Theoretical Economics 15 (2020), 1547–1586 1555-7561/20201547

Collusion and delegation under information control

Andreas Asseyer
School of Business and Economics, Freie Universität Berlin

This paper studies how information control affects incentives for collusion and optimal organizational structures in principal–supervisor–agent relationships. I consider a model in which the principal designs the supervisor's signal on the productive agent's private information, and the supervisor and agent may collude. I show that the principal optimally delegates the interaction with the agent to the supervisor if either the supervisor's budget is large or the value of production is small. The principal prefers direct communication with the supervisor and agent if the supervisor's budget is sufficiently small and the value of production is high.

Keywords. Collusion, information design, delegation.

JEL classification. D73, D83, D86, H57, M55.

1. Introduction

In many organizations, a supervisor advises the principal on how to set the contractual terms for a productive agent with private information. Honest advice from the supervisor lowers the agent's informational advantage and allows the principal to reduce the agent's information rent. However, this creates scope for collusion as the agent is willing to pay the supervisor for biased advice that increases his rent. Public procurement is a prominent example of such a setting.[1] In many countries, procurement officers at central purchasing bodies advise public buyers and corruption is prevalent—often in the form of private suppliers paying bribes in exchange for highly priced public contracts.[2]

The threat of collusion influences the optimal organizational design of principal–supervisor–agent relationships. The extant literature studies whether hierarchical delegation is an optimal response to collusion, i.e., whether the principal can achieve the

Andreas Asseyer: andreas.asseyer@fu-berlin.de
I thank three anonymous referees and Roland Strausz for very helpful comments and suggestions. Furthermore, I thank Helmut Bester, Yves Breitmoser, Gorkem Celik, Jacques Crémer, Eddie Dekel, Matthias Lang, Thomas Mariotti, Vincent Meisner, Martin Pollrich, Sebastian Schweighofer-Kodritsch, and audiences at Haifa, Mannheim, Vienna, the 8th Berlin IO day, EEA–ESEM 2018, EARIE 2017, European Winter Meeting of the Econometric Society 2016, GAMES 2016, the annual meeting of the Verein für Socialpolitik 2016, and the Berlin Micro-Colloquium for useful comments and feedback.

[1]Organization for Economic Cooperation and Development (OECD) governments spend an average of 29% of total expenditure on public procurement (OECD 2017).

[2]In a recent corruption scandal, an employee of the Italian central purchasing body, Consip, was allegedly bribed for the award of a public contract worth 2.7 billion euro (ANSA March 1, 2017). According to the OECD (2014), 57% of cases of foreign bribery payments were made to receive a public contract. See Di Tella and Schargrodsky (2003) for more detailed evidence of this type of procurement fraud in hospitals in Buenos Aires.

© 2020 The Author. Licensed under the Creative Commons Attribution-NonCommercial License 4.0. Available at https://econtheory.org. https://doi.org/10.3982/TE3556

payoff from the optimal centralized mechanism by contracting only with the supervisor who in turn designs the agent's contract (Faure-Grimaud et al. 2003, Celik 2009). The literature assumes that the supervisor's information about the agent is exogenous to the principal and draws a different conclusion of the optimality of delegation depending on the specific information structure that generates the supervisor's information.[3]

In this paper, I study whether delegation is an optimal response to collusion if the principal can influence what the supervisor learns about the agent. In many settings where the principal–supervisor–agent model can be applied, the development of information technology has made the allocation of information endogenous. In public procurement, most advanced economies have digitalized their procurement systems. With an e-procurement system, the allocation of information to the different stakeholders is an important choice variable.[4] These technical developments bring up not only the question of whether information control and delegation are substitutes or complements, but also how information control could be used to fight collusion in supervisory institutions.

To study endogenous information, I consider a standard principal–supervisor–agent model as in Faure-Grimaud et al. (2003) and Celik (2009), and add information control of the principal over the supervisor. The agent can produce a good for the principal at a privately known cost. The supervisor observes a signal of the agent's costs. In the spirit of the literature on Bayesian persuasion (Kamenica and Gentzkow 2011), the principal exerts information control by choosing the information structure that generates the supervisor's signal. Under centralization, the principal offers a grand contract to the supervisor and the agent. This centralized mechanism allows both the supervisor and the agent to send messages to the principal. The supervisor and the agent collude by signing an enforceable side contract that specifies side payments and coordinates their behavior under the grand contract. Under delegation, the principal sets a grand contract under which only the supervisor can send messages. The supervisor and the agent can still sign a side contract that now serves as a subcontract between the lower tiers of the hierarchy.

I show that the principal can implement the optimal centralized outcome by delegation if either the supervisor's budget is large or the principal's value of the good is small. In particular, the principal can extract the full surplus by delegation if the supervisor's budget is large enough. By contrast, centralization outperforms delegation if the supervisor's budget is sufficiently small and the principal's value of the good is large. Moreover, a partially informed supervisor is optimal for the principal whenever the supervisor's budget is strictly positive.

As pointed out by Faure-Grimaud et al. (2003), the key difference between centralization and delegation is the agent's outside option from rejecting the side contract offered by the supervisor. Under centralization, the agent can reject the side contract and still

[3]Faure-Grimaud et al. (2003) prove that delegation is optimal in a model where the agent has one of two types and the supervisor may observe one of two signal realizations. Celik (2009) demonstrates that delegation may be strictly suboptimal when the agent has three types and the supervisor observes the element of a partition of the agent's type space in which the true type lies.

[4]Twenty-nine OECD countries used a national e-procurement system in 2016 (OECD 2017).

participate in the grand contract non-cooperatively. Under delegation, the agent has to accept the side contract to participate in the grand contract. Thus, the principal can use the agent's rent from the non-cooperative equilibrium of the grand contract as an additional instrument under centralization. This instrument might be valuable as the agent's rent in the non-cooperative equilibrium of the grand contract determines his bargaining position within the colluding coalition. A higher rent improves the agent's bargaining position and makes it harder for the supervisor to find a profitable side contract. If the instrument is valuable to the principal, centralization is better than delegation; otherwise, delegation is as good as centralization.

With information control, the principal faces a trade-off between information elicitation and collusion prevention. If the supervisor receives additional information, the agent's informational advantage over the supervisor decreases. As long as the supervisor shares her information truthfully, this is beneficial to the principal. However, it also reduces information asymmetry in the colluding coalition, and, therefore, enables the supervisor and the agent to collude more effectively. Faure-Grimaud et al. (2003) already note that the principal's payoff is maximal if the supervisor's signal has an intermediate precision. In this paper, I analyze how this trade-off can be optimally resolved in a model with an arbitrary type and signal space where precision cannot be captured by a single parameter.

In Section 4, I derive upper bounds on the principal's payoff for the cases of centralization and delegation. These upper bounds exceed the principal's payoff in the absence of a supervisor whenever the supervisor has a strictly positive budget. Thus, a supervisor may only be helpful to the principal if she can absorb some loss.

In Section 5, I present a combination of information structure and grand contract with which the principal attains the upper bound on the payoff with delegation. Under this combination, there is a cutoff cost level such that the good is produced if the cost is (weakly) below the cutoff and the good is not produced otherwise. The information structure generates a different signal realization for each type below the cutoff. The types above the cutoff randomly generate one of these signal realizations. This makes the signal noisy. Under the grand contract, the supervisor and the agent receive for production a total payment equal to the cutoff. The agent is offered a price equal to the unique type below the cutoff that is possible after the signal realization, and the supervisor keeps the difference between the cutoff and this type as a bonus if the agent accepts the offer. Without production, the supervisor has to make a transfer to the principal. Clearly, the agent receives no rent under this grand contract. The supervisor receives a positive rent with production and a negative rent without production. To minimize the supervisor's expected rent for any signal realization, the information structure pools more types above the cutoff into signal realizations that lead to a higher bonus under production. Thus, the bonus and the probability of production are negatively assorted. If the supervisor's budget is large enough, the principal can extract the full surplus in expectation by making the supervisor the residual claimant under production while setting a sufficiently negative payment without production.

In Section 6, I present the main results of this paper. First, delegation is optimal if either the supervisor's budget is large enough or the principal's value of the good is small.

Second, centralization is superior to delegation if the supervisor's budget is sufficiently small and the principal's value of the good is high. As discussed above, delegation is inferior to centralization whenever the principal finds it optimal to use the agent's rent in the non-cooperative equilibrium of the grand contract as an instrument. If this is the case, the agent's rent is strictly positive and the principal's payoff is bounded away from the full surplus. As the principal can extract the full surplus under delegation if the supervisor's budget exceeds a threshold, a continuity argument implies that delegation remains optimal if the supervisor's budget lies in some region below the threshold.

Delegation is also optimal if the principal's value of the good is small. In that case, the cutoff separating producing and nonproducing types is small and the probability of production is low. Even if the supervisor's budget is low, the principal can, therefore, effectively reduce the expected rent of the supervisor by setting negative payments for the supervisor in the relatively frequent case of no production. This is feasible under delegation and can be achieved as in the grand contract described above. In contrast, it is not feasible under delegation to impose losses on the supervisor with production, as the supervisor can always avoid production (and the associated losses) by asking the agent for a prohibitively high bribe for the opportunity to produce. If the principal's value of the good is high, the mass of producing types is large and it is more effective to extract rents from the supervisor by imposing losses with production. I construct a combination of information structure and grand contract under centralization that allows the principal to do this. I show that this combination gives the principal a strictly higher payoff than the optimal payoff under delegation if the principal's value of the good is large and the supervisor's budget is sufficiently small. For this parameter range, the combination of information structure and grand contract is near optimal, as its payoff is a first-order approximation of the upper bound under centralization.

As noted above, this paper contributes to the literature on collusive supervision with adverse selection[5] (Faure-Grimaud et al. 2003, Celik 2009, Mookherjee et al. 2020b) by introducing information control on the principal's side. The literature builds on the approach of Laffont and Martimort (1997, 2000) to mechanism design with collusion by modeling collusion as an enforceable side contract between asymmetrically informed parties.[6,7] In contrast to Faure-Grimaud et al. (2003), Celik (2009), and this paper, Mookherjee et al. (2020b) analyze a model of collusive supervision in which the colluding coalition can enter a side contract before accepting the contract offered by the principal.[8] The participation decision of the agent and supervisor can, therefore,

[5]Collusion and delegation with two productive agents is studied under moral hazard by Baliga and Sjöström (1998) and under adverse selection by Laffont and Martimort (1998).

[6]Che and Kim (2006) study the cost of collusion in a general mechanism design framework.

[7]Green and Laffont (1979) and Tirole (1986) study collusion with symmetric information, Crémer (1996), McAfee and McMillan (1992), and Caillaud and Jéhiel (1998) consider specific mechanisms.

[8]Mookherjee et al. (2020a) apply the model of Mookherjee et al. (2020b) to develop a transaction cost-based theory of international vertical integration. Mookherjee and Tsumagari (2019) consider a variant of this model with a stronger form of collusion.

be part of the collusive agreement in the side contract.[9] They show that delegation is strictly suboptimal in this setting.[10]

In this paper, I model the principal's control over the supervisor's signal in the spirit of the literature on Bayesian persuasion (Kamenica and Gentzkow 2011, Rayo and Segal 2010). Thus, the principal can choose an arbitrary signal design at no cost. In line with the literature on collusive supervision, I assume that the principal cannot himself observe the signal realization, and that the signal is observed by the supervisor and the agent. Thus, the signal realization is public for the players of the mechanism set by the principal.[11]

Bergemann et al. (2015) and Roesler and Szentes (2017) study the implications of information design in models of bilateral trade. Bergemann et al. (2015) analyze the payoffs for the buyer and seller that can result from varying the information the seller possesses on the buyer's valuation. Roesler and Szentes (2017) study the optimal information acquisition of a buyer regarding her valuation. In the current paper, the trading relationship between principal and agent is intermediated by a supervisor whose information can be varied.

This paper is also related to Ortner and Chassang (2018). They analyze a principal–monitor–agent model under moral hazard and show that corruption can be fought by introducing asymmetric information in the colluding coalition through the use of random transfers. In contrast to the present paper, it is therefore the terms of the contract and not the type of the agent that creates asymmetric information within the coalition. Inducing asymmetric information on transfers is costless in their setting, as the principal only cares about expected transfers. Thus, the principal does not face a trade-off between information elicitation and collusion prevention. This trade-off is central to the analysis in this paper.[12]

The remainder of this paper is organized as follows. In Section 2, I illustrate the benefits of information control in a simple example. Section 3 introduces the general model. Section 4 sets up the principal's problem, and provides benchmark and preliminary results. In Section 5, I characterize the optimal combination of information control and grand contract under delegation. Section 6 provides conditions for the optimality of either delegation or centralization under endogenous information. Section 7 discusses several extensions of the model. Section 8 concludes. All proofs can be found in the Appendix.

2. An illustrative example

In this section, I present a simplified version of the general model to show how the principal can benefit from information control while delegating the interaction with the

[9] Further papers that study this form of collusion are Mookherjee and Tsumagari (2004), Dequiedt (2007), Pavlov (2008), Che and Kim (2009), and Che et al. (2018).

[10] Delegation is also often studied based on the model of Crawford and Sobel (1982). See Ivanov (2010) for an analysis of information control in this model.

[11] See also the literature on information design in games (Bergemann and Morris 2013, 2016, Taneva 2019, Mathevet et al. 2020).

[12] Negenborn and Pollrich (2020) study the optimal use of asymmetric information about the contract in a general principal–supervisor–agent model.

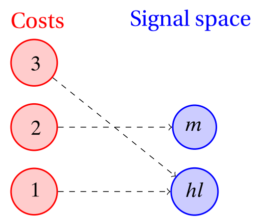

Figure 1. Information structure of a partially revealing signal. Signal realization hl is released if costs are 1 or 3; signal realization m is released if costs are 2.

agent to the supervisor. The agent A can produce a good at cost θ, which is uniformly distributed on $\{1, 2, 3\}$. The principal P values the good by $v \in (2, 3)$. The supervisor S observes a signal σ about θ and is endowed with a budget $\ell \geq 1$. In the absence of S, P optimally offers A the monopsony price $p^* = 1$ and makes an expected payoff of $(v - 1)/3$.

If S is perfectly informed about A's cost and A is unable to pay bribes to S, P can extract the full surplus by delegating the interaction with A to S. In particular, S can be authorized to choose the price that P offers to A. As S is a disinterested party, she finds it optimal to pick $p = \theta$ if $\theta \leq 2$ and $p < 3$ if $\theta = 3$. Given this behavior, P's expected payoff is the expected full surplus of $2/3 \cdot v - 1$.

With collusion and perfect information of S about A's cost, this arrangement is prone to manipulation. In particular, S may promise A always to choose the maximal price P is willing to pay in exchange for a bribe. Under this form of collusion, P's payoff—given the maximal price $\overline{p}$—is $\Pr(\theta \leq \overline{p})(v - \overline{p})$, weakly smaller than the monopsony payoff of $(v - 1)/3$. Thus, collusion destroys all benefits from supervision if S is perfectly informed about A.

Can P be better off if S knows less about A's costs? Suppose S perfectly learns A's cost whenever $\theta = 2$, but cannot distinguish $\theta = 1$ from $\theta = 3$. The information structure underlying S's signal is depicted in Figure 1. Supervisor S either observes the signal realization $\sigma = m$ and knows that $\theta = 2$ or observes $\sigma = hl$ and updates her beliefs to the uniform distribution over $\{1, 3\}$.

Furthermore, suppose P authorizes S to pick a price offer p to A. Principal P pays S a transfer of $2 - p$ if A accepts p, and S pays 1 to P if A rejects p. If the signal realization is m, S optimally offers a price of 2 and receives a payoff of 0 as A always accepts. If the signal realization is hl, S optimally offers a price of 1. Agent A accepts the offer if $\theta = 1$ and rejects if $\theta = 3$. For $\theta = 1$, S makes a profit of 1. For $\theta = 3$, S makes a loss of 1. This loss does not exceed her budget ℓ. Thus, S receives an expected payoff of $1/2 \cdot 1 + 1/2 \cdot (-1) = 0$ after the signal realization hl. It follows that S accepts the delegation contract after both signal realizations. The delegation contract is robust to collusion, as the total payment of P to S and A depends only on the production decision. Supervisor S, therefore, has no interest in increasing the price offer to A, as this comes at her own expense.

Under this combination of a partially revealing information structure and a delegation contract, P extracts the full surplus even though S and A can collude. This follows

	a specialist: $^1/_2$	b specialist: $^1/_2$
a project: $^1/_3$	1	3
m project: $^1/_3$	2	2
b project: $^1/_3$	3	1

FIGURE 2. The determinants of A's cost θ. Cost θ depends on project type and specialization. Each project type and specialization is equally likely.

from the observation that the production decision is efficient and neither S nor A receives a positive rent in expectation. Thus, a partially informative signal and delegation are optimal for P. This result extends to the general model as long as S's budget ℓ is large enough.

Principal P's preferred information structure is non-monotone, as it pools nonadjacent types into the same signal realization. Such non-monotone information structures naturally arise in settings where production costs are determined by two dimensions—a project-specific and an agent-specific dimension—and where the supervisor observes only one dimension.

Consider a procurement project that consists of two tasks. Each task is either of type a or of type b. If both tasks have type a (b), the procurement project is an a project (b project). If the tasks have different types, the procurement project is an m project. Agent A is either specialized in tasks of type a or of type b. Agent A's costs are given by 1 plus the number of tasks in which he is not specialized. Agent A knows his specialization and observes the project type. Principal P knows neither the project type nor A's specialization and believes that the project type is independent from the specialization, and that each project type and each specialization is equally likely. Figure 2 summarizes this description.

If S observes only the project type but not A's specialization, her learning process can be represented by the information structure in Figure 1. If S observes that the project is type a or b, she can infer that costs are going to be high or low. If S observes that the project is of type m, she knows that costs are intermediate. Thus, P benefits from disclosing data about the project to S while hiding data about past performances of A.

3. The model

Principal and agent

The principal, P, seeks to procure a single indivisible good. The agent, A, can produce the good at cost θ. Agent A is privately informed about θ, which is the realization of the random variable $\tilde{\theta}$ with distribution $F(\theta) = \Pr(\tilde{\theta} \leq \theta)$ on $\Theta \subseteq \mathbb{R}$ with $\text{Conv}(\Theta) = [\underline{\theta}, \overline{\theta}]$. Principal P's value of the good is $v \in \mathbb{R}$ with $v > \underline{\theta}$. Given a transfer t from P to A, P's payoff is $vX - t$ and A's payoff is $t - \theta X$, where $X \in \{0, 1\}$ denotes the production decision. Production is efficient for $v \geq \theta$. The resulting expected full surplus is

$$\overline{W} \equiv \int_{\underline{\theta}}^{v} (v - \theta)\, dF(\theta).$$

If P and A are the only players, P cannot do better than to offer A the price[13]

$$p^* = p^*(v) \equiv \max \arg \max_{p \in \Theta} (v - p) F(p).$$

I denote the resulting monopsony payoff for P by $\underline{W}$.

Supervisor and information control

The supervisor S learns about A's cost θ by observing the realization σ of the signal $\tilde{\sigma}$. The signal realization is also observed by A, but not by P.[14] Supervisor S's payoff equals the net transfer she receives and does not depend on the production decision. Supervisor S is endowed with a budget $\ell \in [0, \infty)$. She can never incur a loss that exceeds her budget.

Any signal can be represented by an information structure $I = (\Sigma, \mu)$. The set $\Sigma \subseteq \mathbb{R}$ is the set of all signal realizations with the generic element $\sigma \in \Sigma$, and $\mu \in \Delta(\Sigma \times \Theta)$ is a probability measure on the set of possible realizations of cost and signal. The measure μ induces a conditional distribution $G(\theta|\sigma) = \Pr(\tilde{\theta} \leq \theta|\tilde{\sigma} = \sigma)$ and a marginal distribution $H(\sigma) = \Pr(\tilde{\sigma} \leq \sigma)$. Following the literature on Bayesian persuasion, I impose only the requirement of Bayes consistency on the information structure, i.e.,

$$\int_{\Sigma} G(\theta|\sigma)\, dH(\sigma) = F(\theta) \quad \forall \theta \in \Theta.$$

I denote by $\mathcal{I}$ the set of all Bayes-consistent information structures. For a given I, the support of the random variables $(\tilde{\sigma}, \tilde{\theta})$ and $\tilde{\theta}|\sigma$ are denoted by $\mathrm{Supp}(\tilde{\sigma}, \tilde{\theta}) \subset \Sigma \times \Theta$ and $\mathrm{Supp}(\tilde{\theta}|\sigma) \subset \Theta$, respectively.

Principal P exerts information control by choosing the information structure $I \in \mathcal{I}$ that generates S's signal. This contrasts with the extant literature on collusion in principal–supervisor–agent relationships where some information structure in $\mathcal{I}$ is exogenously given. Following the literature on Bayesian persuasion, P can choose any information structure in $\mathcal{I}$ at zero cost.

Allocations and payoffs

An allocation describes whether the good is produced and what transfers are exchanged between the parties. Formally, an allocation is given by

$$(X, t_S, t_A, \tau) \in \{0, 1\} \times \mathbb{R}^3,$$

where t_i is the transfer from P to $i \in \{A, S\}$ and τ is a side transfer from A to S. The allocation (X, t_S, t_A, τ) leads to payoffs of $vX - t_A - t_S$ for P, $t_A - \tau - \theta X$ for A, and $t_S + \tau$ for S.[15]

[13] See, for instance, Section 2.2. in Börgers (2015) for a proof of the optimality of a posted price.

[14] This assumption follows Faure-Grimaud et al. (2003), Celik (2009), and Mookherjee et al. (2020b).

[15] As A knows σ and θ, all monotone transformations $u(\cdot)$ of A's payoff do not change the results. Thus, one may allow A to be risk-averse or to have a limited budget.

Agent A and supervisor S each have an outside option. If S chooses her outside option, it follows that $t_S = \tau = 0$, resulting in a payoff of zero for S. If A chooses his outside option, $X = \tau = t_A = 0$ and A's payoff is zero.

Centralization and delegation under collusion

Following Faure-Grimaud et al. (2003) and Celik (2009), I consider two forms of organizational design: centralization and delegation.

Under centralization, P directly communicates with A and S. Principal P offers A and S a (deterministic) grand contract

$$\beta = \big(X(m_S, m_A), t_S(m_S, m_A), t_A(m_S, m_A)\big),$$

which determines the production decision and transfers from P to S and A as functions of the messages m_S and m_A chosen by S and A from the sets M_S and M_A. If a party rejects the grand contract, A and S receive their outside options. Closely following the literature on collusion in mechanism design, I model collusion as an enforceable side contract between S and A that coordinates the communication with P and specifies side transfers. As in Faure-Grimaud et al. (2003) and Celik (2009), I assume that S proposes the side contract to A in a take-it-or-leave-it offer.[16] Formally, S offers A a (deterministic) side contract

$$\gamma = \big(\rho\big(m^{sc}; \sigma\big), \tau\big(m^{sc}; \sigma\big)\big),$$

which determines the communication with P and the side transfer through the reporting strategy $\rho : M^{sc} \times \Sigma \to M_S \times M_A$ and the payment rule $\tau : M^{sc} \times \Sigma \to \mathbb{R}$. Both ρ and τ are functions of the message m^{sc} chosen by A from the set M^{sc} and the signal realization σ, which is common knowledge of the colluding parties. If A rejects the side contract, A and S play the grand contract non-cooperatively.

Under delegation, P communicates directly only with S via the grand contract, while S communicates with A via the side contract. Formally, P offers S a grand contract β with $M_A = \{m_A\}$ and S offers A a side contract γ. Under delegation, the side contract γ may be interpreted as a subcontract between P's main contractor S and the subcontractor A. If S rejects the grand contract, A and S receive their outside options. If S accepts the grand contract and A rejects the side contract, A receives his outside option and S is forced to send a message that induces no production.

With this definition of delegation, P can make direct transfers to A. This contrasts with Faure-Grimaud et al. (2003) and Celik (2009), where the grand contract has to satisfy $t_A = 0$ under delegation. However, this difference is not substantial as t_A, t_S, and τ can be interpreted as the net transfers in a setting where P pays $t_A + t_S$ to S and S pays $t_A - \tau$ to A.[17] The advantage of this paper's definition of delegation is that the cases of delegation and centralization can be treated in the same framework by using a "collusion-proofness" principle in both instances. Moreover, this formulation fits the leading example of public procurement where private firms are paid by public buyers and not by the central purchasing body.

[16] I show in Section 7 that this assumption is not crucial for the results.
[17] I further comment on this equivalence after defining P's problem under delegation in Section 4.

Timing and equilibrium concept

The timing of the game for the cases of centralization and delegation is as follows:

$t = 0$: Principal P chooses an information structure $I \in \mathcal{I}$.

$t = 1$: Supervisor S and agent A observe I and σ. Agent A furthermore observes θ.

$t = 2$: Principal P offers a grand contract β.

$t = 3$: Under centralization, S and A each accept or reject P's offer β. Under delegation, only S accepts or rejects β.

$t = 4$: Supervisor S offers a side contract γ to A.

$t = 5$: Agent A accepts or rejects γ.

I focus on *perfect Bayesian equilibria* (PBE) with *passive beliefs* in which S offers direct and truthful side contracts whenever they are optimal.[18] In these equilibria, S does not update her belief about θ if A rejects the side contract off the equilibrium path. Moreover, if the set of S's best responses to a grand contract β contains a direct and truthful side contract, then S offers such a side contract to A. This approach follows the concept of *weak collusion-proofness* in Laffont and Martimort (2000).

Remarks on the model

Principal P exercises information control through public information design as both S and A observe the signal realization. An alternative modeling approach allows P to design signals that are privately observed by A and S.[19] There are two reasons for my modeling decision. First, the model remains close to the literature with exogenous information that assumes that the realization of the exogenously determined signal is observed by both A and S (Faure-Grimaud et al. 2003, Celik 2009, Mookherjee et al. 2020b). While the case of private signals is interesting, it is harder to compare to the literature, which does not cover the case of private signals. Second, the model fits my leading application of public procurement where private firms typically have the right to know what data are collected about them by public authorities.

As in Faure-Grimaud et al. (2003) and Celik (2009), I restrict grand contracts and side contracts to be deterministic. In the context of public procurement, deterministic mechanisms seem to comply better with the aim of rewarding public contracts in a transparent way. Moreover, the restriction to deterministic grand contracts may reflect the practical difficulty to commit to a stochastic mechanism.

As the subsequent analysis reveals, both assumptions turn out to be without loss of optimality if the supervisor has a sufficiently large budget. In this case, the principal can extract the full surplus with public information and deterministic mechanisms.

[18] See Fudenberg and Tirole (1991) for a definition of PBE.
[19] These signals might even be made dependent on a report of A.

4. Preliminary analysis

In this section, I use a collusion-proofness principle to formulate P's contracting problems under centralization and delegation. I then derive upper bounds on P's payoff for centralization and delegation. These upper bounds play a crucial role for the analysis of optimal combinations of information structure and grand contract in the subsequent sections.

Collusion-proofness principle

I first invoke a collusion-proofness principle. This approach follows Laffont and Martimort (1997) and allows me to restrict attention to direct and truthful grand contracts under which S offers A the direct and truthful *null side contract* $\gamma_0 \equiv (\rho_0, \tau_0)$ with $\rho_0(\theta; \sigma) \equiv (\sigma, \sigma, \theta)$ and $\tau_0(\theta; \sigma) \equiv 0$ for all $(\sigma, \theta) \in \mathrm{Supp}(\tilde{\sigma}, \tilde{\theta})$.

Lemma 1. *For any grand contract β, any information structure I, and any signal realization $\sigma \in \Sigma$, S has an optimal side contract γ that is direct and truthful.*

For any equilibrium (with centralization or delegation) in which P offers the grand contract β and S offers the side contract γ, there exists a payoff-equivalent equilibrium in which P offers $\beta_0 = \beta \circ \gamma$ with $M_S = \Sigma$ and $M_A = \Sigma \times \Theta$ under centralization and offers $M_S = \Sigma^2 \times \Theta$ under delegation, and S offers γ_0.[20]

For centralization, the collusion-proofness principle implies that it is optimal for P to offer a direct and truthful grand contract under which S does not benefit from nontrivial collusion with A. For delegation, the collusion-proofness principle implies that it is optimal for P to offer a direct grand contract under which S finds it optimal to offer a side contract that truthfully conveys S's information σ and A's information (σ, θ) without using side transfers.[21]

Centralization

Using the collusion-proofness principle, I now formulate P's optimization problem under centralization. The grand contract is acceptable to both S and A if

$$\mathbb{E}\big[t_S(\sigma, \sigma, \tilde{\theta})|\sigma\big] \geq 0, \tag{PC$_S$}$$

$$t_A(\sigma, \sigma, \theta) - \theta X(\sigma, \sigma, \theta) \geq 0 \tag{PC$_A$}$$

for all $(\sigma, \theta) \in \mathrm{Supp}(\tilde{\sigma}, \tilde{\theta})$. Furthermore, S and A find it optimal to individually report their private information truthfully if

$$\mathbb{E}\big[t_S(\sigma, \sigma, \tilde{\theta})|\sigma\big] \geq \mathbb{E}\big[t_S(\hat{\sigma}_S, \sigma, \tilde{\theta})|\sigma\big], \tag{IC$_S$}$$

$$t_A(\sigma, \sigma, \theta) - \theta X(\sigma, \sigma, \theta) \geq t_A(\sigma, \hat{\sigma}_A, \hat{\theta}) - \theta X(\sigma, \hat{\sigma}_A, \hat{\theta}) \tag{IC$_A$}$$

[20]The usual revelation principle does not apply as the side contract is deterministic. The proof uses a revelation principle in terms of payoffs due to Strausz (2003) and incorporates collusion-proofness.

[21]The terms of the side contract are not observable by P. Thus, P might ask A and S to report the terms of the side contract to the grand contract. The collusion-proofness principle implies that P does not benefit from requesting these reports from A and S.

for all $\sigma, \hat{\sigma}_S \in \Sigma$, $(\sigma, \theta), (\hat{\sigma}_A, \hat{\theta}) \in \operatorname{Supp}(\tilde{\sigma}, \tilde{\theta})$. Supervisor S's limited budget (LB) implies that the transfer from P to S needs to satisfy $t_S(\sigma, \sigma, \theta) \geq -\ell$ for all $(\sigma, \theta) \in \operatorname{Supp}(\tilde{\sigma}, \tilde{\theta})$. Finally, a grand contract is collusion-proof if S does not benefit from a nontrivial side contract that A prefers over the null side contract, that incentivizes A to report θ truthfully, and that respects S's limited budget. Formally, the condition (CP^c) requires

$$\gamma_0 \in \arg\max_{\gamma} \mathbb{E}\big[t_S\big(\rho(\tilde{\theta}; \tilde{\sigma})\big) + \tau(\tilde{\theta}; \tilde{\sigma})\big] \text{ subject to}$$

$$t_A\big(\rho(\theta; \sigma)\big) - \tau(\theta; \sigma) - \theta X\big(\rho(\theta; \sigma)\big) \geq t_A(\sigma, \sigma, \theta) - \theta X(\sigma, \sigma, \theta), \tag{$\text{PC}_A^{\gamma,c}$}$$

$$t_A\big(\rho(\theta; \sigma)\big) - \tau(\theta; \sigma) - \theta X\big(\rho(\theta; \sigma)\big) \geq t_A\big(\rho(\breve{\theta}; \sigma)\big) - \tau(\breve{\theta}; \sigma) - \theta X\big(\rho(\breve{\theta}; \sigma)\big), \tag{IC_A^{γ}}$$

$$t_S\big(\rho(\theta; \sigma)\big) + \tau(\theta; \sigma) \geq -\ell \tag{LB^{γ}}$$

for all $\sigma \in \Sigma$, $\theta, \breve{\theta} \in \operatorname{Supp}(\tilde{\theta}|\sigma)$.[22] A side contract γ is feasible under centralization if it satisfies ($\text{PC}_A^{\gamma,c}$), (IC_A^{γ}), and (LB^{γ}). Principal P's problem under centralization is

$$\mathbf{P}^c : \max_{I,\beta} \mathbb{E}\big[vX(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) - t_S(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) - t_A(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})\big]$$
$$\text{subject to } (\text{PC}_S), (\text{PC}_A), (\text{IC}_S), (\text{IC}_A), (\text{LB}), \big(\text{CP}^c\big).$$

A grand contract β is *feasible under centralization* if it satisfies all constraints in $\mathbf{P}^c$.

Delegation

Under delegation, P offers the grand contract to S only. Thus, the grand contract should respect S's limited budget constraint (LB) and participation constraint (PC_S). A grand contract is collusion-proof under delegation if S picks the null side contract among all side contracts that A prefers over the outside option, incentivize A to report θ truthfully, and respect S's limited budget. The constraint (CP^d) captures this formally:

$$\gamma_0 \in \arg\max_{\gamma} \mathbb{E}\big[t_S\big(\rho(\tilde{\theta}; \tilde{\sigma})\big) + \tau(\tilde{\theta}; \tilde{\sigma})\big] \text{ subject to } \big(\text{IC}_A^{\gamma}\big), \big(\text{LB}^{\gamma}\big)$$

$$t_A\big(\rho(\theta; \sigma)\big) - \tau(\theta; \sigma) - \theta X\big(\rho(\theta; \sigma)\big) \geq 0 \tag{$\text{PC}_A^{\gamma,d}$}$$

for all $\sigma \in \Sigma$, $\theta, \breve{\theta} \in \operatorname{Supp}(\tilde{\theta}|\sigma)$.[23] A side contract γ is feasible under delegation if it satisfies (IC_A^{γ}), (LB^{γ}), and ($\text{PC}_A^{\gamma,d}$). Principal P's problem under delegation is given by

$$\mathbf{P}^d : \max_{I,\beta} \mathbb{E}\big[vX(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) - t_S(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) - t_A(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})\big] \text{ subject to } (\text{PC}_S), (\text{LB}), \big(\text{CP}^d\big).$$

[22]By the principle of optimality, (CP^c) ensures that for all signal realizations $\sigma \in \Sigma$, $\rho_0(\theta; \sigma) = (\sigma, \sigma, \theta)$ and $\tau_0(\theta; \sigma) = 0$ are optimal for S given the posterior belief induced by σ.

[23]By the principle of optimality, (CP^d) ensures that for all signal realizations $\sigma \in \Sigma$, $\rho_0(\theta; \sigma) = (\sigma, \sigma, \theta)$ and $\tau_0(\theta; \sigma) = 0$ are optimal for S given the posterior belief induced by σ.

A grand contract β is *feasible under delegation* if it satisfies the constraints (PC_S), (LB), and (CP^d). The constraints (PC_A) and (IC_A) are not part of problem $\mathbf{P}^d$. Under delegation, A cannot directly agree to participate in the grand contract. Instead, A participates in the grand contract if he accepts the side contract, i.e., if the constraint ($\text{PC}_A^{\gamma,d}$) is satisfied. Thus, (PC_A) is not relevant under delegation. Similarly, A does not report to the grand mechanism under delegation. Therefore, the constraint (IC_A) is not needed.

Finally, I come back to my earlier observation that it is immaterial whether we allow P to make direct transfers to A under delegation.[24] Formally, this observation follows from a change of variable in program $\mathbf{P}^d$ by using $t_{PS} \equiv t_A + t_S$ and $t_{SA} \equiv t_A - \tau$ instead of the variables t_A and t_S. The transfers t_{PS} and t_{SA} can then be interpreted as transfers from P to S and S to A in a setting where P cannot make direct transfers to A.

Difference between centralization and delegation

The key difference between centralization and delegation lies in A's participation constraints ($\text{PC}_A^{\gamma,c}$) and ($\text{PC}_A^{\gamma,d}$) for the side contract. If A rejects the side contract under centralization, he still receives the payoff from the non-cooperative equilibrium of the grand contract. Under delegation, a rejection of the side contract implies that A receives his outside option.

By contrast, the additional constraints (PC_A) and (IC_A) do not harm P under centralization. To see this, note that A's participation constraint for the null side contract under delegation ($\text{PC}_A^{\gamma,c}$) is equivalent to (PC_A). Moreover, the incentive constraint for A under the null side contract ($\text{IC}_A^{\gamma_0}$) is equivalent to (IC_A).

Under centralization, P has one more instrument than under delegation. In particular, P can directly control A's payoff in the non-cooperative equilibrium of the grand contract. If P increases this payoff, the constraint ($\text{PC}_A^{\gamma,c}$) becomes tighter and the collusion-proofness constraint (CP^c) is relaxed. Thus, centralization is always at least as good as delegation. Moreover, delegation is optimal whenever P does not benefit from using A's rent as an instrument.

Benchmarks: Extreme information structures

Before I proceed with the analysis, I consider two benchmarks: the cases of uninformative and fully informative signal design. I show that P's optimal payoff is in both instances the monopsony payoff $\underline{W}$.

Independently of the information structure, P can always ignore S under centralization and guarantee a payoff of $\underline{W}$ by offering A the monopsony price p^*. Formally, this is equivalent to the grand contract[25]

$$\beta_{p^*} \equiv \big(X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}), t_S(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}), t_A(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta})\big) = \big(\mathbf{1}_{\hat{\theta} \le p^*}(\hat{\theta}), 0, p^* \cdot \mathbf{1}_{\hat{\theta} \le p^*}(\hat{\theta})\big).$$

Under the non-cooperative equilibrium of this grand contract, A accepts the price offer if and only if $\theta \le p^*$ and receives the payoff $\max\{p^* - \theta, 0\}$. Note that under any side

[24] Faure-Grimaud et al. (2003) and Celik (2009) do not allow such transfers under delegation.

[25] The indicator function $\mathbf{1}_A(x)$ satisfies $\mathbf{1}_A(x) = 1$ if x satisfies A and satisfies $\mathbf{1}_A(x) = 0$ otherwise.

contract $\gamma = (\rho, \tau)$, the total payoff of A and S is given by $X(\rho(\theta; \sigma))(p^* - \theta)$, which is weakly lower than A's payoff under the non-cooperative equilibrium. As any side contract needs to give A at least the non-cooperative payoff, S's payoff has to be weakly negative under collusion. It follows that collusion is not profitable for S. Principal P can, therefore, achieve at least the monopsony payoff under centralization.

With an uninformative signal, the monopsony payoff is also an upper bound for P's payoff. To see this, note that any uninformative signal is equivalent to the case $\Sigma = \{\sigma\}$. Thus, σ can be omitted as an argument in the grand contract and the side contract. The constraints (PC_A) and (IC_A) are, therefore, equivalent to

$$t_A(\theta) - \theta X(\theta) \geq \max\bigl\{0, t_A(\hat{\theta}) - \theta X(\hat{\theta})\bigr\}.$$

Maximizing P's expected payoff subject to these constraints is equivalent to the monopsony problem and results in a payoff of $\underline{W}$.

The monopsony payoff is also an upper bound for P if the signal is fully informative. In this case, the relationship between signals and types is bijective. Thus, we can again omit σ as an argument from the grand contract and the side contract. As we can ignore the constraint (IC_A^γ) under complete information, the maximal side transfer can be derived from ($\mathrm{PC}_A^{\gamma,c}$) as

$$\tau(\theta) = t_A\bigl(\rho(\theta)\bigr) - \theta X\bigl(\rho(\theta)\bigr) - \bigl(t_A(\theta) - \theta X(\theta)\bigr).$$

Using this and the fact that (CP^c) requires $t_S(\theta) \geq t_S(\rho(\theta)) + \tau(\theta)$, we have

$$t_S(\theta) + t_A(\theta) - \theta X(\theta) \geq t_S\bigl(\rho(\theta)\bigr) + t_A\bigl(\rho(\theta)\bigr) - \theta X\bigl(\rho(\theta)\bigr).$$

Let $t(\theta) \equiv t_S(\theta) + t_A(\theta)$ and $\rho(\theta) = \hat{\theta}$. Note that ($\mathrm{PC}_S$) and ($\mathrm{PC}_A$) imply $t(\theta) - \theta X(\theta) \geq 0$. Maximizing P's expected payoff subject to the resulting constraint

$$t(\theta) - \theta X(\theta) \geq \max\bigl\{0, t(\hat{\theta}) - \theta X(\hat{\theta})\bigr\}$$

is again equivalent to the monopsony problem and gives the payoff $\underline{W}$ to P.

Upper bounds on P*'s payoff for any information structure*

The constraints in P's optimization problems $\mathbf{P}^c$ and $\mathbf{P}^d$ simplify the structure of transfers in any feasible grand contract considerably. The simple structure of transfers allows me to derive upper bounds on P's payoff under centralization and delegation that hold for any combination of an information structure and a grand contract.

LEMMA 2. *For any feasible grand contract under centralization or delegation, there exist the functions* $t_A^0, t_A^1, t_S^0, t_S^1 : \Sigma \to \mathbb{R}$ *such that for all* $(\sigma, \theta) \in \mathrm{Supp}(\tilde{\sigma}, \tilde{\theta})$,

(i) $t_i(\sigma, \sigma, \theta) = X(\sigma, \sigma, \theta)t_i^1(\sigma) + (1 - X(\sigma, \sigma, \theta))t_i^0(\sigma)$ *with* $i \in \{A, S\}$

(ii) $t_A^0(\sigma) \geq 0$, $p(\sigma) \equiv t_A^1(\sigma) - t_A^0(\sigma) \geq \sup\{\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\}$

(iii) $t_S^0(\sigma) \geq -\ell$, $t_S^1(\sigma) \geq -\ell$, $t_S^j(\sigma) < 0 \Rightarrow t_S^{j'}(\sigma) \geq 0$ *for* $j, j' \in \{0, 1\}$ *and* $j \neq j'$

(iv) $t_S^j(\sigma) + t_A^j(\sigma) = t^j$ *for* $j \in \{0, 1\}$.

Any feasible grand contract under delegation furthermore satisfies, for all $\sigma \in \Sigma$,

(v) $t_S^1(\sigma) \geq 0$.

Result (i) states that the transfers to A and S in any feasible grand contract depend only on the signal realization and the production decision. The intuition behind this result is as follows. Agent A's transfer for a given signal realization has to be the same for all cost levels θ that lead to the same production decision. Otherwise, for any signal realization, A would report only the two cost levels $\hat{\theta}_1$ and $\hat{\theta}_0$ that maximize his transfer $t_A(\sigma, \sigma, \theta)$ among all cost levels that lead to production (for $\hat{\theta}_1$) or no production (for $\hat{\theta}_0$). Similarly, if S's transfer was different between two cost levels that lead to the same production decision, S and A could sign a profitable side contract in which, for a given signal realization, they would report only the cost levels that maximize the transfer $t_S(\sigma, \sigma, \theta)$ for the two cases of production and no production.

Result (ii) makes the following point. Decompose A's transfer into a *fixum* $t_A^0(\sigma)$ and a *price* $p(\sigma)$ paid for production. For any feasible grand contract, the fixum has to be positive, as A perfectly anticipates the production decision before deciding whether to participate in the grand contract. Moreover, the price $p(\sigma)$ always needs to cover A's production costs as A would otherwise overstate the costs to avoid production.

Result (iii) states first that S's transfer to P can never exceed the bound ℓ—a direct implication of S's limited budget. Second, point (iii) notes that S can never make a loss both with and without production, as this would obviously violate S's participation constraint (PC_S).

Result (iv) states that P can make the total transfer to S and A contingent only on whether the good is produced. Whenever the total transfer to S and A does not depend only on the production decision, S and A can sign a side contract, which leads to the same production decision as without collusion, but coordinates their messages to P such that the highest total transfer to the colluding parties is generated. As the production decision remains unchanged, such a side contract is always feasible. By fixing the total transfer, P induces a strong conflict of interest between S and A, as any increase in the transfer of one of the parties has to decrease the transfer of the other party by the same amount.

Finally, result (v) holds, as S can avoid production under delegation by offering a side contract that always sends a message that induces no production while no side transfers are exchanged. This side contract replicates A's outside option. Hence, A is willing to accept. Thus, P has to reward S for production under delegation.

The results of Lemma 2 can be used to derive upper bounds on P's payoff under delegation and centralization. Let x be the (ex ante) probability of production induced by some grand contract, i.e.,

$$x \equiv \mathbb{E}\big[X(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})\big],$$

and let $\overline{\theta}(x)$ be the x-quantile of the distribution F, formally defined as[26]

$$\overline{\theta}(x) \equiv \min\{\theta \in \Theta : F(\theta) \geq x\}.$$

The maximal social surplus under a grand contract with production probability x is

$$B_1(x) \equiv \int_{\underline{\theta}}^{\overline{\theta}(x)} (v - \theta)\, dF(\theta),$$

as it is cost-minimizing to let only the types below the x-quantile produce. To state the following lemma concisely, I define[27]

$$B_2(x) \equiv x\big(v - \overline{\theta}(x)\big) + (1 - x)\ell \quad \text{and} \quad B_3(x) \equiv x\big(v - \overline{\theta}(x)\big) + \int_{\overline{\theta}(x)-\ell}^{\overline{\theta}(x)} F(\theta)\, d\theta.$$

Lemma 3. *Under delegation, P's payoff from any information structure and any feasible grand contract with production probability x does not exceed*

$$B_d(x) \equiv \min\{B_1(x), B_2(x)\}.$$

Under centralization, P's payoff from any information structure and any feasible grand contract with production probability x does not exceed

$$B_c(x) \equiv \min\{B_1(x), \max\{B_2(x), B_3(x)\}\}.$$

As the participation of A and S in the grand contract is voluntary, P's expected payoff from any grand contract with probability of production x cannot exceed the maximal social surplus $B_1(x)$.

Next, I explain how the bound $B_2(x)$ arises. Due to Lemma 2, P pays expected total transfers of $\mathbb{E}[t_A(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) + t_S(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})] = xt^1 + (1-x)t^0$ under any feasible grand contract with production probability x. Moreover, Lemma 2 implies that S can never incur a loss with production under delegation. As this also has to hold for some signal realization $\sigma' \in \Sigma$ after which A produces for a cost weakly above $\overline{\theta}(x)$ and A can always guarantee himself a positive payoff, it follows that $t^1 = t_S^1(\sigma') + t_A^1(\sigma') \geq \overline{\theta}(x)$. Without production, S's loss is at most ℓ. Thus, $t^0 \geq -\ell$. Hence, P's payoff under delegation satisfies $x(v - t^1) - (1-x)t^0 \leq x(v - \overline{\theta}(x)) + (1-x)\ell = B_2(x)$.

While P cannot impose a loss on S with production under delegation, P may do so under centralization by setting $t_S^1(\sigma) < 0$ for some $\sigma \in \Sigma$. In the proof of the lemma, I show that P's payoff from imposing a loss on S with production under a grand contract with production probability x cannot exceed the bound $B_3(x)$. Using integration by parts, this bound can be rewritten as

$$B_3(x) = \int_{\underline{\theta}}^{\overline{\theta}(x)} (v - \theta)\, dF(\theta) - \int_{\underline{\theta}}^{\overline{\theta}(x)-\ell} \big(\overline{\theta}(x) - \ell - \theta\big)\, dF(\theta).$$

[26] If F has a strictly positive density $f(\theta) = F'(\theta) > 0$, then F is invertible and $\overline{\theta}(x) = F^{-1}(x)$.

[27] I use the convention $\int_a^b z(x)\, dx = -\int_b^a z(x)\, dx$.

Based on this reformulation, the bound $B_3(x)$ can be interpreted as the difference between the maximal social surplus generated under a grand contract with production probability x and rent payments of $\overline{\theta} - \ell - \theta$ for each type $\theta \leq \overline{\theta} - \ell$.

I explain in three steps why P cannot avoid these rent payments. First, for any grand contract with an ex ante probability of production x, the highest producing type lies weakly above $\overline{\theta}(x)$. Thus, there exists a signal realization $\sigma \in \Sigma$ such that $p(\sigma) \geq \overline{\theta}(x)$. As S can at most incur a loss of ℓ, the total payment to A and S with production satisfies $t^1 \geq \overline{\theta}(x) - \ell$. Second, if S makes a loss in the case of production, P cannot also impose a loss on S without production, as this would violate S's participation constraint. As A always receives weakly positive payments, this implies that the total payment to A and S without production satisfies $t^0 \geq 0$. Third, consider the cost level θ with $\theta \leq \overline{\theta}(x) - \ell$. If A produces for this cost level, A and S receive a joint rent of $t^1 - \theta \geq \overline{\theta}(x) - \ell - \theta$. This rent cannot be extracted, as the joint payment to A and S without production t^0 is weakly positive.

5. Optimal delegation

In this section, I present P's optimal combination of information structure and grand contract under delegation. In particular, I construct an information structure and a feasible grand contract under delegation that allow P to implement any ex ante probability of production x with an expected payoff equal to the upper bound $B_d(x)$. The optimal combination of information structure and grand contract under delegation implements the production probability x_d at which the function $B_d(x)$ attains its maximum.[28]

The key idea behind the optimal combination of information structure and grand contract under delegation is as follows. The signal provides enough information to allow S to extract all rents from A in the case of production. At the same time, the signal maintains S's uncertainty about the production decision and allows P to extract rents from S in the case of no production.

Next, I describe a combination of an information structure and a grand contract that implements a production probability x and leads to a payoff of $B_d(x)$ for P. Figure 3 depicts a *weighted information structure*. The signal space of the weighted information structure is the interval of types below the x-quantile, i.e., $\Sigma = [\underline{\theta}, \overline{\theta}(x)]$. If A's cost satisfies $\theta \leq \overline{\theta}(x)$, the signal realization $\sigma = \theta$ is drawn. If $\theta > \overline{\theta}(x)$, some $\sigma \in \Sigma$ is drawn according to the density function $w(\sigma)$. I refer to $w(\sigma)$ as a weighting function, as it determines the relative weights with which the types above $\overline{\theta}(x)$ are pooled as noise into the signal realizations in Σ. The weighting function $w(\cdot)$ characterizes the weighted

[28] Both functions $B_d(x)$ and $B_c(x)$ are upper semicontinuous and therefore attain a maximum on $[0, 1]$. This follows from $B_1(x)$ being continuous and $B_2(x)$ and $B_3(x)$ being upper semicontinuous as well as the fact that both the minimum as well as the maximum of two upper semicontinuous functions is also upper semicontinuous.

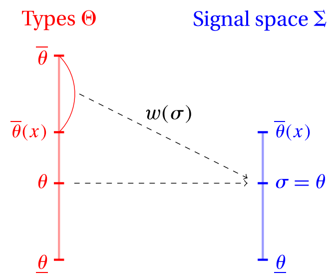

Figure 3. Weighted information structure. A weighted information structure has the signal space $[\underline{\theta}, \overline{\theta}(x)]$. If $\theta \leq \overline{\theta}(x)$, then the signal realization $\sigma = \theta$ is generated. If $\theta > \overline{\theta}(x)$, some $\sigma \in [\underline{\theta}, \overline{\theta}(x)]$ is drawn from the density $w(\sigma)$.

information structure I_w and induces, for any signal realization $\sigma \in \Sigma$, a conditional cumulative distribution function (cdf)[29]

$$G(\theta|\sigma) = \begin{cases} 0 & \text{if } \theta < \sigma, \\ \dfrac{f(\sigma)}{f(\sigma) + (1-x)w(\sigma)} & \text{if } \theta \in [\sigma, \overline{\theta}(x)], \\ \dfrac{f(\sigma)}{f(\sigma) + (1-x)w(\sigma)} + \dfrac{(1-x)w(\sigma)}{f(\sigma) + (1-x)w(\sigma)} \dfrac{F(\theta) - F(\overline{\theta}(x))}{1 - F(\overline{\theta}(x))} & \text{if } \theta > \overline{\theta}(x). \end{cases}$$

I combine a weighted information structure with a grand contract under delegation that induces production if all reports coincide and lie below the threshold $\overline{\theta}(x)$, i.e.,

$$X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = \mathbf{1}_{\hat{\sigma}_S = \hat{\sigma}_A = \hat{\theta} \leq \overline{\theta}(x)}(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}). \tag{1}$$

Under truthful reporting, the project is, therefore, realized whenever costs take the lowest possible value given the signal realization. Furthermore, the transfer to A covers the exact cost of production:

$$t_A(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = \hat{\theta} X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}). \tag{2}$$

As the grand contract has to be feasible, Lemma 2 implies that the transfer to S is

$$t_S(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta})(t^1 - \hat{\theta}) + (1 - X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}))t^0. \tag{3}$$

Proposition 1. *Under delegation, P can reach the payoff $B_d(x)$ for any probability of production $x \in [0, 1]$ through the weighted information structure I_{w^x} with*

$$w^x(\sigma) = \frac{f(\sigma)[\min\{\overline{\theta}(x), \check{\theta}(x)\} - \sigma]_+}{\displaystyle\int_{\underline{\theta}}^{\min\{\overline{\theta}(x), \check{\theta}(x)\}} F(\sigma')\, d\sigma'}$$

[29]The marginal cdf over signal realizations is given by $H(\sigma) = \int_{\underline{\theta}}^{\sigma} (f(\sigma') + (1-x)w(\sigma'))\, d\sigma'$.

and the grand contract β^x defined by (1), (2), and (3) with

$$t^0 = -\min\left\{\frac{\int_{\underline{\theta}}^{\overline{\theta}(x)} F(\sigma)\,d\sigma}{1-x}, \ell\right\} \quad \textit{and} \quad t^1 = \overline{\theta}(x),$$

where $\check{\theta}(x)$ is defined as the unique solution to

$$\int_{\underline{\theta}}^{\theta} F(\sigma)\,d\sigma = (1-x)\big(\overline{\theta}(x) - \theta + \ell\big). \tag{4}$$

The combination of information structure I_{w^x} and grand contract β^x optimally resolves P's trade-off between information elicitation and collusion prevention under delegation. Supervisor S's signal is very informative on A's costs in case the good is produced. This allows S to extract all rents from A by setting a price that equals A's cost under production. Supervisor S earns the *bonus* $t^1 - \sigma$ for production after the signal realization $\sigma \in \Sigma$. At the same time, the probability of a bonus payment $G(\sigma|\sigma)$ decreases in the bonus and enables P to extract S's rent by a payment in the case of no production. In contrast to the exact information on producing types, the optimal signal reveals no information about the relative likelihood of high-cost types. As the type of A does not influence his payoff without production, P does not benefit from information about nonproducing types.

I now explain Proposition 1 in more detail. I start by arguing that the null side contract is feasible given the information structure and the grand contract in Proposition 1. From A's perspective, any grand contract that satisfies (1) and (2) is equivalent to a price offer equal to the signal realization. Thus, under the null side contract, A's participation constraint ($\mathrm{PC}_A^{\gamma_0,d}$) and incentive compatibility constraint ($\mathrm{IC}_A^{\gamma_0}$) are satisfied and A never receives a positive rent.

Next, I show that S optimally responds to the information structure and the grand contract in Proposition 1 by offering the null side contract. Given some weighted information structure I_w, a grand contract that satisfies (1)–(3), and the null side contract, S receives an expected payoff after the signal realization σ of

$$\Pr(\tilde{\theta} = \sigma|\tilde{\sigma} = \sigma)\big(t^1 - \sigma\big) + \Pr(\tilde{\theta} > \sigma|\tilde{\sigma} = \sigma)t^0 = \frac{f(\sigma)\big(t^1 - \sigma\big) + (1-x)w(\sigma)t^0}{f(\sigma) + (1-x)w(\sigma)}.$$

For $t^1 = \overline{\theta}(x)$ and $t^0 \leq 0$, S does not benefit from offering a side contract under which A never produces, as S would then receive the negative payoff t^0. Supervisor S would not benefit from a side contract under which some types in the interval $(\overline{\theta}(x), \overline{\theta}]$ produce. This would require S to pay A a side transfer τ with $\sigma + \tau > \overline{\theta}(x)$. Supervisor S's payoff under production would then be negative, as $t^1 - \sigma - \tau = \overline{\theta}(x) - \sigma - \tau < 0$. Together with $t^0 \leq 0$, this implies that S's payoff under such a deviation cannot exceed 0.

Principal P optimally designs the weighting function to minimize the transfer t^0 while satisfying S's participation constraint. Note that (PC_S) holds for some weighted information structure and grand contract satisfying (1)–(3) if

$$t^0 = -\min_{\{\sigma\in\Sigma: w(\sigma)>0\}}\left\{\frac{f(\sigma)\big(\overline{\theta}(x)-\sigma\big)}{(1-x)w(\sigma)}\right\}.$$

As S's participation constraint is binding after the worst signal realization from S's perspective, P would like to make the worst signal as good as possible. In technical terms, P faces a max-min problem. Ideally, P would like to choose $w(\cdot)$ and t^0 such that S's expected payoff is 0 for all signal realizations. A weighting function that renders S's expected payoff constant across all signal realizations is

$$w(\sigma) = \frac{f(\sigma)\big(\overline{\theta}(x)-\sigma\big)}{(1-x)c},$$

where the constant $c \in \mathbb{R}_+$, pinned down by $\int_\Sigma w(\sigma)\,d\sigma = 1$, is given by

$$c = \frac{\displaystyle\int_{\underline{\theta}}^{\overline{\theta}(x)} F(\sigma)\,d\sigma}{1-x}.$$

If $c \leq \ell$, it is feasible to extract all rents from S by setting $t^0 = -c$. In this case, P's expected payoff equals the maximal social surplus under the ex ante probability x, i.e., $B_1(x)$. If $c > \ell$, P cannot extract the full social surplus. However, as I show in the proof of Proposition 1, P can reach an expected payoff of $B_2(x)$ by setting $t^0 = -\ell$ and $w(\sigma) = 0$ for signal realizations σ between $\check{\theta}(x)$ and the threshold $\overline{\theta}(x)$, and $w(\sigma)$ such that the expected payoff of S is constant for the remaining signal realizations.[30]

Proposition 1 implies three important results as immediate corollaries. First, the optimal combination of information structure and grand contract under delegation follows directly from the proposition.

Corollary 1. *An optimal combination of information structure and grand contract under delegation is given by* $(I_{w^{x_d}}, \beta^{x_d})$ *with* $x_d \in \arg\max_{x\in[0,1]} B_d(x)$.

Under the optimal grand contract with delegation, S receives a positive transfer from P with production and makes a payment to P without production. Note that this feature is shared with the optimal grand contract in Faure-Grimaud et al. (2003).

Second, Proposition 1 allows us to derive necessary and sufficient conditions for partial information revelation to be optimal under both delegation and centralization.

Corollary 2. *Under both centralization and delegation, an optimal information structure partially informs S about A's cost* θ *if and only if* $x_d \in (F(\underline{\theta}), 1)$ *and* $\ell > 0$.

Recall from the analysis of the benchmarks that P can achieve at most a payoff of $\underline{W}$ under either an uninformative or a completely informative signal. Note that

[30]The uniqueness of $\check{\theta}(x)$ follows from the intermediate value theorem. See the proof for details.

$\underline{W} = \max_{x\in[0,1]} x(v - \overline{\theta}(x))$ and that $B_d(x)$ can be expressed as

$$B_d(x) = x\bigl(v - \overline{\theta}(x)\bigr) + \min\left\{\int_{\underline{\theta}}^{\overline{\theta}(x)} F(\theta)\,d\theta, (1-x)\ell\right\}.$$

Thus, $x_d \in (F(\underline{\theta}), 1)$ and $\ell > 0$ is equivalent to $\max_{x\in[0,1]} B_d(x) > \underline{W}$. Under mild conditions, P prefers a partially informative signal under delegation to either of the extreme information structures under centralization. As $B_c(x) = x(v - \overline{\theta}(x))$ for either $x \in \{F(\underline{\theta}), 1\}$ or $\ell = 0$, the same mild conditions are necessary and sufficient for the optimality of a partially informative signal under centralization. Without S, P can at most achieve the payoff $\underline{W}$. Thus, the corollary also provides necessary and sufficient conditions for P to benefit from hiring S under both centralization and delegation.

Third, P can extract the full surplus whenever S's budget is large enough.

Corollary 3. *If S's budget is large enough, P can extract the full surplus under delegation: If $v < \overline{\theta}$ and $\ell \geq \overline{\ell}(v) \equiv \overline{W}/(1 - F(v))$, then* $\max_{x\in[0,1]} B_d(x) = \overline{W}$. *If $v \geq \overline{\theta}$, then* $\lim_{\ell\to\infty} \max_{x\in[0,1]} B_d(x) = \overline{W}$.

The full surplus is generated through the efficient production cutoff $\overline{\theta}(x) = v$. Under efficient production, S becomes the residual claimant of all surplus in the case of production. If $v < \overline{\theta}$, P can extract the full surplus through the payment t^0 whenever

$$\ell \geq \frac{\displaystyle\int_{\underline{\theta}}^{v} (v-\theta)\,dF(\theta)}{1 - F(v)} = \frac{\overline{W}}{1 - F(v)} = \overline{\ell}(v).$$

Thus, P can achieve the full surplus under delegation only if S can absorb at least a loss of $\overline{\ell}(v)$, which is a multiple of the full surplus with a factor of $1/(1 - F(v))$. If $v \geq \overline{\theta}$, P can set a cutoff just below $\overline{\theta}$. As ℓ grows large, the production probability x can be set arbitrarily close to 1, resulting in a payoff of almost $\overline{W}$.

Faure-Grimaud et al. (2003) note that if S is risk-neutral, P can achieve the same profit as in the case where P can observe S's signal. Corollary 3 is reminiscent of this result as a sufficiently large budget for S may be viewed as an analogue to the risk neutrality of S.

Celik (2009) points out that the suboptimality of delegation in his model is linked to an informational *double marginalization* problem that arises under delegation. With information control, P avoids the problem of double marginalization by giving S enough information about A to extract A's rent completely. Principal P uses S's uncertainty over the production decision to extract rents from S and may even achieve the first-best surplus under delegation.

6. When is delegation (sub)optimal?

In this section, I provide conditions under which delegation is optimal. In particular, I show that delegation is optimal if either S's budget is large or the value of the good is small. Furthermore, I show that delegation is suboptimal if the value of the good is high and S's budget is sufficiently strict. I denote the median cost type by $m \equiv \overline{\theta}(1/2)$.

Proposition 2. *Delegation is optimal if either*

(i) S's budget is large, i.e., $\ell \geq \hat{\ell}(v)$*, where* $\hat{\ell} : [0, \infty) \to [0, \overline{\ell}(v))$ *or*

(ii) the value of the good is small, i.e., $v < m$*.*

First, I explain why delegation is optimal if S's budget is large. I start with the following observation.

Lemma 4. *If delegation is suboptimal, P cannot extract the full surplus.*

Recall that the central difference between delegation and centralization is A's outside option when bargaining over the side contract. Under centralization, A can reject the side contract and participate in the grand contract non-cooperatively. Under delegation, A can only participate in the grand contract by accepting the side contract. If centralization is superior to delegation, P finds it optimal to pay a positive rent to A so as to limit the possibility of S finding a profitable side contract. As A's rent is strictly positive, P cannot extract the full surplus.

We know from Corollary 3 that P can (approximately) extract the full surplus under delegation if S's budget is sufficiently large. Moreover, P's optimal payoff under delegation is continuously increasing in ℓ as the bound $B_d(x)$ is continuously increasing in ℓ. Thus, there exists a threshold $\hat{\ell}$ below the threshold $\overline{\ell}$, such that delegation is optimal if $\ell \geq \hat{\ell}$.

Next, I explain why delegation is optimal if the value of the good lies below the median cost type. Corollary 2 shows that P only benefits from the presence of S if S can incur a loss in some states of the world. Under centralization, P can impose a loss on S in the case of either no production or production. Supervisor S's participation constraint prevents P from doing both. Under delegation, P can only impose a loss on S without production. As the maximal loss to S is limited by ℓ, P can reduce expected rent payments through imposing losses under a grand contract with production probability x by at most $x\ell$ if the loss is incurred under production and $(1-x)\ell$ if the loss is incurred without production. For $x < 1/2$, P can reduce rents by more if S incurs a loss without production. As $x < 1/2$ is optimal if $v < m$, delegation is optimal if $v < m$.

Reversing the argument above, centralization may be better than delegation if P wants to implement a high production probability and S's budget is small. Indeed, if $x > 1/2$, P may reduce S's rent up to $x\ell$ if S incurs a loss with production and only up to $(1-x)\ell$ for a loss without production. However, this argument ignores that P has to pay a positive rent to A if S incurs a loss under production as A would otherwise accept any side contract that avoids production. The next proposition shows that P can nevertheless benefit from centralization.

Proposition 3. *Suppose* $f(\theta) = F'(\theta) > 0$ *for all* $\theta \in \Theta$*. Delegation is suboptimal if*

(i) the value of the good is high, i.e., $p^*(v) \in (m, \overline{\theta})$ *and*

(ii) S's budget is sufficiently small, i.e., $\ell \in (0, \varepsilon)$ *for some* $\varepsilon > 0$*.*

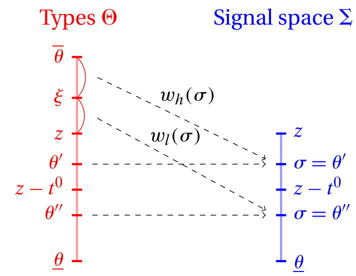

FIGURE 4. Information structure I^c under centralization.

In particular, there exists a combination of an information structure and a feasible grand contract under centralization (I^c, β^c) *for which P's expected payoff exceeds* $\max_x B_d(x)$ *and first-order approximates* $\max_x B_c(x)$ *if (i) and (ii) are satisfied.*

I now describe the information structure I^c and the grand contract β^c, which are formally defined in the proof of the proposition.

Figure 4 depicts the information structure I^c. Given the cutoff cost $z \in \Theta$, the signal space Σ consists of the interval $[\underline{\theta}, z]$. A type $\theta \leq z$ generates the signal realization $\sigma = \theta$. The types $[z, \overline{\theta}]$ are separated into two intervals by the type $\xi > z$. A type in the interval $[\xi, \overline{\theta}]$ generates signal realizations in the interval $\Sigma_h = [z - t^0, z]$ according to the density $w_h(\sigma)$. A type in the interval $[z, \xi)$ generates signal realizations in $\Sigma_l = [\underline{\theta}, z - t^0]$ according to the density $w_l(\sigma)$ with $t^0 > 0$. Note that I^c is a weighted information structure for $\xi \in \{z, \overline{\theta}\}$.

Under the grand contract, A produces β^c for $\theta \leq z$. For a signal realization σ, the transfers are given by $t_A^1(\sigma) = \sigma + t^0$, $t_A^0(\sigma) = 0$, $t_S^1(\sigma) = z - \ell - \sigma$, and $t_S^0(\sigma) = t^0$. Thus, A is effectively offered the price $\sigma + t^0$ and earns a margin of t^0 with production. Nevertheless, only types below z accept the price, as I^c pools high producing types with high nonproducing types and low producing types with low nonproducing types. Supervisor S incurs a loss with production for high signals in $[z - \ell, z]$, but receives a positive transfer without production. In the proof of the lemma, I specify the parameters z, t^0, ξ, $w_h(\sigma)$, and $w_l(\sigma)$ such that I^c is well defined and β^c is feasible under centralization for a small ℓ. In particular, S does not benefit from a collusive agreement under which A is induced to not produce. For any σ, this would require S to pay t^0 to A for any cost level $\theta \in \text{Supp}(\tilde{\theta}|\sigma)$, not only for those types that would produce under the null side contract. The functions $w_h(\sigma)$ and $w_l(\sigma)$ pool sufficiently many nonproducing types into the signal realizations to make such a collusive agreement unattractive for S.

For a small ℓ, I^c and β^c allow P to reduce the total rent of S and A by approximately $x\ell$ with respect to the monopsony case. This implies that P achieves a higher payoff than under delegation if $x > 1/2$. As $x > 1/2$ is optimal if the monopsony price $p^*(v)$ exceeds the median, delegation is suboptimal in this case. Moreover, as $B_3(x) \simeq x(v - \overline{\theta}(x)) + x\ell$ for a small ℓ, (I^c, β^c) is near optimal for a small ℓ and a high v.

There are two notable differences between (I^c, β^c) and the optimum under delegation (I^{x_d}, β^{x_d}). First, there is a difference between the grand contracts, which is in line with the previous discussion. Under β^c, S incurs a loss with production for the types

$\theta \in [z - \ell, z]$ and makes a gain without production. This is reminiscent of the optimal grand contract under centralization in Celik (2009), where S also only incurs a loss after a low-cost realization.In contrast, under the grand contract β^{x_d}, S incurs only a loss without production. Second, whereas I^{x_d} pools all nonproducing types into the same signal realizations, there are two separate intervals of nonproducing types under I^c. This is necessary as, under β^c, A receives a price that exceeds the cost θ by the fixed margin t^0. To avoid A producing for some θ above the cutoff z, I^c needs to ensure that for any signal realization, the distance between producing and nonproducing types exceeds the margin t^0. Nevertheless, I^c is very similar to a weighted information structure. In particular, there is exactly one producing type in the support of each signal realization.

While I identify the near-optimal combination (I^c, β^c) for a small ℓ, I was unable to generally characterize optimal combinations of information structure and grand contract if delegation is suboptimal. If delegation is suboptimal, P prefers to pay a positive rent to A. The key challenge toward a complete characterization of optimal combinations of information structure and grand contract lies in finding the optimal allocation of rents across A and S. This problem is considerably more complicated than under delegation where all rents are extracted from A.

7. Extensions

In this section, I discuss several extensions. Formal statements and proofs of all the results can be found in the last working paper version of this paper (Asseyer 2020).

7.1 *Alternative bargaining protocols*

In the previous sections, the side contract is set by S through a take-it-or-leave-it offer to A. In the context of public procurement, this assumption captures the situation in which a (corrupt) procurement officer has all the bargaining power vis-à-vis the private supplier.

The results of the previous sections also hold under alternative bargaining protocols and different distributions of bargaining power between the supervisor and agent.[31] Following the approach of Laffont and Martimort (1997), one may introduce a third party T that offers the side contract to S and A. The framework with T encompasses and generalizes the setting studied in the previous sections. If T has the same preferences as S, it is as if S offers the side contract herself. If T has the same preferences as A, we are in the polar setting in which A has all the bargaining power. Without specifying the preferences of the third party, assume that T offers a side contract that is Pareto efficient for the colluding coalition.[32] For a given grand β, there is typically a set of Pareto efficient side contracts. Suppose P takes a cautious approach and evaluates the grand contract β by the smallest payoff that may arise for any feasible and Pareto efficient side contract.

[31] This finding is consistent with the robustness results of Faure-Grimaud et al. (2003) and Mookherjee et al. (2020b), who also consider this extension.

[32] This approach follows Che and Kim (2006).

Propositions 1–3 extend to the setting with T. In the model with T, P faces a set of feasible and Pareto efficient side contracts that T may choose in response to some grand contract. The null side contract played in the equilibrium of the model without T lies in this set, as it is feasible and Pareto efficient. As P expects the worst side contract to be offered, the payoff from any grand contract is weakly lower in the model with T than in the model without T. Thus, the payoff bounds $\max_x B_d(x)$ and $\max_x B_c(x)$ remain valid in the model with T. Moreover, one may show that the optimal combination of information structure and grand contract under delegation $(I_{w^{x_d}}, \beta^{x_d})$ and the near-optimal combination under centralization (I^c, β^c) deliver the same payoff in both models. Propositions 1–3 follow from this observation. Under $(I_{w^{x_d}}, \beta^{x_d})$ and (I^c, β^c), a side contract is feasible only if it induces the same production decision as the null side contract.[33] As both grand contracts fix the total payment from P to A and S conditional on the production decision, P is only hurt by a side contract if it induces a different production decision than the null side contract. Thus, the preferences of the player who proposes the side contract does not matter, and the models with and without T deliver the same results.

7.2 *Divisible output*

Delegation performs well under information control even if output is divisible. In particular, suppose A can produce any quantity $q \in Q \subset \mathbb{R}_+$ with $0 \in Q$ and $\max Q < \infty$ at a cost $c(\theta, q)$. The cost function $c(\theta, q)$ is increasing in both arguments, is supermodular, and satisfies $c(\theta, 0) = 0$ for all $\theta \in \Theta$. Principal P's gross utility is given by $v(q)$ with $v(0) = 0$.

Principal P can extract the full surplus with divisible output if S has an unlimited budget and the smallest strictly positive quantity is sufficiently far away from 0. Principal P achieves this in a very similar fashion to the case of indivisible output. Let $q^*(\theta) = \arg\max_q v(q) - c(\theta, q)$ be the efficient production rule. Principal P optimally uses a weighted information structure with a signal space $\Sigma = [\underline{\theta}, \theta_0]$, where θ_0 is the highest type with strictly positive production. Any type $\theta \in [\underline{\theta}, \theta_0]$ generates the signal realization $\sigma = \theta$. Any type $\theta > \theta_0$ generates some signal in Σ according to a density $w(\sigma)$. Principal P optimally combines this information structure with a grand contract under which S is the residual claimant if production is strictly positive and has to make a payment to P if production is 0. In particular, S receives a payment of $\phi^*(\theta) \equiv v(q^*(\theta)) - c(\theta, q^*(\theta))$ if $\theta \leq \theta_0$ and pays $\overline{W}/(1 - F(\theta_0))$ to P if $\theta > \theta_0$, where $\overline{W}$ is the full surplus.

This combination of information structure and grand contract extracts the full surplus if the smallest strictly positive quantity $q^*(\theta_0)$ is sufficiently far away from 0. As S receives the full surplus under production, S has no incentive to collude and distort the production decision if costs are below θ_0. However, S wishes to avoid the payment to P if costs are above θ_0. If the type $\theta > \theta_0$ produces a strictly positive quantity q for some signal realization σ, the type $\theta = \sigma$ receives a rent of at least $c(\theta_0, q) - c(\sigma, q)$. If this

[33]This was formally shown in the proofs of Propositions 1 and 3.

rent exceeds the social surplus $\phi^*(\sigma)$, collusion is not profitable, as S's payoff is negative with and without production. This constraint is hardest to satisfy for the lowest signal realization $\sigma = \underline{\theta}$ and the smallest positive quantity $q^*(\theta_0)$. Thus, P can extract the full surplus if $\phi^*(\underline{\theta}) \leq c(\theta_0, q^*(\theta_0)) - c(\underline{\theta}, q^*(\theta_0))$. This condition holds if $q^*(\theta_0)$ is sufficiently far away from 0. Such a situation may arise if A's cost function exhibits fixed costs or P's gross utility $v(q)$ is negative for small q. For a specific example, suppose the construction company A could construct an airport of size q for the municipality P. It is then likely that A would incur some planning costs independently of the size of the airport and that the airport would only be economically viable for P if it exceeded a certain minimal size.

7.3 *Does information control make delegation optimal?*

Consider a principal who gains information control over a supervisor. When should we expect the principal to change from centralization to delegation after obtaining information control? Put differently, under which conditions is delegation optimal with and suboptimal without information control? The previous sections analyze the first part of this question. Here, I discuss the second part.

Delegation is suboptimal for some exogenously given information structure if there exists a worst signal realization for S. An information structure has a worst signal realization $\overline{\sigma}$ if, for any feasible grand contract, S's payoff with production and the probability of production are lower for $\overline{\sigma}$ than for any other signal realization. Supervisor S's participation constraint is most restrictive for $\overline{\sigma}$. Moreover, the option to report $\overline{\sigma}$ severely limits P's ability to extract rents from S for all other signal realizations. This effect is so strong that P prefers to contract directly with A and to ignore the additional information S might provide.

A worst signal realization exists for many information structures. Two prominent examples are information structures with additive noise and partitional information structures. Under an information structure with additive noise, A's cost is given by $\tilde{\theta} = \tilde{\sigma} + \tilde{\varepsilon}$, where $\tilde{\sigma}$ is S's signal and $\tilde{\varepsilon}$ is an independently distributed noise term. Under a partitional information structure, each signal $\sigma \in \Sigma$ is associated with exactly one element of a partition of Θ. It turns out that information structures with additive noise and partitional information structures have a worst signal realization under relatively mild assumptions. This suggests that the introduction of information control might often be followed by a change from a centralized to a decentralized organizational structure.[34]

7.4 *Ex ante collusion*

Throughout the previous sections, I follow Faure-Grimaud et al. (2003) and Celik (2009), and assume that A and S cannot collude on their participation decisions regarding the grand contract. Mookherjee et al. (2020b) allow for this form of ex ante collusion by

[34] In the model, centralization is always at least as good as delegation. However, if there is a cost of communication for the principal, delegation is strictly optimal under the appropriate conditions.

requiring that any grand contract includes the exit message e that results in no production and no transfers to P. I discuss some of the implications of ex ante collusion for my analysis.

I start with the following key observation. A feasible grand contract is ex ante collusion-proof if and only if $t^0 \geq 0$. Without the exit message e, the colluding coalition can either coordinate on a message that induces production and a total payment of t^1 from P or send a message that induces no production and a total payment of t^0 from P. If the grand contract includes the exit message e, the coalition has the additional option of inducing no production and a total payment of zero from P. As the payoffs of A and S are common knowledge without production, the coalition benefits from the message e if $t^0 < 0$. If $t^0 \geq 0$, the coalition never gains from sending the message e.

It follows that P's payoff under delegation and ex ante collusion lies weakly below the monospony payoff. Under ex ante collusion, P cannot impose losses on S in the case of no production. Under delegation, P cannot impose losses on S with production. As hiring S is only profitable for P if S can incur some loss due to Corollary 2, P does not benefit from the presence of S under delegation and ex ante collusion.[35]

Moreover, given an optimal combination of information structure and grand contract without ex ante collusion, the optimal grand contract is ex ante collusion-proof if and only if delegation is strictly suboptimal. If the optimal grand contract is ex ante collusion-proof, P cannot impose losses on S in the case of no production. So as to benefit from S, P has to impose a loss on S in the case of production. This is only possible with centralization. Thus, delegation is suboptimal. Conversely, if delegation is suboptimal, P imposes a loss on S in the case of production for at least one signal realization. It follows that P cannot impose a loss on S in the case of no production to satisfy S's participation constraint. As S never makes a loss in the case of no production, the exit message e is not beneficial for the colluding parties.

Finally, the combination of information structure and grand contract (I^c, β^c) from Proposition 3 is near optimal with ex ante collusion. In particular, it can be shown that $\max_x B_3(x)$ is a bound on P's payoff under ex ante collusion and that (I^c, β^c) first-order approximates this bound for a small ℓ.

8. Conclusion

I consider a principal–supervisor–agent relationship in which the supervisor and agent can collude and the principal designs the supervisor's signal of the agent's private type. I study how the principal optimally uses information control to fight collusion and whether the principal should delegate the design of the agent's contract to the supervisor. The principal optimally chooses a partially informative signal and can extract the full surplus if the supervisor's budget is large enough. Delegation is an optimal response to collusion under information control if either the supervisor's budget is large or the principal's value from production is small. However, delegation is suboptimal if the agent's production is of high value to the principal and the supervisor's budget is sufficiently small.

[35] This observation is very similar to Proposition 2 in Mookherjee et al. (2020b).

I see two directions for further research. First, this paper studies the joint design of mechanisms and information under collusion with a focus on collusive supervision. It seems worthwhile to explore whether the insights extend to more general mechanism design environments. Second, this paper closely follows the literature on collusive supervision with exogenous information in assuming that the supervisor's signal is also observed by the agent. If the supervisor's budget is large, the principal does not benefit from inducing additional asymmetric information in the colluding coalition through the provision of private information to the supervisor. However, private information on the supervisor's side may be beneficial if the supervisor's budget is small. I leave these questions for future research.

Appendix: Omitted proofs

Proof of Lemma 1: Collusion-proofness principle

I prove that S has an optimal, direct, and truthful side contract for any β, I, and $\sigma \in \Sigma$. Given β, I, and $\sigma \in \Sigma$, S chooses a deterministic side contract. The standard revelation principle does not apply to this setting, as it may be impossible to replicate mixed reporting strategies to an indirect side contract in a deterministic direct side contract (Strausz 2003). However, as the side contract governs the interaction of one principal S and one agent A, a revelation principle in terms of payoffs due to Strausz (2003) applies: For any indirect side contract, there exists a direct and truthful side contract that gives both S and A at least the same payoffs as the indirect side contract. Thus, for any β, there is a direct and truthful side contract γ that is a best response for S. It follows that, given the definition of equilibrium in Section 3, S offers a direct and truthful side contract in any equilibrium, and that some direct and truthful side contract is optimal for S if and only if there does not exist another truthful and direct side contract that is strictly better for S.

I now prove the second part of the lemma. Consider a PBE with passive beliefs under either centralization or delegation in which P offers a grand contract β and S offers a side contract γ. Let $\overline{u}_A(\sigma, \theta)$ denote A's payoff on the equilibrium path for each realization of $(\sigma, \theta) \in \mathrm{Supp}(\tilde{\theta}, \tilde{\sigma})$. I want to argue that there exists another equilibrium in which P offers the grand contract $\beta_0 \equiv \beta \circ \gamma$—with $M_S = \Sigma$ and $M_A = \Sigma \times \Theta$ under centralization, and with $M_S = \Sigma^2 \times \Theta$ under delegation, and S offers the null side contract γ_0. Given β_0, γ_0 is truthful, as A's mapping from reports to payoffs remains the same as with β and γ. Furthermore, the null side contract is optimal for S. Toward a contradiction, suppose that the null side contract is suboptimal. Then there exists a side contract γ^* such that for at least one $\sigma \in \Sigma$, γ^* gives A a payoff of at least $\overline{u}_A(\sigma, \theta)$ for any θ and gives S a strictly higher expected payoff than the null side contract. However, this implies that the side contract $\gamma^{**} \equiv \gamma \circ \gamma^*$ is a profitable deviation from γ if P offers the grand contract β. This leads to a contradiction. Finally, note that both equilibria are payoff-equivalent by construction.

Proof of Lemma 2

I prove the result through a sequence of lemmas.

Lemma 5. *For any feasible grand contract, there exist the functions* $t_S^0, t_S^1, t_A^0, t_A^1 : \Sigma \to \mathbb{R}$ *such that* $t_i(\sigma, \sigma, \theta) = X(\sigma, \sigma, \theta) t_i^1(\sigma) + (1 - X(\sigma, \sigma, \theta)) t_i^0(\sigma)$ *for* $i \in \{A, S\}$.

Proof. I start by proving the result for $i = A$. Fix some signal realization $\sigma \in \Sigma$. For any two types $\theta, \theta' \in \operatorname{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = X(\sigma, \sigma, \theta')$, ($\mathrm{IC}_A^{\gamma_0}$) implies that $t_A(\sigma, \sigma, \theta) = t_A(\sigma, \sigma, \theta')$. Thus, there exist two functions $t_A^0 : \Sigma \to \mathbb{R}$ and $t_A^1 : \Sigma \to \mathbb{R}$ such that $t_A(\sigma, \sigma, \theta) = t_A^j(\sigma)$ whenever $\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma)$ satisfies $X(\sigma, \sigma, \theta) = j$ for $j \in \{0, 1\}$.

I prove the result for $i = S$ by contradiction. Suppose the statement does not hold. Then there exists a feasible grand contract, a signal $\sigma' \in \Sigma$, and two types $\theta', \theta'' \in \operatorname{Supp}(\tilde{\theta}|\sigma')$ such that $X(\sigma', \sigma', \theta') = X(\sigma', \sigma', \theta'')$ and $t_S(\sigma', \sigma', \theta') > t_S(\sigma', \sigma', \theta'')$. Consider a side contract γ' with $\rho(\theta; \sigma) = (\sigma, \sigma, \theta)$ for all $(\sigma, \theta) \neq (\sigma', \theta'')$, $\rho(\theta''; \sigma') = (\sigma', \sigma', \theta')$, and $\tau(\theta; \sigma) = 0$. Due to the argument for the case $i = A$, γ' leads to the same payoffs for A as γ_0. Thus, A reports truthfully to γ'. However, S has a strictly higher payoff under γ' than under γ_0. Thus, the grand contract is not feasible. □

Lemma 6. *Any feasible grand contract satisfies, without loss of optimality,* $t_A^0(\sigma) \geq 0$ *and* $p(\sigma) \equiv t_A^1(\sigma) - t_A^0(\sigma) \geq \sup\{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\}$.

Proof. I start by showing that (a) $t_A^0(\sigma) \geq 0 \; \forall \sigma \in \Sigma$ with $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] < 1$, (b) $p(\sigma) \geq \sup\{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\} \; \forall \sigma \in \Sigma$ with $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] \in (0, 1)$, and (c) $t_A^1(\sigma) \geq \sup\{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\} \; \forall \sigma \in \Sigma$ with $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] = 1$.

The relationship $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] < 1$ implies that there exists some $\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = 0$. Equations ($\mathrm{PC}_A$) for centralization and ($\mathrm{PC}_A^{\gamma_0}$) for delegation imply that $t_A(\sigma, \sigma, \theta) \geq 0$. Result (a) then follows from Lemma 5. The relationship $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] \in (0, 1)$ implies that there exist $\theta, \theta' \in \operatorname{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = 0$ and $X(\sigma, \sigma, \theta') = 1$. It follows from ($\mathrm{IC}_A^{\gamma_0}$) that $p(\sigma) \geq \theta''$ for all θ'' such that $X(\sigma, \sigma, \theta'') = 1$, which proves (b). For $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] = 1$, $X(\sigma, \sigma, \theta) = 1$ for all $\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma)$. Equations ($\mathrm{PC}_A$) for centralization and ($\mathrm{PC}_A^{\gamma_0}$) for delegation imply (c). The lemma then follows from setting $t_A^0(\sigma) = 0$ for $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] = 1$ and $p(\sigma) = 0$ for $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] = 0$. □

Lemma 7. *Any feasible grand contract satisfies, without loss of optimality,* $t_S^j(\sigma) \geq -\ell$ *for* $j \in \{0, 1\}$ *and* $t_S^j(\sigma) < 0 \Rightarrow t_S^{j'}(\sigma) \geq 0$ *for* $j, j' \in \{0, 1\}, j \neq j'$.

Proof. That $t_S^j(\sigma) \geq -\ell \; \forall j \in \{0, 1\}$ follows directly from (LB) for all $\sigma \in \Sigma$ such that $X(\sigma, \sigma, \theta)$ is not constant in θ. For any $\sigma \in \Sigma$ with $X(\sigma, \sigma, \theta)$ constant in θ, (LB) requires either $t_S^0(\sigma) \geq -\ell$ or $t_S^1(\sigma) \geq -\ell$. However, the other condition can be satisfied without affecting payoffs as they are never realized. Finally, note that (PC_S) and Lemma 5 imply

$$\mathbb{E}\big[t_S(\sigma, \sigma, \tilde{\theta})|\sigma\big] = \mathbb{E}\big[X(\sigma, \sigma, \tilde{\theta})|\sigma\big] t_S^1(\sigma) + \big(1 - \mathbb{E}\big[X(\sigma, \sigma, \tilde{\theta})|\sigma\big]\big) t_S^0(\sigma) \geq 0.$$

Thus, either $t_S^0(\sigma)$ or $t_S^1(\sigma)$ has to be weakly positive. □

Lemma 8. *Any feasible grand contract satisfies, without loss of optimality,* $t_S^j(\sigma) + t_A^j(\sigma) = t^j$ *for* $j \in \{0, 1\}$.

Proof. I prove this result by contradiction. Suppose the statement does not hold. Then there exists a feasible grand contract, two signal realizations $\sigma', \sigma'' \in \Sigma$ with $t_S^j(\sigma') + t_A^j(\sigma') < t_S^j(\sigma'') + t_A^j(\sigma'')$, and two types $\theta', \theta'' \in \Theta$ with $X(\sigma', \sigma', \theta') = X(\sigma'', \sigma'', \theta'') = j$. Define a side contract γ such that

$$\rho(\theta; \sigma) = \begin{cases} (\sigma'', \sigma'', \theta'') & \text{if } \sigma = \sigma', \theta \in \{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma') : X(\sigma', \sigma', \theta) = j\}, \\ (\sigma, \sigma, \theta) & \text{otherwise,} \end{cases}$$

$$\tau(\theta; \sigma) = \begin{cases} t_A^j(\sigma'') - t_A^j(\sigma') & \text{if } \sigma = \sigma', \theta \in \{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma') : X(\sigma', \sigma', \theta) = j\} \\ 0 & \text{otherwise.} \end{cases}$$

For A, γ is equivalent to γ_0. Thus, γ satisfies (PC_A^γ) and (IC_A^γ). For $\sigma = \sigma'$ and $\theta \in \{\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma') : X(\sigma', \sigma', \theta) = 1\}$, S's payoff is strictly larger under γ than under γ_0: $t_S^j(\sigma'') + t_A^j(\sigma'') - t_A^j(\sigma') > t_S^j(\sigma')$. For all other combinations of signal realization and type (σ, θ), S's total payoff is the same under γ and γ_0. Thus, γ is feasible, as (LB^γ) is satisfied, and the original grand contract is not feasible, as (CP^z) is violated for $z \in \{c, d\}$. □

Lemma 9. *Any grand contract that is feasible under delegation satisfies, without loss of optimality, $t_S^1(\sigma) \geq 0$ for all $\sigma \in \Sigma$.*

Proof. I prove the result by contradiction. Suppose there exists a feasible grand contract β under delegation and a signal realization $\sigma \in \Sigma$ such that $t_S^1(\sigma) < 0$. As (PC_S) is equivalent to

$$\mathbb{E}[t_S(\sigma, \sigma, \tilde{\theta})|\sigma] = \mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] t_S^1(\sigma) + (1 - \mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma]) t_S^0(\sigma) \geq 0,$$

$t_S^1(\sigma) < 0$ implies $\mathbb{E}[X(\sigma, \sigma, \tilde{\theta})|\sigma] < 1$ and $t_S^0(\sigma) > 0$. Thus, there exists a type $\theta' \in \operatorname{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta') = 0$. Supervisor S could, therefore, offer a feasible side contract with $\rho(\theta; \sigma) = (\sigma, \sigma, \theta')$ and $\tau(\theta; \sigma) = 0$ for all $\theta \in \operatorname{Supp}(\tilde{\theta}|\sigma)$, giving S a payoff of $t_S^0(\sigma) > \mathbb{E}[t_S(\sigma, \sigma, \tilde{\theta})|\sigma]$. Thus, ($\text{CP}^d$) is not satisfied and β is not feasible under delegation. □

Proof of Lemma 3

Fix a grand contract with production probability x. The participation constraints (PC_A) and (PC_S) imply that P's payoff cannot exceed $B_1(x)$. Points (ii) and (v) of Lemma 2 imply that under delegation, $t^1 = t_S^1(\sigma) + p(\sigma) + t_A^0(\sigma) \geq \overline{\theta}(x)$. Points (ii) and (iii) of Lemma 2 imply $t^0 = t_A^0(\sigma) + t_S^0(\sigma) \geq -\ell$. Thus, P's payoff under delegation satisfies $x(v - t^1) - (1 - x)t^0 \leq x(v - \overline{\theta}(x)) + (1 - x)\ell$. It follows that P's payoff is bounded by $B_d(x) = \min\{B_1(x), B_2(x)\}$ under delegation.

Under centralization, P may set a grand contract with $t_S^1(\sigma) \geq 0$ for all $\sigma \in \Sigma$. Using the arguments from the last paragraph, P's payoff cannot exceed $B_d(x)$. Instead, consider a grand contract β with production probability x and $t_S^1(\sigma') < 0$ for some signal $\sigma' \in \Sigma$. I show that $t^0 \geq 0$ and $t^1 \geq \overline{\theta}(x) - \ell$. Point (iii) of Lemma 2 implies $t_S^0(\sigma') \geq 0$. Point (ii) of Lemma 2 implies $t_A^0(\sigma') \geq 0$. Thus, $t^0 = t_S^0(\sigma') + t_A^0(\sigma') \geq 0$.

From point (ii) of Lemma 2, it follows that there exists a signal realization $\sigma'' \in \Sigma$ such that $t_A^1(\sigma'') \geq p(\sigma'') \geq \overline{\theta}(x)$. Point (iii) of Lemma 2 implies $t_S^1(\sigma'') \geq -\ell$. Thus, $t_S^1(\sigma'') + t_A^1(\sigma'') = t^1 \geq \overline{\theta}(x) - \ell$. Equation ($\mathrm{PC}_S$) and point (ii) of Lemma 2 imply

$$\begin{aligned}
&\mathbb{E}\bigl[t_S(\sigma, \sigma, \tilde{\theta}) + t_A(\sigma, \sigma, \tilde{\theta})|\sigma\bigr] \\
&\quad \geq \mathbb{E}\bigl[t_A(\sigma, \sigma, \tilde{\theta})|\sigma\bigr] \\
&\quad = \mathbb{E}\bigl[X(\sigma, \sigma, \tilde{\theta})t_A^1(\sigma) + \bigl(1 - X(\sigma, \sigma, \tilde{\theta})\bigr)t_A^0(\sigma)|\sigma\bigr] \geq \mathbb{E}\bigl[X(\sigma, \sigma, \tilde{\theta})p(\sigma)|\sigma\bigr]
\end{aligned}$$

and $t^0 \geq 0$ implies

$$t_A(\sigma, \sigma, \theta) + t_S(\sigma, \sigma, \theta) = X(\sigma, \sigma, \theta)t^1 + \bigl(1 - X(\sigma, \sigma, \theta)\bigr)t^0 \geq X(\sigma, \sigma, \theta)t^1.$$

These inequalities imply

$$\mathbb{E}\bigl[t_S(\sigma, \sigma, \tilde{\theta}) + t_A(\sigma, \sigma, \tilde{\theta})|\sigma\bigr] \geq \mathbb{E}\bigl[X(\sigma, \sigma, \tilde{\theta}) \max\bigl\{t^1, p(\sigma)\bigr\}\bigr].$$

Using the law of iterated expectations, I find

$$\mathbb{E}\bigl[t_S(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) + t_A(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})\bigr] \geq \mathbb{E}\bigl[X(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta}) \max\bigl\{t^1, p(\tilde{\sigma})\bigr\}\bigr].$$

Due to $t^1 \geq \overline{\theta}(x) - \ell$ and $X(\sigma, \sigma, \theta)p(\sigma) \geq X(\sigma, \sigma, \theta)\theta$, P's payoff is bounded by

$$\begin{aligned}
&\mathbb{E}\bigl[X(\tilde{\sigma}, \tilde{\sigma}, \tilde{\theta})\bigl(v - \max\bigl\{\overline{\theta}(x) - \ell, \tilde{\theta}\bigr\}\bigr)\bigr] \\
&\quad \leq \int_\Sigma \int_{\underline{\theta}}^{\overline{\theta}(x)-\ell} \bigl(v - \overline{\theta}(x) + \ell\bigr)\, d\mu(\sigma, \theta) + \int_\Sigma \int_{\overline{\theta}(x)-\ell}^{\overline{\theta}(x)} (v - \theta)\, d\mu(\sigma, \theta) \\
&\quad = \int_{\underline{\theta}}^{\overline{\theta}(x)-\ell} \bigl(v - \overline{\theta}(x) + \ell\bigr)\, dF(\theta) + \int_{\overline{\theta}(x)-\ell}^{\overline{\theta}(x)} (v - \theta)\, dF(\theta) \\
&\quad = x\bigl(v - \overline{\theta}(x)\bigr) + \int_{\overline{\theta}(x)-\ell}^{\overline{\theta}(x)} F(\theta)\, d\theta.
\end{aligned}$$

The second line follows from the fact that the first line is maximal under the constraint $\int_\Sigma \int_\Theta X(\sigma, \sigma, \theta)\, d\mu(\sigma, \theta) = x$ for $X(\sigma, \sigma, \theta) = \mathbf{1}_{\theta \leq \overline{\theta}(x)}$. The third line follows from a change in the order of integration and $\int_\Sigma d\mu(\sigma, \theta) = dF(\theta)$. The last line follows from integration by parts.

Proof of Proposition 1

To accommodate the possibility of mass points, let

$$\Delta_F(\theta) \equiv F(\theta) - \lim_{\theta' \nearrow \theta} F\bigl(\theta'\bigr) \quad \text{and} \quad f(\theta) = \begin{cases} F'(\theta) & \text{if } F'(\theta) \text{ exists,} \\ \Delta_F(\theta) & \text{if } \Delta_F(\theta) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof

I need to prove that (i) I_{w^x} is a well defined information structure, (ii) β^x is feasible under delegation, and (iii) P achieves an expected payoff of $B_d(x)$ using I_{w^x} and β^x.

Before doing so, I argue that $\check{\theta}(x)$ is uniquely defined. For $x = 1$, the right-hand side of the equation in brackets in (4) is 0 and the left-hand side is 0 for $\theta = \underline{\theta}$ and strictly increasing for $\theta \in \Theta$. For $x < 1$, the left-hand side is weakly increasing for $\theta \geq \underline{\theta}$ and the right-hand side is strictly decreasing in θ. For $\theta = \underline{\theta}$, the left-hand side is 0, whereas the right-hand side is weakly positive as $\underline{\theta} \leq \overline{\theta}(x)$ and $\ell \geq 0$. For $\theta > \underline{\theta}$, the left-hand side is strictly positive and the right-hand side becomes strictly negative as θ grows large. Thus, $\check{\theta}(x)$ is unique for each $x \in [0, 1]$ according to the intermediate value theorem.

Next, I prove (i)–(iii). Let $\underline{\theta}(x) \equiv \min\{\overline{\theta}(x), \check{\theta}(x)\}$. If $w^x(\cdot)$ is a well defined weighting function, then (i) is satisfied. Note that

$$\int_\Sigma w^x(\sigma)\, d\sigma = \frac{\displaystyle\int_{\underline{\theta}(x)}^{\overline{\theta}(x)} 0\, d\sigma}{\displaystyle\int_{\underline{\theta}}^{\underline{\theta}(x)} F(\sigma)\, d\sigma} + \frac{\displaystyle\int_{\underline{\theta}}^{\underline{\theta}(x)} f(\sigma)\big(\underline{\theta}(x) - \sigma\big)\, d\sigma}{\displaystyle\int_{\underline{\theta}}^{\underline{\theta}(x)} F(\sigma)\, d\sigma} = 1.$$

Together with $w^x(\cdot) \geq 0$, this implies that I_{w^x} is well defined.

Condition (ii) holds if β^x satisfies the constraints (LB), (PC_S), (IC_S), and (CP^d). Equation (LB) is satisfied as $t^1 - \sigma = \overline{\theta}(x) - \sigma \geq 0 \geq -\ell$ and $t^0 \geq -\ell$. Supervisor S's expected payoff for $\sigma > \underline{\theta}(x)$ is $\overline{\theta}(x) - \sigma > 0$. For $\sigma \leq \underline{\theta}(x)$, it is

$$\begin{aligned}
&\frac{f(\sigma)\big(\overline{\theta}(x) - \sigma\big) + (1-x)w^x(\sigma)t^0}{f(\sigma) + (1-x)w^x(\sigma)} \\
&\quad = \frac{f(\sigma)\big(\overline{\theta}(x) - \underline{\theta}(x)\big) + f(\sigma)\big(\underline{\theta}(x) - \sigma\big) + (1-x)w^x(\sigma)t^0}{f(\sigma) + (1-x)w^x(\sigma)} \\
&\quad = \frac{f(\sigma)\big(\overline{\theta}(x) - \underline{\theta}(x)\big) + w^x(\sigma)\displaystyle\int_{\underline{\theta}}^{\underline{\theta}(x)} F(\sigma)\, d\sigma + (1-x)w^x(\sigma)t^0}{f(\sigma) + (1-x)w^x(\sigma)} = \overline{\theta}(x) - \underline{\theta}(x) \geq 0,
\end{aligned}$$

where the second equality follows from the definition of $w^x(\sigma)$ and the last equality follows from

$$\begin{aligned}
-t^0 &= \min\left\{\frac{\displaystyle\int_{\underline{\theta}}^{\overline{\theta}(x)} F(\sigma)\, d\sigma}{1-x}, \ell\right\} \\
&= \min\left\{\frac{\displaystyle\int_{\underline{\theta}}^{\overline{\theta}(x)} F(\sigma)\, d\sigma}{1-x}, \frac{\displaystyle\int_{\underline{\theta}}^{\check{\theta}(x)} F(\sigma)\, d\sigma}{1-x} - \overline{\theta}(x) + \check{\theta}(x)\right\} \\
&= \frac{\displaystyle\int_{\underline{\theta}}^{\underline{\theta}(x)} F(\sigma)\, d\sigma}{1-x} - \overline{\theta}(x) + \underline{\theta}(x).
\end{aligned}$$

Thus, (PC_S) is satisfied. Equation (IC_S) is satisfied, as any report with $\hat{\sigma}_S \neq \sigma$ leads with certainty to the worst possible payoff for S. It remains to check whether (CP^d) is satisfied. Note that under β and γ_0, A receives a take-it-or-leave-it price offer of σ. Thus, ($PC_A^{\gamma_0}$) and ($IC_A^{\gamma_0}$) are satisfied. Furthermore, (LB) implies (LB^{γ_0}). Given that the total transfers to S and A are fixed to t^0 and t^1 conditional on the production decision, any side contract γ that is strictly better than γ_0 for S needs to change the probability of production. For a given signal $\sigma \in \Sigma$, S may either lower the probability of production to 0 or increase it to $q > G(\sigma|\sigma)$. In the first case, S's payoff is $t^0 \leq 0$. In the latter case, S needs to pay A a side transfer of τ such that $\sigma + \tau > \overline{\theta}(x)$ whenever production takes place. Thus, S's expected payoff is $q(t^1 - \sigma - \tau) + (1 - q)t^0 \leq 0$. Thus, ($CP^d$) is satisfied.

Finally, (iii) holds, as P's expected payoff from β^x and I_{w^x} is

$$x\big(v - t^1\big) - (1 - x)t^0 = x\big(v - \overline{\theta}(x)\big) + \min\left\{(1 - x)\ell, \int_{\underline{\theta}}^{\overline{\theta}(x)} F(\sigma)\, d\sigma\right\} = B_d(x).$$

Proof of Proposition 2 and Lemma 4

First, I show that delegation is optimal if $\ell \geq \hat{\ell}(v)$ for some $\hat{\ell} : [0, \infty) \to [0, \overline{\ell})$. Let V_c and V_d denote P's optimal payoffs under centralization and delegation. Denote by x_c and x_d the optimal production probabilities under centralization and delegation. The following three lemmas prove Lemma 4.

Lemma 10. *If $V_c > V_d$, then*

$$\Pr\big((\tilde{\sigma}, \tilde{\theta}) \in \big\{\Sigma \times \Theta : t_S^1(\sigma) < 0, X(\sigma, \sigma, \theta) = 1\big\}\big) > 0. \tag{5}$$

Proof. I prove the result by contradiction. Suppose $V_c > V_d$ and $\Pr((\tilde{\sigma}, \tilde{\theta}) \in \{\Sigma \times \Theta : t_S^1(\sigma) < 0, X(\sigma, \sigma, \theta) = 1\}) = 0$. Recall that $V_c \leq B_1(x_c)$. Due to Lemma 2, it holds that

$$t^1 = t_S^1(\sigma) + p(\sigma) + t_A^0(\sigma) \geq \sup_{\sigma \in \Sigma : t_S^1(\sigma) \geq 0} \big\{\sup\big\{\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\big\}\big\}.$$

For $\Pr((\tilde{\sigma}, \tilde{\theta}) \in \{\Sigma \times \Theta : t_S^1(\sigma) < 0, X(\sigma, \sigma, \theta) = 1\}) = 0$, it holds that

$$\sup_{\sigma \in \Sigma : t_S^1(\sigma) \geq 0} \big\{\sup\big\{\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma) : X(\sigma, \sigma, \theta) = 1\big\}\big\} \geq \overline{\theta}(x_c).$$

Together with $t^0 \geq -\ell$, this implies $x_c(v - t^1) - (1 - x_c)t^0 \leq B_2(x_c)$. As $V_c \leq B_1(x_c)$,

$$x_c\big(v - t^1\big) - (1 - x_c)t^0 \leq \min\big\{B_1(x_c), B_2(x_c)\big\} = B_d(x_c) \leq \max_{x \in [0,1]} B_d(x) = V_d,$$

which gives a contradiction. $\square$

Lemma 11. *If $\overline{W} - V_c < \varepsilon$ for all $\varepsilon > 0$, then for any $\delta > 0$,*

$$\Pr\big((\tilde{\sigma}, \tilde{\theta}) \in \big\{\Sigma \times \Theta : X(\sigma, \sigma, \theta) = 1, p(\sigma) - \theta > \delta\big\}\big) = 0. \tag{6}$$

Proof. I prove the result by contradiction. Suppose $\overline{W} - V_c < \varepsilon$ for all $\varepsilon > 0$ and $\Pr((\tilde{\sigma}, \tilde{\theta}) \in \{\Sigma \times \Theta : X(\sigma, \sigma, \theta) = 1, p(\sigma) - \theta > \delta\}) = \alpha > 0$ for some $\delta > 0$. Let $X_c(\sigma, \sigma, \theta)$ be the optimal production rule. Define $\Delta \equiv \{\mathrm{Supp}(\tilde{\sigma}, \tilde{\theta}) : X_c(\sigma, \sigma, \theta) = 1, p(\sigma) - \theta > \delta\}$ and $\Delta^c \equiv \mathrm{Supp}(\tilde{\sigma}, \tilde{\theta}) \setminus \Delta$. Note that

$$
\begin{aligned}
V_c &= \iint \big(X_c(\sigma, \sigma, \theta)\big(v - t_A^1(\sigma) - t_S^1(\sigma)\big) - \big(1 - X_c(\sigma, \sigma, \theta)\big)\big(t_A^0(\sigma) + t_S^0(\sigma)\big)\big)\, d\mu(\sigma, \theta) \\
&\le \iint X_c(\sigma, \sigma, \theta)\big(v - p(\sigma)\big)\, d\mu(\sigma, \theta) \\
&\le \iint_{\Delta^c} X_c(\sigma, \sigma, \theta)(v - \theta)\, d\mu(\sigma, \theta) + \iint_{\Delta} (v - \theta - \delta)\, d\mu(\sigma, \theta) \\
&\le \int_{\underline{\theta}}^{v} (v - \theta)\, dF(\theta) - \iint_{\Delta} \delta d\mu(\sigma, \theta) = \overline{W} - \alpha\delta < \overline{W},
\end{aligned}
$$

where the first inequality in the second line follows from (PC_S) and (ii) of Lemma 2, and the second inequality follows from point (ii) of Lemma 2 and $p(\sigma) > \theta + \delta$ for all $(\sigma, \theta) \in \Delta$. Thus, $\overline{W} - V_c > \alpha\delta$, which leads to a contradiction. □

Lemma 12. *Suppose the information structure I, and suppose the grand contract β satisfy conditions (5) and (6). Then β is not feasible under centralization.*

Proof. I prove the lemma by contradiction. Suppose (I, β) satisfy (5) and (6) and β is feasible under centralization. From (5), it follows that there exists a set of signals $\underline{\Sigma}$ with positive mass such that $\sigma \in \underline{\Sigma}$ implies $t_S^1(\sigma) < 0$ and $\Pr(X(\sigma, \sigma, \tilde{\theta}) = 1 | \tilde{\sigma} = \sigma) > 0$. If $X(\sigma, \sigma, \theta) = 1$ for all $\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma)$, then S's expected payoff after σ would be negative. Thus, (PC_S) implies that for all $\sigma \in \underline{\Sigma}$, there exists $\theta_\sigma^0 \in \mathrm{Supp}(\tilde{\theta}|\sigma)$ such that $X(\sigma, \sigma, \theta_\sigma^0) = 0$ and $t_S^0(\sigma) > 0$. Furthermore, (6) implies that for all $\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = 1$ and all $\delta > 0$, it holds that $p(\sigma) - \theta < \delta$.

Consider now the side contract γ given by

$$
\rho(\theta; \sigma) = \begin{cases} \big(\sigma, \sigma, \theta_\sigma^0\big) & \text{if } \sigma \in \underline{\Sigma}, \\ (\sigma, \sigma, \theta) & \text{otherwise,} \end{cases} \quad \text{and} \quad \tau(\theta; \sigma) = \begin{cases} -\varepsilon & \text{if } \sigma \in \underline{\Sigma}, \\ 0 & \text{otherwise,} \end{cases}
$$

where $\varepsilon > 0$. This side contract is identical to the null side contract for $\sigma \notin \underline{\Sigma}$. Thus, the side contract is feasible for $\sigma \notin \underline{\Sigma}$. For $\sigma \in \underline{\Sigma}$, the side contract leads to a constant production decision $X = 0$ and a constant payment to A of $t_A^0(\sigma) + \varepsilon$. Thus, it satisfies ($\mathrm{IC}_A^\gamma$). Any type $\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = 0$ wants to accept γ, as it increases A's payoff from $t_A^0(\sigma)$ to $t_A^0(\sigma) + \varepsilon$. Any type $\theta \in \mathrm{Supp}(\tilde{\theta}|\sigma)$ with $X(\sigma, \sigma, \theta) = 1$ wants to accept this side contract, as it leads to a payoff of $t_A^0(\sigma) + \varepsilon \ge t_A^0(\sigma) + p(\sigma) - \theta$, as $p(\sigma) - \theta < \delta$ for any $\delta > 0$. Finally, note that γ leads to the same expected payoff for S as the null side contract for $\sigma \notin \underline{\Sigma}$ and leads to an expected payoff of $t_S^0(\sigma) - \varepsilon$ for $\sigma \in \underline{\Sigma}$ with

$$
t_S^0(\sigma) - \varepsilon > \mathbb{E}\big[X(\sigma, \sigma, \tilde{\theta})|\sigma\big] t_S^1(\sigma) + \big(1 - \mathbb{E}\big[X(\sigma, \sigma, \tilde{\theta})|\sigma\big]\big) t_S^0(\sigma)
$$

for ε sufficiently close to 0. Thus, γ is feasible and strictly profitable for S. It follows that (CP^c) is not satisfied and, therefore, β is not feasible. This gives a contradiction. □

Point (i) of Proposition 2 results from Lemma 4 and the following two observations. First, $\max_{x\in[0,1]} B_d(x)$ is continuous and increasing in ℓ as $B_2(x)$ is continuous and increasing in ℓ. Second, due to Corollary 3, $\max_{x\in[0,1]} B_d(x) = \overline{W}$ for $v < \overline{\theta}$ and $\ell \geq \overline{\ell}$, and $\lim_{\ell\to\infty} \max_{x\in[0,1]} B_d(x) = \overline{W}$ for $v \geq \overline{\theta}$. Thus, that there exists a function $\hat{\ell} : [0, \infty) \to [0, \overline{\ell}(v))$ that gives, for each v, a bound $\hat{\ell}(v)$ such that delegation is optimal if $\ell \geq \hat{\ell}(v)$.

Second, I prove that delegation is optimal if $v < m$. Let $x_3 \in \arg\max_{x\in[0,1]} B_3(x)$. Note that $x_3 \leq F(v)$ as $B_3(x) = B_1(x) - \int_{\underline{\theta}}^{\overline{\theta}(x)-\ell} F(\theta)\,d\theta$ is strictly decreasing for $x > F(v)$ due to $B_1(x)$ being strictly decreasing for $x > F(v)$ and $\int_{\underline{\theta}}^{\overline{\theta}(x)-\ell} F(\theta)\,d\theta$ being weakly increasing. Due to $v < m$, we have $F(\overline{\theta}(x_3)) \leq 1/2$ and

$$
\begin{aligned}
B_3(x_3) &= x_3\bigl(v - \overline{\theta}(x_3)\bigr) + \int_{\overline{\theta}(x_3)-\ell}^{\overline{\theta}(x_3)} F(\theta)\,d\theta \leq x_3\bigl(v - \overline{\theta}(x_3)\bigr) + \int_{\overline{\theta}(x_3)-\ell}^{\overline{\theta}(x_3)} F\bigl(\overline{\theta}(x_3)\bigr)\,d\theta \\
&\leq x_3\bigl(v - \overline{\theta}(x_3)\bigr) + \ell\bigl(1 - F\bigl(\overline{\theta}(x_3)\bigr)\bigr) = B_2(x_3).
\end{aligned}
$$

Now I can prove $\max_{x\in[0,1]} B_c(x) = \max_{x\in[0,1]} B_d(x)$ for $v < m$ by contradiction. Suppose $\max_{x\in[0,1]} B_c(x) > \max_{x\in[0,1]} B_d(x)$ and let $x_z \in \arg\max_{x\in[0,1]} B_z(x)$ with $z \in \{c, d\}$. The relationship $B_d(x_d) < B_c(x_c)$ implies $B_2(x_c) < B_3(x_c)$, as $B_2(x_c) \geq B_3(x_c)$ would imply $B_c(x_c) = \min\{B_1(x_c), B_2(x_c)\} = B_d(x_c) \leq B_d(x_d)$, which contradicts $B_d(x_d) < B_c(x_c)$. The relationship $B_2(x_c) < B_3(x_c)$ implies $B_c(x_c) = B_3(x_c)$, as $B_3(x) \leq B_1(x)$ for all $x \in [0, 1]$. It follows that $B_d(x_d) < B_c(x_c) = B_3(x_c) \leq B_3(x_3) \leq B_2(x_3) \leq B_d(x_d)$, which is a contradiction. Thus, $V_c = V_d$ for $v < m$.

Proof of Proposition 3

I construct the combination of information structure and grand contract (I^c, β^c) described in the main text. Define β^c by

$$
X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = \mathbf{1}_{\hat{\sigma}_S=\hat{\sigma}_A=\hat{\theta}\leq z}(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}), \qquad t_A(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = p(\hat{\theta})X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}),
$$

$$
t_S(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta}) = \begin{cases} X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta})\bigl(t^1 - p(\hat{\theta})\bigr) + \bigl(1 - X(\hat{\sigma}_S, \hat{\sigma}_A, \hat{\theta})\bigr)t^0 & \text{if } \hat{\sigma}_S = \hat{\sigma}_A, \\ -\ell & \text{if } \hat{\sigma}_S \neq \hat{\sigma}_A. \end{cases}
$$

$$
\text{with } t^0 = \frac{\displaystyle\int_{z-\ell}^{z} \bigl(F(z) - F(\theta)\bigr)\,d\theta}{1 - F(z)}, t^1 = t^0 + z - \ell,\ p(\sigma) = \sigma + t^0, z \in \Theta.
$$

So as to specify the information structure, I define the weighting function

$$
w(\sigma) = \frac{f(\sigma)[\sigma - z + \ell]_+}{\displaystyle\int_{z-\ell}^{z} \bigl(F(z) - F(\theta)\bigr)\,d\theta},
$$

which satisfies $w(\cdot) \geq 0$ and

$$
\int_{\underline{\theta}}^{z} w(\sigma)\,d\sigma = \int_{z-\ell}^{z} \frac{f(\sigma)(\sigma - z + \ell)}{\displaystyle\int_{z-\ell}^{z} \bigl(F(z) - F(\theta)\bigr)\,d\theta}\,d\sigma = 1,
$$

and define the type $\xi \in \Theta$, which satisfies

$$\int_{z-t^0}^{z} w(\sigma)\big(1-F(z)\big)\,d\sigma = 1-F(\xi).$$

Note that $\xi = z$ if $t^0 \geq \ell$ or $z - t^0 \leq \underline{\theta}$; otherwise, $\xi \in (z, \overline{\theta})$. The information structure I^c is now constructed as follows. For any $\theta \leq z$, the signal realization $\sigma = \theta$ is generated. Any $\theta \in [\xi, \overline{\theta}]$ generates some $\sigma \in \Sigma_h = [z - t^0, z]$ according to the density $w_h(\sigma) = (1 - F(z))/(1 - F(\xi)) \cdot w(\sigma)$. Any $\theta \in [z, \xi)$ generates some $\sigma \in \Sigma_l = [\underline{\theta}, z - t^0]$ according to the density $w_l(\sigma) = (1 - F(z))/(F(\xi) - F(z)) \cdot w(\sigma)$.[36] Thus, I^c is well defined for a small ℓ, as $w_l(\sigma), w_h(\sigma) \geq 0$,

$$\int_{z-t^0}^{z} w_h(\sigma)\,d\sigma = \int_{z-t^0}^{z} \frac{1-F(z)}{1-F(\xi)} w(\sigma)\,d\sigma = 1,$$

$$\begin{aligned}\int_{\underline{\theta}}^{z-t^0} w_l(\sigma)\,d\sigma &= \int_{\underline{\theta}}^{z-t^0} \frac{1-F(z)}{F(\xi)-F(z)} w(\sigma)\,d\sigma = \frac{1-F(z)}{F(\xi)-F(z)}\left(1 - \int_{z-t^0}^{z} w(\sigma)\,d\sigma\right)\\ &= \frac{1-F(z)}{F(\xi)-F(z)}\left(1 - \frac{1-F(\xi)}{1-F(z)}\right) = 1.\end{aligned}$$

Next, I prove that β^c is feasible under centralization given I^c. Equations (PC_A) and (IC_A) are satisfied due to β^c being equivalent to a price offer $p(\sigma)$ to A whenever $\operatorname{Supp}(\tilde{\theta}|\sigma) \cap (\sigma, \sigma + t^0) = \varnothing$ under I^c. The equality $\operatorname{Supp}(\tilde{\theta}|\sigma) \cap (\sigma, \sigma + t^0) = \varnothing$ holds if $\xi \geq z + t^0$. For $\ell = 0$, $t^0 = 0 < \xi = \overline{\theta}$. As ξ is continuously decreasing in t^0, it follows that $\xi \geq z + t^0$ for ℓ small.

Next I show that (PC_S) is satisfied. As cost levels below $z - \ell$ are perfectly revealed, the expected payoff of S for $\sigma < z - \ell$ is $z - \ell - \sigma \geq 0$. For $\sigma \geq z - \ell$, the expected payoff is

$$\begin{aligned}&\Pr(\tilde{\theta} = \sigma|\tilde{\sigma} = \sigma)\big(t^1 - p(\sigma)\big) + \big(1 - \Pr(\tilde{\theta} = \sigma|\tilde{\sigma} = \sigma)\big)t^0\\ &= \frac{f(\sigma)(z-\ell-\sigma)}{f(\sigma)+w(\sigma)\big(1-F(z)\big)} + \frac{w(\sigma)\big(1-F(z)\big)}{f(\sigma)+w(\sigma)\big(1-F(z)\big)} \frac{\int_{z-\ell}^{z}\big(F(z)-F(\theta)\big)\,d\theta}{1-F(z)}\\ &= \frac{f(\sigma)(z-\ell-\sigma)+f(\sigma)(\sigma-z+\ell)}{f(\sigma)+w(\sigma)\big(1-F(z)\big)} = 0.\end{aligned}$$

Furthermore, (IC_S) is satisfied, as S never has an incentive to unilaterally misreport the signal realization. This would lead to $\hat{\sigma}_A \neq \hat{\sigma}_S$ and a certain payoff of $-\ell$ to S.

So as to prove that β^c is feasible, it remains to show that (CP^c) is satisfied. Note first that any side contract γ with $\hat{\sigma}_A \neq \hat{\sigma}_S$ can never be better than the null side contract γ_0, as it leads to the minimal total payoff of $-\ell$ to S and A. Some γ with $\hat{\sigma}_A = \hat{\sigma}_S$ can only be better for S than γ_0 if it induces a different production decision than γ_0. This follows from the fact that the sum of transfers to S and A is fixed conditional on the production decision (t^1 if good is produced; t^0 otherwise). Given γ_0 and some signal

[36] If $\xi = z$, costs in $[\underline{\theta}, z - t^0]$ are perfectly revealed and $w_l(\sigma)$ can be ignored.

realization σ, only the type $\theta = \sigma$ produces and receives a transfer of $p(\sigma) = \sigma + t^0$. A side contract γ that induces the type $\theta = \sigma$ to not produce requires S to make a side transfer of at least $-\tau = t^0$ to A. Supervisor S's expected payoff from γ is, therefore, given by $t^0 + \tau = 0$, which is not better than the payoff from γ_0. For any signal realization $\sigma \geq z - \ell$, a side contract may also change the production decision by inducing some type $\theta' \in \mathrm{Supp}(\tilde{\theta}|\sigma)$ with $\theta' > \sigma$ to produce the good. This requires that S pays A a side transfer $-\tau \geq \theta' - p(\hat{\sigma})$. Thus, S's payoff under production satisfies $t^1 - p(\hat{\sigma}) + \tau \leq t^1 - \theta' = t^0 + z - \ell - \theta' < t^0 + z - \ell - \sigma \leq t^0$, which is worse than under γ_0. It follows that β^c satisfies (CP^c).

Finally, I show that P's maximal expected payoff from (I^c, β^c) exceeds the optimal payoff under delegation and approximates the upper bound under centralization if ℓ is close to 0 and $p^*(v) \in (m, \overline{\theta})$. Principal P's maximal expected payoff from (I^c, β^c) is

$$\max_z F(z)\big(v - t^1\big) - \big(1 - F(z)\big)t^0 = \max_z F(z)(v - z + \ell) - \frac{\displaystyle\int_{z-\ell}^{z} \big(F(z) - F(\theta)\big)\, d\theta}{1 - F(z)}.$$

I denote the solution to this problem by $z_c(\ell)$ and denote the value by $V_c(\ell)$. Equivalently, I denote P's maximal payoff under delegation by $V_d(\ell) = \max_z \tilde{B}_d(z)$, where

$$\tilde{B}_d(z) \equiv B_d\big(F(z)\big) = F(z)(v - z) + \min\left\{\int_{\underline{\theta}}^{z} F(\theta)\, d\theta, \big(1 - F(z)\big)\ell\right\},$$

and I denote by $z_d(\ell)$ the optimal cutoff under delegation. First, I prove $V_c(\ell) > V_d(\ell)$ for $p^* \in (m, \overline{\theta})$ and small ℓ. Note that $\tilde{B}_d(z) = F(z)(v - z) + (1 - F(z))\ell$ for ℓ sufficiently small. Note, furthermore, that $V_c(\ell)$ and $V_d(\ell)$ are continuous in ℓ, $z_c(0) = z_d(0) = p^*$, and $V_c(0) = V_d(0) = \underline{W}$. From the envelope theorem, it follows that $V_c'(0) = F(z_c(0)) = F(p^*)$ and $V_d'(0) = 1 - F(z_d(0)) = 1 - F(p^*)$. From $p^* > m$, it follows that $V_c(\ell) > V_d(\ell)$ for $\ell \in (0, \varepsilon)$ if ε is sufficiently small.

Second, I prove that (I^c, β^c) is approximately optimal for $p^* > m$ and $\ell \in (0, \varepsilon)$ for ε sufficiently small. As $V_c(\ell) > V_d(\ell)$, it is enough to prove that $V_c(\ell)$ approximates the upper bound $V_3(\ell) \equiv \max_z B_3(F(z))$. As $V_3(0) = \underline{W}$ and $V_3'(0) = F(p^*)$, $V_c(\ell)$ is a first-order approximation of $V_3(\ell)$ around $\ell = 0$. Thus, (I^c, β^c) is near optimal for $p^* > m$ and ℓ close to 0.

References

ANSA March 1, 2017, ANAC inspected CONSIP in 2016. retrieved from http://www.ansa.it/english/news/business/2017/03/01/anac-inspected-consip-in-2016-3 on March 7, 2018. [1547]

Asseyer, Andreas (2020), "Collusion and delegation under information control." FUB Discussion Paper 2020/3, Economics, 10.17169/refubium-26337. [1570]

Baliga, Sandeep and Tomas Sjöström (1998), "Decentralization and collusion." *Journal of Economic Theory*, 83, 196–232. [1550]

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris (2015), "The limits of price discrimination." *American Economic Review*, 105, 921–957. [1551]

Bergemann, Dirk and Stephen Morris (2013), "Robust predictions in games with incomplete information." *Econometrica*, 81, 1251–1308. [1551]

Bergemann, Dirk and Stephen Morris (2016), "Bayes correlated equilibrium and the comparison of information structures in games." *Theoretical Economics*, 11, 487–522. [1551]

Börgers, Tilman (2015), *An Introduction to the Theory of Mechanism Design*. Oxford University Press, New York. [1554]

Caillaud, Bernard and Philippe Jéhiel (1998), "Collusion in auctions with externalities." *RAND Journal of Economics*, 29, 680–702. [1550]

Celik, Gorkem (2009), "Mechanism design with collusive supervision." *Journal of Economic Theory*, 144, 69–95. [1548, 1550, 1554, 1555, 1556, 1559, 1567, 1570, 1572]

Che, Yeon-Koo, Daniele Condorelli, and Jinwoo Kim (2018), "Weak cartels and collusion-proof auctions." *Journal of Economic Theory*, 178, 398–435. [1551]

Che, Yeon-Koo and Jinwoo Kim (2006), "Robustly collusion-proof implementation." *Econometrica*, 74, 1063–1107. [1550, 1570]

Che, Yeon-Koo and Jinwoo Kim (2009), "Optimal collusion-proof auctions." *Journal of Economic Theory*, 144, 565–603. [1551]

Crawford, Vincent P. and Joel Sobel (1982), "Strategic information transmission." *Econometrica*, 50, 1431–1451. [1551]

Crémer, Jacques (1996), "Manipulations by coalitions under asymmetric information: The case of groves mechanisms." *Games and Economic Behavior*, 13, 39–73. [1550]

Dequiedt, Vianney (2007), "Efficient collusion in optimal auctions." *Journal of Economic Theory*, 136, 302–323. [1551]

Di Tella, Rafael and Ernesto Schargrodsky (2003), "The role of wages and auditing during a crackdown on corruption in the city of buenos aires." *Journal of Law and Economics*, 46, 269–292. [1547]

Faure-Grimaud, Antoine, Jean-Jacques Laffont, and David Martimort (2003), "Collusion, delegation and supervision with soft information." *Review of Economic Studies*, 70, 253–279. [1548, 1549, 1550, 1554, 1555, 1556, 1559, 1566, 1567, 1570, 1572]

Fudenberg, Drew and Jean Tirole (1991), "Perfect Bayesian equilibrium and sequential equilibrium." *Journal of Economic Theory*, 53, 236–260. [1556]

Green, Jerry and Jean-Jacques Laffont (1979), "On coalition incentive compatibility." *Review of Economic Studies*, 46, 243–254. [1550]

Ivanov, Maxim (2010), "Informational control and organizational design." *Journal of Economic Theory*, 145, 721–751. [1551]

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian persuasion." *American Economic Review*, 101, 2590–2615. [1548, 1551]

Laffont, Jean-Jacques and David Martimort (1997), "Collusion under asymmetric information." *Econometrica*, 65, 875–911. [1550, 1557, 1570]

Laffont, Jean-Jacques and David Martimort (1998), "Collusion and delegation." *RAND Journal of Economics*, 29, 280–305. [1550]

Laffont, Jean-Jacques and David Martimort (2000), "Mechanism design with collusion and correlation." *Econometrica*, 68, 309–342. [1550, 1556]

Mathevet, Laurent, Jacopo Perego, and Ina Taneva (2020), "On information design in games." *Journal of Political Economy*, 128, 1370–1404. [1551]

McAfee, R. Preston and John McMillan (1992), "Bidding rings." *American Economic Review*, 82, 579–599. [1550]

Mookherjee, Dilip, Alberto Motta, and Masatoshi Tsumagari (2020a), "Supply chain integration: A theory based on collusion and communication costs." Unpublished paper, Department of Economics, Boston University. [1550]

Mookherjee, Dilip, Alberto Motta, and Masatoshi Tsumagari (2020b), "Consulting collusive experts." *Games and Economic Behavior*, 122, 290–317. [1550, 1554, 1556, 1570, 1572, 1573]

Mookherjee, Dilip and Masatoshi Tsumagari (2004), "The organization of supplier networks: Effects of delegation and intermediation." *Econometrica*, 72, 1179–1219. [1551]

Mookherjee, Dilip and Masatoshi Tsumagari (2019), "Hierarchical control rights and strong collusion." Unpublished paper, Department of Economics, Boston University. [1550]

Negenborn, Colin von and Martin Pollrich (2020), "Sweet Lemons: Mitigating collusion in organizations." *Journal of Economic Theory*, 189, 105074. [1551]

OECD (2014), "OECD foreign bribery report: An analysis of the crime of bribery of foreign public officials." OECD publishing. [1547]

OECD (2017), "Government at a glance." OECD publishing. [1547, 1548]

Ortner, Juan and Sylvain Chassang (2018), "Making corruption harder: Asymmetric information, collusion, and crime." *Journal of Political Economy*, 126, 2108–2133. [1551]

Pavlov, Gregory (2008), "Auction design in the presence of collusion." *Theoretical Economics*, 3, 383–429. [1551]

Rayo, Luis and Ilya Segal (2010), "Optimal information disclosure." *Journal of Political Economy*, 118, 949–987. [1551]

Roesler, Anne-Katrin and Balázs Szentes (2017), "Buyer-optimal learning and monopoly pricing." *American Economic Review*, 107, 2072–2080. [1551]

Strausz, Roland (2003), "Deterministic mechanisms and the revelation principle." *Economics Letters*, 79, 333–337. [1557, 1574]

Taneva, Ina (2019), "Information design." *American Economic Journal: Microeconomics*, 11, 151–185. [1551]

Tirole, Jean (1986), "Hierarchies and bureaucracies: On the role of collusion in organizations." *Journal of Law, Economics and Organization*, 2, 181–214. [1550]

Co-editor Dilip Mookherjee handled this manuscript.

Manuscript received 14 December, 2018; final version accepted 15 February, 2020; available online 24 February, 2020.