leq E[u_S(\cdot, \eta) | \eta] \), so S’s participation constraint is also weaker under strong collusion.

We can now provide a formal definition of strong collusion proof allocations. In this we use the following definitions: the strategy of reporting truthfully in state \( \eta \) (i.e. \( P(\eta | \eta) = 1 \) and \( P(\eta' | \eta) = 0 \) for any \( \eta' \neq \eta \)) is denoted by \( I(\eta) \). Similarly, let \( I(\theta, \eta) \) denote the strategy of reporting the state truthfully in state \((\theta, \eta)\).

**Definition 2**

Allocation \((u_A, u_S, q)\) is strong collusion-proof (or SCP) for \( \alpha \in [0, 1] \), if \((u_A, u_S, q)\) is IIC, and there exists a payoff \( \omega \geq 0 \) for S, and an augmentation \((u^e_A, u^e_S, q^e)\) of \((u_A, u_S, q)\) on \( \bar{K} \) with \( u^e_S(\theta, \eta_0) = \omega \) for any \( \theta \in \Theta \) and \((u^e_A(\theta, \eta_0), q^e(\theta, \eta_0))\) satisfying \((IC_A)\) and \((PC_A)\), such that for any \( \eta \in \Pi \),

\[
(\mu(\theta, \eta), \tilde{u}_A(\theta, \eta), \tilde{u}_S(\theta, \eta), P(\cdot | \eta)) = (I(\theta, \eta), u_A(\theta, \eta), u_S(\theta, \eta), I(\eta))
\]

solves problem \( PS(\alpha : \eta) \):

\[
\max E[\alpha \tilde{u}_A(\theta, \eta) + (1 - \alpha)\tilde{u}_S(\theta, \eta) | \eta]
\]

subject to \((\mu(\theta, \eta), \tilde{u}_A(\theta, \eta), \tilde{u}_S(\theta, \eta), P(\cdot | \eta))\) satisfies for all \( \theta \in \Theta \):

(i) \( \mu(\theta, \eta) \in \Delta(\bar{K} \cup \{e\}), \tilde{u}_A(\theta, \eta) \in \Re, \tilde{u}_S(\theta, \eta) \in \Re, P(\cdot | \eta) \in \Delta(\bar{\Pi}) \)

(ii) \( \tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta', \eta) + (\theta' - \theta)q^e(\mu(\theta', \eta)) \) for any \( \theta' \in \Theta \)

(iii) \( \tilde{u}_A(\theta, \eta) + \tilde{u}_S(\theta, \eta) \leq X^e(\mu(\theta, \eta)) - \theta q^e(\mu(\theta, \eta)) \)

\(^{20}\)Owing to the Minimax Theorem which applies to this setting, S’s minmax payoff is well defined.
We now present the main result of this section: that the Collusion Proofness Principle extends to the context of strong collusion.

Lemma 1 An allocation \((u_A, u_S, q)\) is achievable in strong collusion with \(\alpha\) if and only if it is strongly collusion-proof for \(\alpha\).

The proof is provided in the Supplementary Appendix. It shows that every achievable allocation resulting from some PBE(sc) in any given GC can also be achieved as the outcome of a revelation mechanism augmented to incorporate the additional message option \(\eta_0\) as represented in the above definition of strong collusion proofness (i.e. where the coalition has no incentive to deviate from truthful reporting with a non-null side contract). Here is a brief outline of the argument. Starting with the original mechanism, observe first that S’s minmax payoff \(\omega\) is well defined as a consequence of the Minmax Theorem. This is because S’s payoff function is independent of the true state and is common knowledge between S and A. Now construct in the usual manner the direct revelation mechanism that (i) corresponds to (i.e, is the composition of) equilibrium reporting strategies and associated equilibrium side contract in state \((\theta, \eta)\), i.e., assigns this outcome when A reports \(\theta\) and both S and A report \(\eta\), and (ii) when their reports are inconsistent: \(\eta_A \neq \eta_S\), select the outcome resulting in state \((\theta, \eta_S)\) and in addition impose a large financial penalty only on S. This strategy of punishing only S for inconsistent reports provides A some protection against threats by S to select a mixed reporting strategy if A refuses the side contract. A’s outside option in side contracting is then exactly analogous to constraint (v) in the preceding definition of SCP.

On the other hand, S is then exposed to the reverse threat by A of selecting a mixed reporting strategy if S should refuse the side contract. (If A were to commit
to a pure strategy report $\eta_A$, S could protect himself from being levied the extra penalty for inconsistent reports by also reporting $\eta_S = \eta_A$. But if A commits instead to a mixed strategy over $\eta_A$, S would not be able to avoid these penalties.) Moreover even if A were to restrict herself to a pure reporting strategy when S refuses the side contract, S could still be vulnerable to possible misreporting of $\theta$ by A since the assigned allocation is based on this report. In particular, it could enable A to lower S’s maxmin payoff below $\omega$, what it was in the original game. It is to prevent this that an auxiliary message option $\eta_0$ is provided by P.

Specifically, the mechanism is extended as follows: $u^S_S(\theta, \eta_0) = \omega$ for all $\theta$, while $(u^e(\theta, \eta_0), q^e(\theta, \eta_0))$ is set equal to the corresponding outcomes in the original game when S plays his maxmin strategy and type $\theta$ selects a best response. It thereby does not expand the scope for profitable collusion beyond what was already available in the original mechanism. Moreover, the scope for either party to extort from the other has not been enlarged. The rest of the argument is then straightforward: since the original allocation already incorporated the outcome of optimal side contracting by the coalition, they will not have any incentive to deviate now in the augmented revelation mechanism to a non-null side contract.

4 Divisible Output: Results

One class of allocations that can always be attained by P irrespective of collusion corresponds to not utilizing reports regarding the supervisor’s signal $\eta$ at all. This is the No Supervision (NS) organization, in which the class of attainable allocations is as follows. There exists a nonnegative constant $c$ and non-increasing real-valued functions $(X_A(\theta), Q(\theta))$ defined on $\Theta$ such that for any $(\theta, \eta)$:

(a) $u_S(\theta, \eta) = c$

(b) $u_A(\theta, \eta) = X_A(\theta) - \theta Q(\theta) = \max_{\theta' \in \Theta} [X_A(\theta') - \theta Q(\theta')]$. 

19
We can interpret this as a mechanism where $S$ is not asked to submit any report, and is paid a fixed amount $c$. $u_A(\theta, \eta)$ is the payoff attained by type $\theta, \eta$ of $A$ in the optimal contract designed by $P$ when she contracts (only) with $A$ on the basis of her prior information. In the latter setting, asking $A$ to report $\eta$ is worthless — in the absence of any scope for cross-checking with a $\eta$ report from $S$, $A$ would submit whichever $\eta$ report would maximize her ex post payoff. Hence NS is not vulnerable to weak or strong collusion: any feasible allocation in NS is also feasible with weak or strong collusion (irrespective of $\alpha$).

We now present our first main result.

**Proposition 1** An allocation which is strong collusion proof for any $\alpha \geq \frac{1}{2}$ is also attainable in NS.

**Proof of Proposition 1:** Consider any allocation $(u_A(\theta, \eta), u_S(\theta, \eta), q(\theta, \eta))$ which is strong collusion proof for $\alpha \geq \frac{1}{2}$. By Lemma 1, there exists $\omega \geq 0$ and an incentive compatible augmentation $(u^e_A, u^e_S, q^e)$ of this allocation satisfying $u_S(\theta, \eta_0) = \omega$, such that for any $\eta$, $(I(\theta, \eta), u_A(\theta, \eta), u_S(\theta, \eta), I(\eta))$ solves $P^S(\alpha : \eta)$. Let the corresponding coalitional incentive scheme be $(X^e(\mu), q^e(\mu))$. Define

$$\mu^*(\theta) \in \arg \max_{\mu \in \Delta(K \cup \{e\})} [X^e(\mu) - \theta q^e(\mu)]$$

i.e., a reporting strategy that maximizes the ex post joint payoff of $A$ and $S$ in every state.

We claim that

$$(\mu(\theta, \eta), u_A(\theta, \eta), u_S(\theta, \eta)) = (\mu^*(\theta), X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) - \omega, \omega)$$

is a solution of $P^S(\alpha : \eta)$ for any $\eta$. Upon setting $\omega = c$, $X_A(\theta) = X^e(\mu^*(\theta)) - \omega$ and $Q(\theta) = q^e(\mu^*(\theta))$, this claim will imply that the allocation is attainable in NS.

To establish the claim, we first derive an upper bound for the objective function in the problem $P^S(\alpha : \eta)$. From the constraint $E[\tilde{u}_S(\theta, \eta) | \eta] \geq \omega$ and the assumption
that $\alpha \geq \frac{1}{2}$, for any reporting strategy $\mu(\theta, \eta)$ the following is true:

$$E[\alpha \tilde{u}_A(\theta, \eta) + (1 - \alpha)\tilde{u}_S(\theta, \eta) | \eta]$$

$$\leq E[\alpha\{X^e(\mu(\theta, \eta)) - \theta q^e(\mu(\theta, \eta))\} + (1 - 2\alpha)\tilde{u}_S(\theta, \eta) | \eta]$$

$$\leq \alpha E[X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) | \eta] + (1 - 2\alpha)\omega.$$  

This upper bound can be attained in $P^S(\alpha : \eta)$ by choosing $\mu(\theta, \eta) = \mu^*(\theta)$,

$$\tilde{u}_A(\theta, \eta) = X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) - \omega$$

and $\tilde{u}_S(\theta, \eta) = \omega$ for any $\theta \in \Theta$, and $P(\eta_0 | \eta) = 1$ and $P(\eta' | \eta) = 0$ for any $\eta' \neq \eta_0$. This allocation satisfies A’s participation constraint (v), since

$$\tilde{u}_A(\theta, \eta) = X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) - \omega$$

$$\geq X^e(\theta, \eta_0) - \theta q^e(\theta, \eta_0) - u^e_S(\theta, \eta_0) = u^e_A(\theta, \eta_0).$$

As the other constraints are obviously satisfied, the claim is established. 

When A has at least as much bargaining power as S, it is optimal for the coalition to pin S down to her (constant) minmax payoff and provide all residual rents to A. Reports by the coalition are then chosen to maximize A’s payoffs (i.e, S’s role is restricted to rubber-stamping whatever report is in A’s interest). Hence P cannot derive any benefit from appointing S. The one-sided asymmetric information within the coalition implies absence of any frictions in collusion when the informed party A has more bargaining power than S. For P to derive some value from appointing S, she has to exploit some frictions in coalitional bargaining.

Now consider the case where S has higher bargaining power than A. We impose some structure on A’s type space and S’s information. One condition is that S should not be ‘too well informed’ about A’s cost; for instance in the extreme case where S is perfectly informed about $\theta$, there will again be no frictions in coalitional bargaining and appointing S will not yield any value to P. For the rest of this section, we focus on the following case:
Context C: S’s cost signal has two possible signal realizations \( \eta_1, \eta_2 \) satisfying a Monotone Likelihood Ratio Property (MLRP) such that \( a(\eta_1 \mid \theta) \) is decreasing (while \( a(\eta_2 \mid \theta) \) is increasing) in \( \theta \).

Our main result is that in Context C, P can derive positive value from appointing S if S has greater bargaining weight than A, for a generic set of information structures. Given the previous result, this implies that (generically) P is better off when S has strictly higher bargaining weight than A, compared to when this is not true.

**Proposition 2** Consider Context C and assume \( \alpha \in [0, 1/2) \). If there do not exist \((\rho, \nu, \gamma) \in \mathbb{R}^3 \) such that \( a(\eta_1 \mid \theta) = \rho + \nu F(\theta)\gamma \) for all \( \theta \in \Theta \), P can attain a strictly higher expected payoff by appointing S, compared to not appointing S.

As the proof is relegated to the Appendix, we outline the main steps in the argument here.

First we show that S’s participation constraint can be dropped in problem \( P^S(\alpha : \eta) \) when \( \alpha \in [0, 1/2) \).

**Lemma 2** S’s participation condition \( E[\tilde{u}_S(\theta, \eta) \mid \eta] \geq \omega \) can be dropped in \( P^S(\alpha : \eta) \) without affecting the solution, if \( \alpha \in [0, 1/2) \).

If the lemma were false, the solution to the relaxed version of problem \( P^S(\alpha : \eta) \) when S’s participation constraint is dropped, must violate this constraint, implying that S ends up with an expected payoff below his minmax payoff \( \omega \). The coalition has the option of switching to the ‘A-residual-claimant’ (ARC) side-contract (used in the proof of Proposition 1) in which S receives a constant payoff of \( \omega \) and A receives the rest of the aggregate coalitional rent. ARC induces ex post efficient reporting strategies, thereby (weakly) expanding the aggregate rent in every state. Given \( \alpha < \frac{1}{2} \), A must also benefit from deviating to ARC. But this leads to a contradiction, as S and A both benefit from deviating to ARC.
Now consider problem $P^S(\alpha : \eta)$ in which S’s participation constraint is dropped. P augments the mechanism in the manner described in Definition 2, where the auxiliary message $\eta_0$ is identified with the high-cost signal report $\eta_2$ (i.e., results in the same outcomes). Hence we can confine attention to two possible signal reports $\eta_1, \eta_2$ for A and S. If both report $\eta_2$, P selects the optimal allocation $(u_A^{NS}(\theta), u_S^{NS}(\theta), q^{NS}(\theta))$ in NS:

$$u_A^{NS}(\theta) = \int_{\bar{\theta}}^{\theta} \bar{q}(y) dy, u_S^{NS}(\theta) = 0$$

$$q^{NS}(\theta) = \bar{q}(\theta) \equiv \arg \max_q [V(q) - H(\theta; k) q],$$

where $H(\theta; k) \equiv \theta + \frac{2k-1}{k} \frac{F(\theta)}{f(\theta)}$ and $k$ denotes the weight assigned to P’s profit. Then $P$ attains the same payoff as in NS: $W^{NS} \equiv E[V(\bar{q}(\theta)) - H(\theta; k) \bar{q}(\theta)].$

When both S and A report the low-cost signal $\eta_1$, let P select the following variation on the optimal allocation in NS. Let $\beta \equiv \frac{1-2\alpha}{1-\alpha}$, which lies in the interval $(0, 1)$. Let $\Lambda(\cdot) : \Theta \rightarrow \mathbb{R}$ such that (i) $\Lambda(\cdot)$ is non-decreasing in $\theta$ with $\Lambda(\theta) = 0$ and $\Lambda(\bar{\theta}) = 1$, and (ii) the function $z_{\beta}(\theta)$ defined by

$$z_{\beta}(\theta) = \theta + \beta \frac{F(\theta \mid \eta_1) - \Lambda(\theta)}{f(\theta \mid \eta_1)}$$

is nondecreasing. The choice of $\Lambda(\cdot)$ is akin to choosing a set of outside option payoffs for A when A and S bargain over a side-contract. As this function takes non-negative values, A’s outside option is raised. This lowers the virtual cost to the coalition of asking A to produce one more unit of the good, from $\theta + \beta \frac{F(\theta \mid \eta_1)}{f(\theta \mid \eta_1)}$ to $z_{\beta}(\theta)$. Hence ceteris paribus the coalition is induced to deliver a larger quantity of output. This adjustment enables P to control the ‘double marginalization of rent’ problem (wherein the coalition under-produces the good owing to S’s desire to limit A’s information rents).

Below we explain in further detail the exact manner in which P can select $\Lambda(\cdot)$. Given this function and thereby the coalitional virtual cost function $z_{\beta}(\cdot)$, the corresponding output schedule is set at $q(\theta, \eta_1) = \bar{q}(z_{\beta}(\theta))$. $z_{\beta}(\theta)$ exceeds or falls below $\theta$ according as $\Lambda(\theta)$ is smaller or larger than $F(\theta \mid \eta_1)$, implying in turn that $q(\theta, \eta_1)$ is
smaller or larger than \( q^{NS}(\theta) \). So (conditional on \( \eta = \eta_1 \)) the contracted output with A will be expanded over some ranges of reported cost \( \theta \), and shrunk for other values of \( \theta \). Specifically, the output and payoffs are altered as follows:

\[
q(\theta, \eta_1) = \bar{q}(z_\beta(\theta)) \tag{13}
\]

\[
u_A(\theta, \eta_1) = \int_\theta^\bar{\theta} \bar{q}(z_\beta(y))dy \tag{14}
\]

\[
u_S(\theta, \eta_1) = \bar{X}(z_\beta(\theta)) - \theta \bar{q}(z_\beta(\theta)) - \int_\theta^{\bar{\theta}} \bar{q}(z_\beta(y))dy \tag{15}
\]

where

\[
\bar{X}(z) \equiv z\bar{q}(z) + \int_z^{\bar{\theta}} \bar{q}(y)dy. \tag{16}
\]

This can be interpreted as follows. P behaves ‘as if’ she is contracting with a single composite agent (representing the coalition) with a unit cost \( z \) of delivering the good. The coalition submits a report of \((\theta, \eta_1)\) which then determines a report of \( z \) given by (12) and an output order of \( q(\theta, \eta_1) = \bar{q}(z_\beta(\theta)) \). The corresponding total payment from P to the coalition is given by (16), so as to induce the coalition to report \( \theta \) and hence \( z \) truthfully. The contract also specifies the division of this aggregate payment between A and S as per (14) and (15) to insure that A’s individual incentive constraint to report \( \theta \) is satisfied, with the rest going to S.

Finally, when S and A submit different reports \( \eta^S \neq \eta^A \), A is offered the same allocation as in the case where the submitted \( \eta \) reports are \( \eta^S \) for both S and A, while S receives a payment equal to what he would have received if their \( \eta \) reports had been \( \eta^A \) for both S and A, minus a large penalty. This will ensure that the side contract will always involve submission of a common report by S and A.

The aim is to construct \( \Lambda(\cdot) \) with the properties stated above, such that the resulting allocation is SCP and improves \( P \)'s payoff in state \( \eta_1 \) relative to the allocation resulting when S is not hired:

\[
E[V(\bar{q}(z_\beta(\theta)))] - \frac{2k-1}{k}X(z_\beta(\theta)) - \frac{1-k}{k} \theta \bar{q}(z_\beta(\theta)) | \eta_1] \]

\[
> E[V(q^{NS}(\theta)) - \frac{2k-1}{k} \bar{X}(\theta) - \frac{1-k}{k} \theta q^{NS}(\theta) | \eta]. \tag{17}
\]
Since the allocation is unchanged in state $\eta_2$, $P$ will achieve a higher payoff than in NS.

The proof shows that such a variation is indeed SCP provided the following two conditions are satisfied:

(a) $E[u_A(\theta, \eta_1) - u_A(\theta, \eta_2) \mid \eta_2] \geq 0$

(b) $\int_\theta^\theta [u_A(\theta, \eta_2) - u_A(\theta, \eta_1)] d\Lambda(\theta) \geq 0$.

These two conditions are shown to jointly imply that the coalition does not benefit from manipulating $A$’s outside option by $S$ threatening to report $\eta_j$, $j \neq i$ different from the true signal $\eta_i$, if $A$ were to refuse the offered side contract.

Conditions (a) and (b) can be rewritten as follows:

$$E\left[ \{\bar{q}(\beta(\theta)) - q(\theta)\} \frac{F(\theta \mid \eta_2)}{f(\theta \mid \eta_2)} \mid \eta_2 \right] \geq 0$$

and

$$E[(\beta(\theta) - \bar{h}(\theta \mid \eta_1))(\bar{q}(\beta(\theta)) - q(\theta)) \mid \eta_1] \geq 0$$

where $h_{\beta}(\theta \mid \eta) = \theta + \beta \frac{F(\theta \mid \eta)}{f(\theta \mid \eta)}$. \(^{21}\)

Consider a small variation of the $\beta(\theta)$ function around the identity map $\theta'(\theta) = \theta$.

The corresponding point-wise variations in the left-hand-sides of (17)-(19) are as follows\(^{22}\)

$$[V'(\bar{q}(z))\bar{q}'(z) - \frac{2k - 1}{k} \bar{X}'(z) - \frac{1 - k}{k} \theta \bar{q}'(z)]_{z=\theta} f(\theta \mid \eta_1) = \bar{q}'(\theta) \frac{2k - 1}{k} \frac{F(\theta)}{f(\theta)} f(\theta \mid \eta_1),$$

and

$$[\bar{q}'(z) \frac{F(\theta \mid \eta_2)}{f(\theta \mid \eta_2)}]_{z=\theta} f(\theta \mid \eta_2) = \bar{q}'(\theta) F(\theta \mid \eta_2),$$

and

$$[(\bar{q}(z) - \bar{q}(\theta)) + (z - h_{\beta}(\theta \mid \eta_1))\bar{q}'(z)]_{z=\theta} f(\theta \mid \eta_1) = -\beta \bar{q}'(\theta) F(\theta \mid \eta_1).$$

\(^{21}\)We use $E[\int_\theta^\theta \bar{q}(y) dy \mid \eta] = E[\frac{F(\theta \mid \eta)}{f(\theta \mid \eta)} \bar{q}(\theta) \mid \eta]$ to derive these equations.

\(^{22}\)We use $V'(\bar{q}(z)) = H(\theta; k)$ to obtain (20).
A sufficient condition for a variation which locally preserves the value of the left-hand-sides of (18) and (19), while increasing the value of the left-hand-side of (17), is that \( F(\theta) f(\theta) \) does not lie in the space spanned linearly by \( F(\theta \mid \eta_2) \) and \( -F(\theta \mid \eta_1) \). This is ensured by the generic property stated in Proposition 2. So the variation ends up expanding the output procured over some ranges and contracting it over others (compared to the NS allocation) when the low-cost signal \( \eta_1 \) is reported. As is well known, the optimal allocation in NS involves ‘under-procurement’ owing to standard adverse selection distortions: hence expanding (resp. contracting) output procured increases (resp, decreases) P’s ex post payoff. The ranges over which expansion and contraction respectively take place can be chosen to ensure that P’s ex ante payoff increases (the difference being proportional to the ex ante expected value of (17) which will be positive), while ensuring that SCP conditions are preserved, i.e., (18) and (19) are preserved.

5 Procuring an Indivisible Good

The analysis of the previous Section left open the question how P’s optimal payoff varies with \( \alpha \) the bargaining power of A over the range where it is smaller than \( \frac{1}{2} \). Moreover it made no effort to analyze the optimal contracting problem with strong collusion. In this Section we address both these concerns in a specific context, where the good being procured is indivisible. In this context we also show that P is strictly worse off compared with weak collusion.

We maintain the same assumptions as in the divisible good model, with the following exceptions. The quantity to be procured is \( q \in \{0, 1\} \). P obtains a zero gross benefit if \( q = 0 \), and a benefit of \( V > 0 \) otherwise. There are two possible signal realizations \( \eta_i, i = L, H \), and we are in Context C so \( a_i(\theta) \equiv a(\eta_i \mid \theta) > 0 \) on \( \Theta \) for \( i \in \{L, H\} \), and \( a_H(\theta) \) is increasing in \( \theta \). \( F_i(\theta) \equiv F(\theta \mid \eta_i) \) now denotes the distribution of \( \theta \) conditional on signal realization \( i \), which has a positive density \( f_i(\theta) \equiv f(\theta \mid \eta_i) \).
\( \kappa_i \equiv p(\eta_i) = \int_\theta^\bar{\theta} f(\theta)a_i(\theta) d\theta \in (0,1) \) denotes the probability of signal \( i \). Our assumptions imply \( F_L(\theta) > F(\theta) > F_H(\theta) \) and \( F_L(\theta)/f_L(\theta) > F(\theta)/f(\theta) > F_H(\theta)/f_H(\theta) \) on \((\theta, \bar{\theta})\). In addition, we assume that (i) \( h_i(\theta) = \theta + \frac{F_i(\theta)}{F_i(\theta)} \) and \( l_i(\theta) = \theta + \frac{F_i(\theta)-1}{f_i(\theta)} \) are increasing in \( \theta \) for each \( i = L, H \), and (ii) \( V \in (\theta, H(\bar{\theta})) \). Finally, \( k = 1 \) so P’s objective is to maximize her own profit.

Then in the No Supervision (NS) case, P offers a non-contingent price \( p^{NS} \) to maximize \( F(p)[V - p] \). Assumption (ii) above guarantees an interior solution \( p^{NS} \in (\theta, \bar{\theta}) \) in the optimal NS contract. Let \( W^{NS} \equiv F(p^{NS})[V - p^{NS}] \) denote the resulting expected payoff of P, which is positive. The second-best allocation results when P can costlessly access S’s signal, removing any scope for collusion between S and A. Here for each \( i \in \{L, H\} \), P offers A a price \( p^{SB}_i \) which maximizes \( (V - p_i)F_i(p_i) \) subject to \( p_i \in \Theta \). Then (ii) implies \( p^{SB}_H > p^{NS} > p^{SB}_L \).

We characterize the strong collusion-proof allocation based on Definition 2. It can be verified that Proposition 1 holds in this setting; hence we confine attention to the case \( \alpha < 1/2 \).

Since the output is indivisible, the A’s contract in any state (following reports sent to P) is described by a pair \((p, u)\) where \( u \) is a lump-sum payment to A and \( p \) is an incremental transfer to A when the output is delivered. As A always has the option to refuse the contract, \( u \geq 0 \) without loss of generality. There are three possible reports that could be sent by each party: \( i = L, H, 0 \) where 0 is the auxiliary message. Given a contract \((p_i, u_i)\) in signal state \( i \in \{L, H, 0\} \), A delivers the output if and only if \( p_i \geq \theta \), and earns payoff \( u_A(\theta, i) = \max\{p_i - \theta, 0\} + u_i \).

For each output \( q \in \{0, 1\} \), the coalition can coordinate the reports sent to P that maximizes their total payment \( u_A + u_S + \theta q \). The total payment is denoted \( X_0 \) if \( q = 0 \) and \( X_1 \) if \( q = 1 \). Since the coalition have the option of a zero total payment for \( q = 0 \) in any state by jointly deciding to exit from grand contract, \( X_0 \) cannot be lower than 0. Define \( b \equiv X_1 - X_0 \). Collusion-proofness requires the aggregate output-contingent payments \( b, X_0 \) to be independent of reports submitted. The mechanism
can be thought of as a ‘collective’ output-contingent contract \((b, X_0)\) offered to the coalition, combined with suggested subcontracts for A: \((p_i, u_i)\) in state \(i = L, H, 0\) respectively, with S being the residual claimant. Then in state \((\theta, i)\) such that \(\theta \leq p_i\), S receives \(b + X_0 - p_i - u_i\) and A is paid \(p_i + u_i\), while for \((\theta, i)\) such that \(\theta > p_i\), S ends up with \(X_0 - u_i\) and A with \(u_i\). In a collusion proof mechanism, truthful reporting occurs on the equilibrium path, so the resulting equilibrium allocation can be represented by the vector \((b, X_0, p_L, u_L, p_H, u_H)\), while the mechanism is defined additionally by the subcontract \((p_0, u_0)\) corresponding to the off-equilibrium paths where S submits the auxiliary message 0.

Since \(\alpha < 1/2\), Lemma 2 continues to apply. Hence we can drop S’s participation constraint, and set \(\omega = 0\) without loss of generality. Then we can define the equilibrium allocation vector \((b, X_0, p_L, u_L, p_H, u_H)\) to be strong collusion-proof (as per Definition 2) if there exists \((p_0, u_0)\) with \(u_0 \geq 0\) such that

(a) for each \(i \in \{L, H\}\), \((\tilde{p}_i, \tilde{u}_i) = (p_i, u_i)\) and \(\gamma_{ii} = 1\) maximize the coalition’s payoff:

\[
(1 - \alpha)[F_i(\tilde{p}_i)(b - \tilde{p}_i) + X_0 - \tilde{u}_i] + \alpha\int_{\theta}^{\tilde{p}_i}(\tilde{p}_i - \theta)dF_i(\theta) + \tilde{u}_i \tag{23}
\]

subject to

\[
\max\{\tilde{p}_i - \theta, 0\} + \tilde{u}_i \geq \Sigma_{j \in \{L, H, 0\}}\gamma_{ij}[\max\{p_j - \theta, 0\} + u_j] \tag{24}
\]

for any \(\theta \in \Theta\), \(\gamma_{ij} \in [0, 1]\) for \(j = L, H, 0\) and \(\Sigma_{j \in \{L, H, 0\}}\gamma_{ij} = 1\),

(b) \(p_0 + u_0 \leq b + X_0\) and \(u_0 \leq X_0\).

Condition (a) requires that following signal \(i\), the coalition does not benefit from manipulating A’s contract from \((p_i, u_i)\) to \((\tilde{p}_i, \tilde{u}_i)\), combined with the threat that S will report \(j \in \{L, H, 0\}\) with probability \(\gamma_{ij}\) if A refuses the side contract. As in the previous Section, when the state reports of A and S do not match, penalties are imposed only on S, while A is offered a contract corresponding to the state reported by S. Hence a report of \(j\) by S will cause A to end up with a payoff of \(\max\{p_j - \theta, 0\} + u_j\)
if A’s cost is $\theta$. This explains A’s participation constraint (24). Condition (b) states that the coalition does not ever benefit from jointly coordinating on the auxiliary report $i = 0$.

By contrast, with weak collusion S would not be able to commit to an arbitrary reporting threat $\{\gamma_{ij}\}$ to coerce A into accepting the offered side-contract. Instead S and A would both report the true signal realization $i$ if A were to refuse the side contract, so $\gamma_{ii}$ would be constrained to equal 1, and the right-hand-side of (24) would be replaced by $\max\{p_i - \theta, 0\} + u_i$.

The optimal allocation in strong collusion is therefore selected to maximize P’s expected payoff

$$\Sigma_i \kappa_i F_i(p_i)[V - b] - X_0$$

(25)

over the set of strong collusion-proof allocations (described by conditions (a) and (b) above).

Our main results in this section are summarized as follows.

**Proposition 3** Suppose P procures an indivisible good and all the assumptions stated in this Section hold.

(1) Suppose $\alpha = 0$ and $F(p^{NS})a_H(p^{NS}) > F_H(p^{NS})$. Then hiring S is valuable.

(2) If S is valuable at $\alpha \in [0, \frac{1}{2})$, P’s payoff is locally decreasing in $\alpha$, and converges to $W^{NS}$ as $\alpha$ approaches 1/2. Otherwise, P’s payoff equals $W^{NS}$ for all $\alpha$.

(3) P’s payoff in strong collusion is always strictly lower than in weak collusion.

Part (1) provides sufficient conditions for S to be valuable when $\alpha = 0$. Part (2) then states that P’s payoff is decreasing in $\alpha$ over the range where S is valuable and has greater bargaining power than A. It then follows that P’s payoff is non-increasing in $\alpha$ over the entire range $[0, 1]$ and that it is maximized at $\alpha = 0$.

Note that in the divisible good case, Proposition 2 showed that S is generically valuable in Context C provided $\alpha < \frac{1}{2}$. This result does not necessarily hold in the
current context where the good is indivisible. As in the context of weak collusion (see Mookherjee et al. (2020)), the value of $S$ is smaller when there is less scope for varying the quantity procured, thereby reducing the advantage secured from accessing $S$’s information. Part (1) of the above Proposition thus provides sufficient conditions for $S$ to be valuable when the good is indivisible. A specific example where these conditions hold is that of a uniform prior of $\theta$ on $[0, 1]$, with $a_H(\theta)$ satisfying $a_H(\theta) = a_L(1 - \theta)$, $a_H(\theta) = 0$ and $a_H(\theta)$ is strictly convex on $[0, 1/2]$.\footnote{\textit{a}_H(\theta) also needs to satisfy assumption (i) listed at the beginning of this section. These conditions are satisfied with \textit{a}_H(\theta) = 2\theta^2 \text{ for } \theta \in [0, 1/2] \text{ and } 1 - 2(1 - \theta)^2 \text{ for } \theta \in (1/2, 1].}$ With a uniform prior, the condition stated in (a) reduces to

\[
p^{NS}a_H(p^{NS})/2 > \int_0^{p^{NS}} a_H(\theta)d\theta.
\]

Figure 1 shows an example of $a_H(\theta)$ satisfying these sufficient conditions. For some $p^{NS} \in (0, 1)$, $p^{NS}a_H(p^{NS})/2$ is equal to the area of the triangle $A$ in Figure 1, and $\int_0^{p^{NS}} a_H(\theta)d\theta$ is the area under $a_H(\theta)$ between 0 and $p^{NS}$. We can confirm that the latter is less than the area $A$ for any choice of $p^{NS} = V/2 \in (0, 1)$. Thus, in this example, $S$ is valuable for any $V \in (0, 2)$ with $H(1) = 2$.\footnote{\textit{a}_H(\theta) also needs to satisfy assumption (i) listed at the beginning of this section. These conditions are satisfied with \textit{a}_H(\theta) = 2\theta^2 \text{ for } \theta \in [0, 1/2] \text{ and } 1 - 2(1 - \theta)^2 \text{ for } \theta \in (1/2, 1].}
The rest of this section explains various steps in the analysis of the optimal contracting problem, which provides a flavor of the arguments underlying Proposition 3. Detailed proofs are provided in the Supplementary Appendix.

We start by explaining how the strong collusion proofness property can be characterized by a set of inequality constraints in the optimization exercise. Note first that the coalition incentive compatibility (CIC) conditions (a) and (b) can be simplified as follows, given a choice of \((p_L, p_H, b)\) where \(p_H \geq p_L\) and \(p_H \geq b\) (rationalized by the observation that allocations not satisfying these properties are dominated by the optimal NS allocation).\(^{24}\) In particular, the condition \(p_H \geq b\) implies the use of ‘countervailing incentives’, wherein S and A’s interests (with regard to supply) are opposed in state H: S would be worse off while A would be better off if the good is supplied (which would happen whenever \(\theta < p_H\)).

Given P’s objective function (25), we first need to find the lowest lump sum compensation \(X_0\) which can implement the given choice of \((p_L, p_H, b)\). Then at the second stage, we can optimize over \((p_L, p_H, b)\). We start with the former step: how all other dimensions of the mechanism \((u_H, u_L, u_0, p_0)\) can be chosen to minimize the cost \(X_0\) of implementing the desired allocation \((p_L, p_H, b)\) as a strong collusion-proof allocation. Contrasting this with the corresponding constraint for \(X_0\) in weak collusion helps provide some insight into the extra costs entailed by strong collusion.

Condition (a) can be restated as follows. Recall the definition of \(\beta \equiv \frac{1 - 2\alpha}{1 - \alpha}\), so that \(1 - \beta = \frac{\alpha}{1 - \alpha}\) is the ratio of welfare weight of A relative to S. The coalition’s objective in (a) can be represented by the Lagrangean expression

\[
\mathcal{L} = X_0 + \int_\theta \hat{p}_i \left[ b - W_i(\theta; \beta) \right] dF_i(\theta) - \beta \sum_j \gamma_{ij} \left[ \int_\theta \Lambda_i(\theta) d\theta + u_j \right]
\] (26)

\(^{24}\)Given the distributional assumptions, the ‘supply elasticity’ is lower in state L, so it is intrinsically desirable for A to be offered \(p_H\) in state H which is not lower than \(p_L\). And any allocation where \(b\) is larger than \(p_H\) is dominated by the NS allocation where A is offered \(\hat{p} = p_H\), as it would raise the likelihood of supply in state L, leave it unchanged in state H, while lowering the price paid by P for delivery of the good.
where $\Lambda_i(\theta)$ is non-decreasing with $\Lambda_i(\theta) = 0, \Lambda_i(\bar{\theta}) = 1$. $\beta \Lambda_i'(\theta)$ can be interpreted as the ‘shadow cost’ multiplier associated with the participation constraint of type $\theta$ of $A$, and $W_i(\theta : \beta) \equiv \theta + \beta \frac{F_i(\theta) - \Lambda_i(\theta)}{f_i(\theta)} = \theta + \frac{F_i(\theta)}{f_i(\theta)} - \{(1 - \beta) \frac{F_i(\theta)}{f_i(\theta)} + \beta \Lambda_i'(\theta)\}$ denotes a welfare-weighted virtual cost to the coalition of supplying the good. $W_i$ lowers the virtual cost $\theta + \frac{F_i(\theta)}{f_i(\theta)}$ of $S$ by the sum of the following two terms: (i) the associated gain in $A$’s information rent $\frac{F_i(\theta)}{f_i(\theta)}$, adjusted by $A$’s relative welfare weight $(1 - \beta)$, and (ii) $\beta \Lambda_i'(\theta)$ as required to satisfy $A$’s type dependent participation constraint. The expression $\beta \left[ \int_0^{p_i} \Lambda_i(\theta)d\theta + u_j \right]$ can be interpreted as the total shadow cost of meeting $A$’s participation constraint (aggregating across all relevant types $\theta \leq p_j$ of $A$, in state $i$ when report $j$ is submitted by the coalition).25

The control variables of the coalition are the deviating contract offered to $A$ in state $i = L, H$: delivery prices $\tilde{p}_i$ and fixed payments $\tilde{u}_i$, besides the probabilities $\gamma_{ij}$ corresponding to the threats of $S$ to report $j$ in true state $i$ should $A$ veto the offered side contract. Three alternate reports can be submitted: corresponding to states $H, L$ and the auxiliary message $0$.

Expression (26) generates necessary conditions for CIC. First, for the coalition to not want to deviate from the prices $p_i$ set by $P$, it is necessary that in each state $i = L, H$:

$$\tilde{p}_i = p_i \quad \text{maximizes} \quad \int_{\bar{\theta}}^{\tilde{p}_i} [b - W_i(\theta : \beta)]dF_i(\theta) \quad (27)$$

i.e., the coalition selects a price to maximize its net surplus defined by the expected difference between the bonus $b$ earned for delivery of the good, and the welfare-weighted virtual cost $W_i$ of supplying it. This generates (for interior $p_H, p_L \in (\bar{\theta}, \tilde{\theta})$) the following necessary condition: $b = W_i(p_i : \beta) \equiv p_i + \beta \frac{F_i(p_i) - \Lambda_i(p_i)}{f_i(p_i)}$. Since $\Lambda_i(p_i) \in [0, 1]$, this requires

$$p_i - \beta \frac{1 - F_i(p_i)}{f_i(p_i)} \leq b \leq p_i + \beta \frac{F_i(p_i) - \Lambda_i(p_i)}{f_i(p_i)} \quad (28)$$

for $i = L, H$. This condition restricts the extent to which the coalitional bonus $b$ can deviate from the price $p_i$ offered to $A$ for delivering the good, i.e., the extent to which

25This follows from the fact that $\beta \int_0^{p_i} \Lambda_i(\theta)d\theta$ equals $\int_0^{p_i} (p_j - \theta)d(\beta \Lambda_i(\theta))$. 32
S and A’s interim supply incentives diverge.

Next, CIC requires that the coalition does not benefit from manipulating A’s outside option with a threat of S sending a false report if A refuses the offered side contract. Hence in state $i \gamma_{ii} = 1, \gamma_{ij} = 0$ if $j \neq i$ should maximize (26), which requires:

$$\int_{\underline{a}}^{p_i} \Lambda_i(\theta)d\theta + u_i \leq \int_{\underline{a}}^{p_j} \Lambda_i(\theta)d\theta + u_j$$

(29)

for all $j \neq i$. In particular, (29) must hold for the auxiliary ‘exit’ report $j = 0$. Since this auxiliary report is never actually sent in equilibrium, it is optimal for P to raise the associated shadow cost (i.e., $u_0, p_0$) as far as possible.

The limit to which $u_0, p_0$ can be raised is given by condition (b) of CIC: $u_0$ can be set equal to $X_0$ and $p_0$ to $b$. In other words, if the auxiliary report is submitted by S, the latter exits and the entire coalitional contract $(X_0, b)$ is offered by P to A.

Therefore (29) requires for each $i = L, H$:

$$\int_{\underline{a}}^{p_i} \Lambda_i(\theta)d\theta + u_i \leq \int_{\underline{a}}^{b} \Lambda_i(\theta)d\theta + X_0$$

(30)

which in turn provides a lower bound to the fixed payment $X_0 \geq 0$:

$$X_0 \geq \max\{0, \max_{i=L,H} \{\int_{\underline{a}}^{p_i} \Lambda_i(\theta)d\theta + u_i\}\}$$

(31)

Next, we obtain a lower bound for the shadow cost of A’s participation constraint $\int_{b}^{p_i} \Lambda_i(\theta)d\theta$. CIC implies the coalition should not benefit from altering the price $p_i$ to any alternative price $p'$:

$$\int_{p'}^{p_i} [b - W_i(\theta : \beta)]dF_i(\theta) \geq 0$$

(32)

which reduces to

$$\int_{p'}^{p_i} \Lambda_i(\theta)d\theta \geq -\frac{1}{\beta} \int_{p'}^{p_i} [b - \theta - \frac{F_i(\theta)}{f_i(\theta)} + (1 - \beta)\frac{F_i(\theta)}{f_i(\theta)}]dF_i(\theta)$$

$$= L_i(p_i, b : \beta) - L_i(p', b : \beta)$$

(33)

where $\beta L_i(p, b : \beta) \equiv [(p - b)F_i(p) - (1 - \beta)\int_{b}^{p} F_i(\theta)d\theta]$ represents the interim (welfare weighted) loss of rents to the coalition from offering delivery bonus of $p$ to A (rather
than \( b \) the bonus received by the coalition), measured in units of S’s rents.\(^{26}\) Condition (33) can thus be interpreted as saying that the reduction in coalitional rent loss resulting from a deviation from \( p_i \) to \( p' \) should be outweighed by the corresponding increase in shadow cost of meeting A’s participation constraint. Putting \( p' = b \) we obtain the following lower bound for \( \int_b^{p_i} \Lambda_i(\theta)d\theta \):

\[
\int_b^{p_i} \Lambda_i(\theta)d\theta \geq L_i(p_i, b : \beta)
\]  

(34)

for each \( i = L, H \) since \( L_i(b, b : \beta) \equiv 0 \).

It still remains to find lower bounds for \( u_i, i = L, H \). Condition (29) combined with (33) then implies the lowest values that we can set for these lump-sum payments to A are \( u_H = 0, u_L = \int_{p_L}^{p_H} \Lambda_H(\theta)d\theta \geq L_H(p_H, b : \beta) - L_L(p_L, b : \beta) \). Combining this with the expressions obtained above, we see that the lower bound on the fixed payment \( X_0 \) needed to implement the required allocation as a strong collusion proof outcome is

\[
X_0(p_L, p_H, b : \beta) = \max\{0, L_H(p_H, b : \beta), L_L(p_L, b : \beta) + L_H(p_H, b : \beta) - L_L(p_L, b : \beta)\}
\]  

(35)

It is instructive to compare this with the corresponding expression for the minimum \( X_0 \) in weak-collusion (see Mookherjee et al. (2020, Online Appendix)\(^{27}\)):

\[
X_0(p_L, p_H, b) = \max\{0, F_H(p_H)(p_H - b), F_L(p_L)(p_L - b)\}
\]  

(36)

Intuitively, \( F_i(p_i)(p_i - b) \) is the interim loss borne by S owing to countervailing incentives \( p_i > b \) in state \( i \). To ensure S wants to participate, P must provide at least this much lump-sum compensation to the coalition in every state under weak collusion. For state \( H \), this minimum compensation \( F_H(p_H)(p_H - b) \) is weakly smaller than the corresponding expression \( L_H(p_H, b : \beta) \) appearing in the second argument

\(^{26}\)S’s interim loss is \((p - b)F_i(p)\), from which we subtract the associated welfare-weighted expected rent \( \int_b^{p_i} F_i(\theta)d\theta \) earned by A.

on the right-hand-side of (35).\textsuperscript{28} The lower bound on required compensation in weak collusion results from the greater ability of P to manipulate A’s outside options (because this problem does not include constraint (29), and neither can the coalition coordinate on an auxiliary report).

Return now to the second step of the strong collusion analysis. The Supplementary Appendix shows that the lower bound (35) can indeed be achieved, provided the following additional ‘monotonicity’ condition is satisfied:

\[ L_L(p_H, b : \beta) - L_L(p_L, b : \beta) \geq L_H(p_H, b : \beta) - L_H(p_L, b : \beta) \quad (37) \]

which also happens to be necessary for coalition incentive compatibility.\textsuperscript{29}

Therefore the second stage problem of finding an optimal allocation \((p_L, p_H, b)\) reduces to maximization of (25) with required compensation equal to \(X_0(p_L, p_H, b : \beta)\), with attention restricted to \(p_H \geq \max\{b, p_L\}\) in addition to constraints (28) and (37).

This problem can be solved as follows. Fix a value of \(\beta\). Let the optimum allocation be denoted \((p^*_L, p^*_H, b^*)\) and suppose S is valuable (so \(p^*_H > \max\{b^*, p^*_L\}\)). Moreover suppose the maximum for \(X_0(p^*_L, p^*_H, b^* : \beta)\) is achieved at \(L_H(p^*_H, b^* : \beta)\), which strictly exceeds \([L_L(p^*_L, b^* : \beta) + L_H(p^*_H, b^* : \beta) - L_L(p^*_L, b^* : \beta)]\). Also for convenience suppose that the monotonicity constraint (37) does not bind. Now suppose A’s welfare weight falls, causing \(\beta\) to go up to \(\beta' > \beta\). The proof is based on showing that (i) the same allocation \((p^*_L, p^*_H, b^*)\) continues to be feasible at \(\beta'\), while (ii) the cost \(X_0\) of implementing it goes down.

Part (1) of Proposition 3 follows from observing that the allocation continues to satisfy constraint (28) at \(\beta'\), since this constraint is weakened further. Intuitively, the decrease in A’s welfare weight increases the conflict of interest between S and A, thus widening the CIC-feasible gap between their respective interim supply incentives. And

\textsuperscript{28}If \(\beta = 1\) these expressions are exactly the same. When \(\beta < 1\), it follows from recalling that we can restrict attention to allocations with countervailing incentives in state \(H\) (i.e., \(p_H \geq b\)), and then observing that \(L_H(p_H, b : \beta) \geq F_H(p_H)(p_H - b)\) if and only if \(F_H(p_H)(p_H - b) \geq \int_{p_L}^{p_H} F_H(\theta)d\theta\).

\textsuperscript{29}The latter follows from the necessity of \(\int_{p_L}^{p_H} \Lambda_H(\theta)d\theta \leq \int_{p_L}^{p_H} \Lambda_L(\theta)d\theta\), which is implied by (29).
part (2) follows from observing that $L_H(p^*_H, b^*: \beta') < L_H(p^*_H, b^*: \beta)$, as $L_H(p, b : \beta)$ is strictly decreasing in $\beta$ whenever there are countervailing incentives at state $H$, i.e., $b < p_H$ (since the slope of $L_H$ with respect to $\beta$ has the same sign as $\int_b^{p_H} [F_H(\theta) - F_H(p_H)]d\theta < 0$). As $S$ shares rents with $A$ when $A$ has greater bargaining power, the compensation that $P$ must pay the coalition for the costs imposed by countervailing incentives in state $H$ goes up.

Part (3) of Proposition 3 states that strong collusion is always more costly to $P$ than weak collusion. The proof is based on showing that the optimal allocation in weak collusion involves provision of countervailing incentives in both states. In particular, this includes state $L$ (i.e., $p_L > b$), since starting with $p_H > b, p_L \leq b$ the marginal cost of raising $p_L$ slightly equals the associated increase in $X_0 = \max\{0, F_H(p_H)(p_H - b), F_L(p_L)(p_L - b)\}$. Given $p_H > b, p_L \leq b$, the expression for $X_0$ for local changes in $p_L$ reduces to $F_H(p_H)(p_H - b)$, which is independent of $p_L$. Hence the marginal cost of raising $p_L$ slightly is zero, while the marginal benefit (resulting from a higher likelihood of the good being supplied in state $L$) is positive, making it worthwhile to raise $p_L$ above $b$ in weak collusion. The proof is completed by showing that this is not the case in strong collusion, owing to the different expression for $X_0$ in that setting.\footnote{From part (b), it suffices to show this when $\alpha = 0$. If the result is false, there is an allocation which is optimal in both weak and strong collusion contexts. We show that the optimal strong collusion mechanism always sets $p_L$ equal to $b$, owing to a positive and sufficiently large marginal cost (in terms of higher compensation $X_0$ required) of raising $p_L$ slightly above $b$.}

References


30


Appendix: Proof of Proposition 2

Proof of Lemma 2: Suppose that for some \( \eta \), \((I(\theta, \eta), u_A(\theta, \eta), u_S(\theta, \eta), I(\eta))\) does not solve the relaxed version of \( \mathcal{PS}(\alpha : \eta) \) where the constraint \( E[u_S(\theta, \eta) \mid \eta] \geq \omega \) is dropped. It implies \( E[u^*_S(\theta, \eta) \mid \eta] < \omega \) in the optimal solution of the relaxed problem represented by 
\[
(\mu^*(\theta, \eta), \tilde{u}_A(\theta, \eta), \tilde{u}_S(\theta, \eta), P_r(\cdot \mid \eta)).
\]

As shown in the proof of Proposition 1, side contract \( \tilde{SC} \) defined as follows is feasible in \( \mathcal{PS}(\alpha : \eta) \), hence also in the relaxed problem:

- \( \bar{\mu}(\theta, \eta) = \mu^*(\theta) \) which maximizes \( X^e(\mu) - \theta q^e(\mu) \) subject to \( \mu \in \Delta(\bar{K} \cup \{e\}) \)
- \( P(\eta_0 \mid \eta) = 1 \) and \( P(\eta' \mid \eta) = 0 \) for any \( \eta' \neq \eta_0 \)
- \( \bar{u}_A(\theta, \eta) = X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) - \omega \) (denoted by \( u^+_A(\theta, \eta) \) in later part)
- \( \bar{u}_S(\theta, \eta) = \omega \)

Hence
\[
E[\alpha \bar{u}_A(\theta, \eta) + (1 - \alpha) \bar{u}_S(\theta, \eta) \mid \eta] \\
= E[(1 - \alpha)\{X^e(\mu^*(\theta, \eta)) - \theta q^e(\mu^*(\theta, \eta))\} + (1 - 2\alpha)\bar{u}_A(\theta, \eta) \mid \eta] \\
\geq E[(1 - \alpha)\{X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta))\} + (1 - 2\alpha)u^+_A(\theta, \eta) \mid \eta].
\]

But since \( E[X^e(\mu^*(\theta, \eta)) - \theta q^e(\mu^*(\theta, \eta)) \mid \eta] \leq E[X^e(\mu^*(\theta)) - \theta q^e(\mu^*(\theta)) \mid \eta] \) by the definition of \( \mu^*(\theta) \), \( \alpha < \frac{1}{2} \) implies that \( E[u^+_A(\theta, \eta) \mid \eta] \geq E[\bar{u}_A(\theta, \eta) \mid \eta] \). This implies that the side contract \( \tilde{SC} \) creates a Pareto improvement over the solution to the relaxed problem, yielding a strictly higher value of the third party’s expected payoff, a contradiction.

The next step in the proof of Proposition 2 is to consider the specific mechanism described in the text; we establish this allocation is SCP provided conditions (a) and (b) are satisfied.
Owing to the previous lemma, we can drop S’s participation constraint (iv) from problem $P^S(\alpha : \eta)$. So consider the relaxed problem denoted by $\bar{P}^S(\alpha : \eta)$, for this allocation defined on $\Theta \times \{\eta_1, \eta_2\}$, which selects $(\mu(\theta, \eta), \bar{u}_A(\theta, \eta), p(\eta))$ to maximize

$$E[X^\alpha(\mu(\theta, \eta)) - \theta q^c(\mu(\theta, \eta)) - \beta \tilde{u}_A(\theta, \eta) | \eta]$$

subject to $\mu(\theta, \eta) \in \Delta(\Theta \times \{\eta_1, \eta_2\} \cup \{e\})$ and $p(\eta) \in [0, 1]$,

$$\tilde{u}_A(\theta, \eta) \geq p(\eta)u_A(\theta, \eta) + (1 - p(\eta))u_A(\theta, \eta')$$

and

$$\tilde{u}_A(\theta, \eta) \geq \tilde{u}_A(\theta', \eta) + (\theta' - \theta)q^c(\mu(\theta', \eta))$$

for any $\theta, \theta' \in \Theta$.

Specifically, we aim to show that $(\mu(\theta, \eta), \bar{u}_A(\theta, \eta), p(\eta)) = ((\theta, \eta), u_A(\theta, \eta), 1)$ solves $\bar{P}^S(\alpha : \eta)$, if

(a) $E[u_A(\theta, \eta_1) - u_A(\theta, \eta_2) | \eta_2] \geq 0$

(b) $\int_\theta^{\theta'} [u_A(\theta, \eta_2) - u_A(\theta, \eta_1)]d\Lambda(\theta) \geq 0$.

Upon choosing $\Lambda(\cdot, \eta_1) \equiv \Lambda(\cdot)$ and $\Lambda(\cdot, \eta_2) \equiv F(\cdot | \eta_2)$, we can combine (a) and (b) into the following single condition

$$\int_\theta^{\theta'} [u_A(\theta, \eta') - u_A(\theta, \eta)]d\Lambda(\theta, \eta) \geq 0$$

when $\eta, \eta' \in \{\eta_1, \eta_2\}$ and $\eta \neq \eta'$.

Since $\Lambda(\theta, \eta)$ is non-decreasing in $\theta$, this condition implies that

$$0 \leq \int_\theta^{\theta'} [\tilde{u}_A(\theta, \eta) - p(\eta)u_A(\theta, \eta) - (1 - p(\eta))u_A(\theta, \eta')]d\Lambda(\theta, \eta)$$

$$\leq \int_\theta^{\theta'} [\tilde{u}_A(\theta, \eta) - u_A(\theta, \eta)]d\Lambda(\theta, \eta)$$
for any $(\tilde{u}_A(\theta, \eta), p(\eta))$ satisfying constraints of $\bar{P}^S(\alpha : \eta)$. This result can be used to obtain an upper bound of the objective function in $\bar{P}^S(\alpha : \eta)$. First note that

$$E[X^e(\mu(\theta, \eta)) - \theta q^e(\mu(\theta, \eta)) - \beta \tilde{u}_A(\theta, \eta) | \eta]$$

$$\leq E[X^e(\mu(\theta, \eta)) - \theta q^e(\mu(\theta, \eta)) - \beta \tilde{u}_A(\theta, \eta) | \eta] + \beta \int_\theta^\bar{\theta} [\tilde{u}_A(\theta, \eta) - u_A(\theta, \eta)]d\Lambda(\theta, \eta)$$

$$= E[X^e(\mu(\theta, \eta)) - z_\beta(\theta, \eta)q^e(\mu(\theta, \eta)) | \eta] - \beta \int_\theta^\bar{\theta} u_A(\theta, \eta)d\Lambda(\theta, \eta).$$

The second equality uses the fact that

$$\tilde{u}_A(\theta, \eta) = \tilde{u}_A(\bar{\theta}, \eta) + \int_\theta^\bar{\theta} q^e(\tilde{\mu}(y, \eta))dy.$$

Next, note that $\tilde{\mu} = (\theta, \eta)$ maximizes $X^e(\tilde{\mu}) - z_\beta(\theta, \eta)q^e(\tilde{\mu})$. This implies that an upper bound to the value of the objective function is given by:

$$E[X(\theta, \eta) - \theta \tilde{q}(z_\beta(\theta, \eta)) - \beta u_A(\theta, \eta) | \eta].$$

But this is attainable with $(\mu(\theta, \eta), \tilde{u}_A(\theta, \eta), p(\eta)) = ((\theta, \eta), u_A(\theta, \eta), 1)$ (which satisfies all constraints) in $\bar{P}^S(\alpha : \eta)$, implying that it is the optimal solution of this problem. This implies the allocation is SCP.

Let $Z(\eta_1)$ denote the set of non-decreasing functions $z : \Theta \to \mathbb{R}$ such that $z(\theta) = \theta + \beta \frac{F(\theta|m) - \Lambda(\theta)}{f(\theta|m)}$ for some $\Lambda(\theta)$ which is non-decreasing in $\theta$ with $\Lambda(\theta) = 0$ and $\Lambda(\bar{\theta}) = 1$. In order to prove Proposition 2, it suffices to construct $z_\beta(\cdot) \in Z(\eta_1)$ where (17), (18) and (19) are satisfied at the same time. The rest of the proof is devoted to this construction.

**Step 1:** Under the hypothesis of Proposition 2, there exist $(\lambda_1, \lambda_2)$ and closed intervals on $\Theta$ $(\Theta_1 = [\theta_1, \bar{\theta}_1], \Theta_2 = [\theta_2, \bar{\theta}_2]$ and $\Theta_3 = [\theta_3, \bar{\theta}_3]$) such that $\underline{\theta} < \theta_i < \bar{\theta}_i$.

\[\text{By definition of } (X^e(\mu), q^e(\mu)), \]

$$X^e(\theta, \eta) - z_\beta(\theta, \eta)q^e(\theta, \eta) = \bar{X}(z_\beta(\theta, \eta)) - z_\beta(\theta, \eta)\tilde{q}(z_\beta(\theta, \eta)).$$
\( \theta_{i+1} < \bar{\theta}_{i+1} < \bar{\theta} \ (i = 1, 2) \), and the sign of

\[
\frac{F(\theta)}{f(\theta)} f(\theta | \eta_1) + \lambda_1 F(\theta | \eta_2) - \lambda_2 F(\theta | \eta_1)
\]

alternates among the interiors of \( \Theta_1, \Theta_2 \) and \( \Theta_3 \). The proof is provided in the Supplementary Appendix. In later analysis, our focus is restricted to the case that there exists \((\lambda_1, \lambda_2)\) such that

\[
\frac{F(\theta)}{f(\theta)} f(\theta | \eta_1) + \lambda_1 F(\theta | \eta_2) - \lambda_2 F(\theta | \eta_1)
\]

is negative on the interior of \( \Theta_1 \) and \( \Theta_3 \), and positive on the interior of \( \Theta_2 \). We can adopt the same analysis for the opposite case.

**Step 2:** For any closed interval \([\theta', \theta''] \subset \Theta\) such that \( \bar{\theta} < \theta' < \theta'' < \bar{\theta} \), there exists \( \delta > 0 \) so that \( z(\cdot) \in Z(\eta_1) \) for any function \( z(\cdot) \) satisfying the following properties:

(i) \( z(\cdot) \) is increasing and differentiable with \( |z(\theta) - \theta| < \delta \beta \) and \( |z'(\theta) - 1| < \delta \beta \) for any \( \theta \in \Theta \)

(ii) \( z(\theta) = \theta \) for any \( \theta \notin [\theta', \theta''] \).

The proof is provided in the Supplementary Appendix.

**Step 3:** There exists \( z_\beta(\cdot) \in Z(\eta_1) \) satisfying (17) – (19), and for which

\[
E[V(\bar{q}(z_\beta(\theta))) - \frac{2k-1}{k} \bar{X}(z_\beta(\theta)) - \frac{1-k}{k} \theta \bar{q}(z_\beta(\theta)) | \eta]
\]

\[
> E[V(\bar{q}(\theta)) - \frac{2k-1}{k} \bar{X}(\theta) - \frac{1-k}{k} \theta \bar{q}(\theta) | \eta].
\]

The proof is provided in the Supplementary Appendix.

Step 3 implies that \( P\)'s payoff is improved over the optimal \( NS \), and the proof is completed. 

\[\square\]