The Effect of Exit Rights on Cost-based Procurement Contracts*

Rodrigo Andrade[†], Henrique Castro-Pires[‡], Humberto Moreira[§]

August 15, 2023

This paper studies optimal procurement contracts in an environment with dynamic information arrival and ex-post exit rights. A procuring agency designs contracts for a firm that receives information over time. In the first period, the firm gets a private non-fully informative signal about the project's cost. In the second period, the firm fully learns the cost and decides whether to keep the contract or take an exogenous ex-post outside option. We show that if the ex-post outside option value is sufficiently close to the ex-ante, the optimal mechanism takes a static form: all first-period signal reports are pooled into a single contract, and payments depend solely on the second-period reports. The interpretation is that optimal contracts do not condition transfers on ex-ante self-reported cost estimates but only on realized verifiable costs. Finally, we study if competition among a large number of firms allows the procuring agency to screen the first-period information and implement the second-best allocation. We show that the answer depends on the value of the ex-post outside option: for low values, the procuring agency can screen the first-period information and implement the second-best allocation at approximately no additional cost, while for high values, the cost of implementing the second-best allocation diverges. We relate our findings to firms' incentives to under-report expected costs, and the ubiquitous cost-overruns observed empirically in public projects.

Keywords: cost-based procurement, cost-overrun, dynamic mechanism design, ex-post participation constraints, limited liability

*For their helpful comments and suggestions, we thank Jeroen Swinkels, Alessandro Pavan, George Georgiadis, David Besanko, Michael Riordan, Fernando Barros, Jefferson Bertolai, Fernando Barbosa, and audiences at 2022 SEA Annual Meeting (Fort Lauderdale, FL), UFRGS, and FEARP/USP. Moreira acknowledges FAPERJ and CNPq for financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The views expressed in this article are those of the authors. They do not necessarily represent those of the World Bank. All errors and omissions are our own.

[†]World Bank, (email: rbandrade@worldbank.org).

[‡]University of Surrey, (email: h.castro-pires@surrey.ac.uk).

[§]Fundação Getulio Vargas (FGV/EPGE), Brazilian School of Economics, (email: humberto.moreira@fgv.br).

1 Introduction

In procurement relationships, the information regarding a project's cost often arrives gradually. Due to their superior knowledge of their own technology, contractors are usually better positioned to estimate a project's cost than the procurer. However, even the firms themselves face uncertainty regarding what are going to be the actual realized costs. The law often protects firms against ex-post unfavorable cost realizations by limiting their losses to their equity stake in the project or other more generous forms of limited liability. Such protection limits the ability of procurers to extract payments or seize the assets of insolvent suppliers. When designing procurement contracts, procurers must consider this dynamic arrival of information while ensuring the contractor is still solvent and willing to complete the project.

This paper studies optimal procurement contracts under dynamic asymmetric information and ex-post exit rights. A procuring agency designs contracts for a firm that receives information about a project's costs over time. The firm can exit the contractual arrangement and take an exogenous outside option at any time.[1] The introduction of exit rights in a dynamic procurement model captures the effect of limited liability protection on contract design and allows us to understand its interplay with asymmetric information. In practice, we often observe such constraints. For instance, the filing for bankruptcy without completing the project represents an ex-post outside option to the firm. Our main result shows that when this ex-post outside option is sufficiently attractive, the procuring agency should offer a single cost-plus contract without trying to elicit the firm's ex-ante cost estimates.

In many infrastructure sectors with large projects, it is essential to ensure project completion even when realized costs are higher than expected. Regulators in charge of managing concession contracts, concerned with supplier insolvency and service continuity, often advocate for the "principle of financial equilibrium." Under this principle, the firm is entitled to a change in contract terms when unfavorable states of nature materialize (Guasch, 2004).[2] On a similar note, infrastructure concessionaires in Brazil recently had enshrined in legislation a right to "relinquish their contracts" and have their investments reimbursed if they demonstrate the non-viability of the concession at any point in time.[3] In this paper, at first we focus on the consequences of such strong form of liability protection for incentives design in the concession context. In particular, we set up a two-period cost-based procurement model based on Laffont and Tirole (1986) and introduce an ex-post participation constraint. The analysis of the model entails a theoretical contribution concerning the optimality of static vs. dynamic contracts, connecting the findings to an ongoing discussion in the literature.

[1]The requirement of performance bonds, which are paid upfront and returned to the firm upon project completion, is often suggested as a solution to this problem but has limited scope due to cash constraints or imperfect capital markets.

[2]In principle, compensation for unfavorable events should only be sought in case these events were of extreme nature or lie completely outside the firm's agency, such as *force majeure* events, changes in law, or unilateral variations imposed by the authority. In practice, however, queries for financial equilibrium amendments pertain to a wider and grayer range of sources.

[3]Law 13,448, of 2017, which sets rules for re-tendering concession contracts, was enacted in a context in which a number of local infrastructure concessions were undergoing severe financial distress.

In our model, a procuring agency designs contracts for a firm to execute a project. The project is assumed to be valuable enough that the agency wants to implement the project in any cost scenario. The procurer designs a contract to implement the project at the lowest possible expenditure. The model has two periods: first, the firm receives a private non-fully informative signal about the project's intrinsic cost. Second, the firm learns whether the intrinsic costs are high or low, then decides whether to stay in the relationship or take its outside option.[4] In case it stays, the firm chooses an effort to reduce the cost. Cost-reducing efforts generates disutility for the firm as in the classic Laffont and Tirole (1986) procurement model. Realized costs are verifiable and observed by the procurer; however, neither the intrinsic cost nor effort are observable.

Revisiting the classic static Laffont and Tirole (1986) procurement model with two types is helpful to gain an intuition for our results. In Laffont and Tirole (1986), the optimal mechanism does not distort the effort choice of the low-cost type and distorts downward the effort of the high-cost type. The distortion reduces the information rents accrued by the most efficient (low-cost) type. In particular, the larger the prior belief that the firm is low-cost, the more critical it is to reduce the information rents and the bigger the distortion for high-cost types. In our model with ex-ante signal realizations, the better the signal, the higher the chance of a low-cost firm. Hence, if the procuring agency could directly observe the first-period signal realization, it would distort the effort levels of high-cost firms more after better first-period signal realizations. However, first-period signals are the firm's private information. We show that in a truthful direct revelation mechanism, the first-period incentive compatibility constraints can be characterized as a monotonicity and an envelope condition, which, together with the ex-post participation constraint, require lower distortions after better signals — the opposite monotonicity direction of what the principal wants absent incentive compatibility concerns.[5] As a result, the procuring agency is better off not screening the first-period information and offering a contract that conditions the transfers only on realized verifiable costs.

The paper then characterizes optimal contracts for lower values of the firm's ex-post outside option. First, we establish a threshold $\bar{u}_1$ for which any ex-post outside option value above it implies that the ex-ante participation constraint is slack, and the optimal contract takes a static form as in the classic Laffont-Tirole procurement model. Second, we establish that there exists a second lower threshold $\bar{u}_2$ for which pooling of the first-period signals remain optimal for any ex-post outside option $\bar{u} \in [\bar{u}_2, \bar{u}_1]$, even though the ex-ante participation now binds. Finally, we characterize a third threshold $\bar{u}_3$ for which for any ex-post outside option $\bar{u} \leq \bar{u}_3$, the ex-post participation constraint is slack, and — under a regularity condition — the optimal mechanism fully screens first-period signals.

We then relate our findings to the ubiquitous cost overruns observed in procurement projects. In practice, cost overruns are observed worldwide in many different sectors ranging from transport and energy infrastructure to IT projects (see Siemiatycki (2015) for a survey). For instance, Flyvbjerg et al. (2003) study cost overruns on a sample of 258 major roads, tunnels, bridges, urban transit, and interurban rail projects in 20

[4]We refer to the first-period private information as "signals" and the second-period as "types".

[5]This conflict is a form of non-responsiveness. See Castro-Pires and Moreira (2021) for further references.

countries and found that 86% of the projects in their sample had a higher cost than initially estimated. We show that cost overruns are consistent with a setting in which a miscalibrated procurer underestimates the firm's ex-post outside option value and tries to screen the ex-ante information but fails. She offers a mechanism that incentivizes the firm to report the lowest cost estimate regardless of its private signal and take the ex-post outside option if the cost realization is high. Hence, the reported ex-ante costs underestimate actual cost realizations, and the firm often exercises its exit rights.

Finally, we extend the analysis to the case with multiple firms competing for the project. In period 1, n firms observe private non-fully informative signals about their own intrinsic costs. The procuring agency designs contracts and selects a single firm to execute the project. In period 2, the selected firm learns its intrinsic cost, chooses whether to leave the project, and, in case it does not leave, the amount of cost-reducing effort. The key distinction with the baseline setting is that the procuring agency can use the project allocation decision as an extra screening instrument. We define as the second-best allocation the procuring agency's choices if it could directly observe first-period signals. That is, the second-best allocation describes which firm would be selected for the project and what efforts would be recommended if the procuring agency knew all first-period signal realizations. We, then, show that when the number of firms is sufficiently large, the procuring agency can elicit first-period signals and implement the second-best allocation at approximately no additional cost.[6] However, if the ex-post outside option value is sufficiently high, the cost of implementing the second-best allocation diverges to infinity when the number of firms increases.

The results regarding multiple firms imply that competition solves the asymmetric information issue only when the procuring agency can sufficiently punish firms that have reported a low expected intrinsic cost but have a high realized verifiable cost. To build the intuition, suppose that the procuring agency offers a mechanism that selects the firm that reports the lowest expected cost (which is the selection rule in the second-best). In that case, reporting lower expected costs increase the chance of being selected for the project and might generate incentives for firms to under-report their expected costs. For instance, any firm can pretend to have received the best first-period signal, win the project with probability one, and take their ex-post outside option if it has a high cost. When the ex-post outside option is valuable enough (as good as the ex-ante, for instance), this potential deviation implies that the procuring agency must leave strictly positive rents to every firm. As the number of firms grows, the cost of implementing the second-best allocation explodes.

However, when the ex-post outside option value is low, the procuring agency can punish under-reporting by offering lower payments to firms who have reported a low expected cost but have a high realized one. Such a punishment allows the procuring agency to increase the dispersion of ex-post utilities and screen first-period signals without leaving additional rents to firms. Moreover, increasing the number of firms reduces expected information rents and, in the limit, allows the procuring agency to elicit first-period signals essentially for free.

Outline. This paper is organized as follows. In the next section, we discuss the related

[6] This result is in the spirit of Riordan and Sappington (1987), McAfee and McMillan (1987), and Laffont and Tirole (1987a).

literature; Section 2 presents the model and formally states the procurement design problem under ex-post individual rationality. Section 3 provides the optimal solution for the "absolute exit rights" case, in which the firm's ex-post outside option is equal to the ex-ante. Section 4 analyzes the problem for lower levels of the ex-post outside option and relates the results to cost overruns. Section 6 discusses extensions beyond binary types and the role of commitment. Lastly, Section 7 concludes and proposes directions for future research. Appendix B presents a further discussion on the optimality of pooling mechanisms under exit rights and alternative information structures. Omitted proofs are left to Appendix A.

Related literature

Our paper contributes to the literature on dynamic procurement contracts. Baron and Besanko (1984) pioneered the analysis with a multiperiod model of a continuing relationship between a regulator and a firm. In a sequence of papers, Laffont and Tirole (1987a, 1988, 1990) analyze the effects of limited commitment on optimal contracts. In contrast, our article focuses on the case of full commitment and the effects of an ex-post participation constraint. We also show that our results remain valid if the procurer has limited commitment power. In another related paper on dynamic procurement contracts, Krähmer and Strausz (2011) study a moral hazard problem with pre-project planning. They analyze how to provide incentives for information acquisition prior to the execution of the project. We instead focus on how the ex-ante information relates to ex-post exit rights and how they jointly affect optimal contracts.

Our article is also related to Courty and Li (2000), who studied the optimal sales of an indivisible unit good under risk neutrality and sequential private information arrival through the lenses of a sequential screening model. In that model, the agent is locked in the relationship after accepting to participate in the mechanism. Hence, the mechanism offered by the principal had to respect only an ex-ante individual rationality constraint. As a result, the optimal mechanism consists of a menu of option-contracts.[7] In a related paper, Pavan et al. (2014) study mechanism design in dynamic quasilinear environments where the information arrives over time. For a recent dynamic mechanism design literature survey, we refer to Bergemann and Välimäki (2019).

On top of the dynamic nature, our article belongs to the literature that studies sequential screening in settings in which the agent's losses in each period are bounded below by the ex-ante outside option, e.g., Krishna et al. (2013), Krasikov and Lamba (2021), Ashlagi et al. (2022), and Krähmer and Strausz (2022). Those articles study a framework in which trade (or equivalently, production) happens in all periods, while our model focuses on the case in which the first period refers only to information, and production occurs only in the second period.

In the context of the dynamic arrival of information and a single trading period, Krähmer and Strausz (2015) introduced an ex-post individual rationality constraint to a sequential screening model motivated by withdrawal rights regulation in e-commerce

[7] Option contracts refer to an arrangement where the buyer pays an upfront fee, and when he later discovers his true valuation, he may decide to keep the good or withdraw from the contract with a pre-specified refund.

markets[8]. They find sufficient conditions for an optimal sales policy to be a *static* contract, i.e., a mechanism that ignores any first-period information the agent might report and conditions the allocation solely on ex-post private valuations. Bergemann et al. (2020) furthered their result, finding necessary and sufficient conditions for optimal pooling *versus* sequential separation with two ex-ante signals. Both papers have environments where players' utilities depend linearly on payoff-relevant private information. Linear environments give rise to *threshold* mechanisms, i.e., mechanisms that allocate the good with probability one to types above a certain threshold on the type space.[9]

Krähmer and Strausz (2016) examine nonlinear pricing with ex-post participation in an environment with curvature, binary signals, and a continuum of types. The authors find that the optimality of static contracts breaks down if the principal can sell multiple units and the distributions of types conditional on signals satisfy a cross-hazard rate condition. They argue that variable quantities afford the principal an additional screening instrument, so discriminating ex-ante information becomes profitable again. In contrast to Krähmer and Strausz (2016), we re-establish the optimality of full-pooling of the first-period signals in a setting with binary types and arbitrary signal distributions. In Appendix B, we discuss the role of the cardinality of signals and types and establish the importance of a cross-hazard rate condition to extend the results to a two-signals-three-types setting.

We also contribute to the literature studying the effect of limited liability on optimal contracts. In related work, Calveras et al. (2004) motivates the occurrence of abnormally low bids in procurement auctions if bidders can file for bankruptcy ex-post. Moreover, Ollier and Thomas (2013), Gottlieb and Moreira (2021), and Castro-Pires and Moreira (2021) show that the lack of screening might stem from the interplay between limited liability, moral hazard, and adverse selection. Our results show that a limited liability constraint also generates pooling in a dynamic setting without moral hazard.

Our paper also contributes to the literature under which multiple agents compete for incentive contracts as in Riordan and Sappington (1987), and McAfee and McMillan (1987). Differently from previous work, we introduce exit rights that allow the firm to take an ex-post outside option after fully learning the project's cost. Then, we show that competition solves the asymmetric information problem only if the ex-post outside option value is low enough. In a recent article, Chakraborty et al. (2021) show that competition in procurement might be detrimental when there are moral hazard concerns after the project has been assigned to a given firm. Despite having a similar conclusion, their results stem from different sources than ours: they focus on distortions caused by hidden actions, while we show that competition might be detrimental due to the existence of attractive ex-post outside options.

[8] Although we focus on the procurement application, our model can be recast as a monopolist selling a good in which buyers have private information about their preferences for quality (or quantity). While Krähmer and Strausz (2015) focus on the case with linear preferences (or single-unit demand), we analyze the case with preference curvature (or multi-unit demand).

[9] Hence the relevance of the discussion on the optimality of "deterministic" *versus* "stochastic" contracts, where the latter means a mechanism that allocates the good randomly to specific types.

2 Environment

The procuring agency (principal) delegates the execution of a project to a firm (agent). The project has a social benefit level assumed to be high enough to ensure the project's viability in all cost scenarios. The project's total cost is given by

$$C = \beta - e,$$

where $\beta \in \{\beta_L, \beta_H\}$ are the two possible levels for the project's intrinsic cost (low and high), and $e \in [0, \beta_H]$ denotes the effort exerted by the firm to reduce the cost to level C. Let $\Delta\beta = \beta_H - \beta_L > 0$ be the incremental cost.

The interaction happens in two periods. In the first period, the procuring agency and the firm sign the contract; in the second period, the firm executes the project. At the contracting stage, the procuring agency announces the incentive scheme to the firm, which accepts it if the expected profit from the partnership exceeds the initial reservation utility, normalized to zero. The firm is uncertain about the project's intrinsic cost β, but observes a private signal $s \in [\underline{s}, \overline{s}]$ which conveys information about the likelihood of β_H. Specifically, a firm who observes a signal realization s has probability $\Pr(\boldsymbol{\beta} = \beta_L|s) = 1 - s$ of facing a low-cost project and $\Pr(\boldsymbol{\beta} = \beta_H|s) = s$ of a high-cost project.[10] The signal s is distributed according to a cumulative distribution F with a strictly positive density f bounded away from 0.

In the production period, the firm privately learns the project's intrinsic cost β. Then, the firm has the right to exit the partnership and get an exogenous outside option $\bar{u} \leq 0$. Such a right imposes an ex-post participation constraint on contract design.

If the firm decides to complete the project, it chooses its optimal cost-reducing effort. The firm has an effort disutility $\psi : \mathbb{R}_+ \to \mathbb{R}_+$, which is assumed to be twice continuously differentiable, strictly increasing ($\psi' > 0$), strictly convex ($\psi'' > 0$), and $\psi'(\beta_H) > 1$. After the project is delivered, the procuring agency observes the final cost $C = \beta - e$, but cannot distinguish between the intrinsic cost β and cost-reduction effort e. A firm with type β and final cost C that receives a transfer T has payoff

$$V(\beta, C, T) = T - C - \psi(\beta - C).$$

It is helpful to denote payoffs in terms of the net transfer $t = T - C$, that is, the payments above the cost reimbursement. The procuring agency can fully commit to the offered contract. The procuring agency designs a transfer $T : [\underline{s}, \overline{s}] \times \mathbb{R}_+ \to \mathbb{R}_+$, specifying a remuneration schedule that depends on the firm's reported signal and on the final cost C, to minimize expected procurement costs.

By the dynamic revelation principle (Myerson (1986)), the problem is equivalent to finding an optimal direct mechanism in an incentive-compatible and individually-rational set. A direct mechanism assigns a net transfer $t_j(\hat{s})$ and a cost requirement $C_j(\hat{s})$ for each sequence of announcements $(\hat{s}, \beta_j)$.

We use the terms *ex-ante* and *ex-post* to refer to each stage of the model. *Ex-ante* refers to after the signal has been observed but before the type is realized. Finally, *ex-post* refers to after both signal and type have been realized.

[10] We denote random variables in bold.

Incentive compatibility. In the second period, after the signal has been reported, the firm's only relevant decision is which $\beta \in \{\beta_L, \beta_H\}$ to report. For $j, i \in \{L, H\}$ denote by $u_j(s|i)$ the ex-post profit of a firm of type β_i who has reported signal s and type β_j:

$$u_j(s|i) = t_j(s) - \psi(\beta_i - C_j(s)).$$

Let $u_i(s) = u_i(s|i)$ denote the firm's profit from reporting their true type in the second period. For every signal s, ex-post incentive compatibility is given by:

$$\begin{aligned} u_L(s) &\geq t_H(s) - \psi(\beta_L - C_H(s)) \\ u_H(s) &\geq t_L(s) - \psi(\beta_H - C_L(s)). \end{aligned} \tag{IC-2_s}$$

Lemma 2.1. *For all s, ex-post incentive constraints* (IC-2_s) *are equivalent to*

$$u_L(s) \geq \phi(\beta_H - C_H(s)) + u_H(s) \tag{IC-2_s^L}$$
$$u_H(s) \geq u_L(s) - \phi(\beta_H - C_L(s)), \tag{IC-2_s^H}$$

where $\phi(e) \equiv \psi(e) - \psi(e - \Delta\beta)$. Moreover, (IC-$2_s^L$) *and* (IC-$2_s^H$) *jointly imply that $C_H(s) \geq C_L(s)$, and $u_L(s) > u_H(s)$ are necessary conditions for ex-post incentive compatibility.*

In Appendix

For any mechanism that satisfies all (IC-2_s)'s, the second-period optimal continuation strategy for the firm is to report her type truthfully no matter what signal it has reported in period 1. That is, regardless of whether the firm has reported truthfully or not in the first period, (IC-2_s) ensures incentives for truthful reporting in the second-period.[11] Hence, from the first-period viewpoint, the expected payoff of a firm that has observed signal s but reports signal $\hat{s}$ is:

$$U(\hat{s}|s) = (1-s)u_L(\hat{s}) + s u_H(\hat{s}).$$

Let $U(s) = U(s|s)$ denote the firm's expected profit from reporting its signal accurately in the first period. Ex-ante incentive compatibility is thus given by:

$$U(s) \geq U(\hat{s}|s), \quad \forall \hat{s}, s. \tag{IC-$1_{s,\hat{s}}$}$$

Individual rationality. The direct mechanism satisfies multi-period voluntary participation if it secures ex-ante and ex-post profits larger than the firm's opportunity cost in each period, i.e.[12]

$$U(s) \geq 0, \quad \forall s \tag{IR-1}$$
$$u_H(s) \geq \bar{u}, \quad \forall s. \tag{IR-2}$$

[11] Note that Myerson (1986)'s revelation principle only requires truth-telling in the second-period after a truthful report in the first. However, in our setting, the firm's period two utility is independent of the signal. Hence, incentive compatibility in period two automatically ensures truth-telling even when the firm has misreported its first-period signal.

[12] Under (IC-2_s), ex-post profits are decreasing with β, i.e., $u_L(s) \geq u_H(s)$, $\forall s$, so the only relevant ex-post IR constraint concerns the expensive project type β_H.

The ex-post constraint (IR-2) captures the concept of limited liability in procurement contracting, as the firm is entitled to stop its loss at $\bar{u} \leq 0$ in unfavorable states of the nature. Moreover, the case $\bar{u} = 0$ is interpreted as a strong form of limited liability protection, under which the firm is entitled to its ex-ante opportunity cost in all states.[13]

Design problem. Denote by $T_i(s) = t_i(s) + C_i(s)$ the ex-post transfer to the firm with reported signal s and final cost $C_i(s)$. The procuring agency wishes to minimize the expected procurement cost that guarantees the project is executed:

$$\mathcal{P}: \quad \min_{\{T_i(\cdot),C_i(\cdot)\}_{i\in\{L,H\}}} \quad \int_{\underline{s}}^{\overline{s}} \{(1-s)T_L(s) + sT_H(s)\}\, dF(s)$$
$$\text{s.t.} \qquad (\text{IC-1}_{s,\hat{s}}), (\text{IC-2}_s), (\text{IR-1}), (\text{IR-2}).$$

A direct mechanism $\{(T_i(\cdot), C_i(\cdot))\}_{i\in\{L,H\}}$ can be equivalently represented by the recommended efforts and ex-post utilities for each sequence of reports, that is, $\{(e_i(\cdot), u_i(\cdot))\}_{i\in\{L,H\}}$. Therefore, the procuring agency's design problem can be restated as follows:

$$\mathcal{P}: \quad \min_{\{(e_i(\cdot),u_i(\cdot))\}_{i\in\{L,H\}}} \int_{\underline{s}}^{\overline{s}} \{(1-s)\left[\beta_L - e_L(s) + u_L(s) + \psi(e_L(s))\right]$$
$$+s\left[\beta_H - e_H(s) + \psi(e_H(s)) + u_H(s)\right]\} f(s)ds$$

$$\begin{aligned}
\text{s.t.} \qquad & (1-s)u_L(s) + su_H(s) \geq (1-s)u_L(\hat{s}) + su_H(\hat{s}), \forall \hat{s}, s && (\text{IC-1}_{s,\hat{s}})\\
& u_H(s) \geq \bar{u}, \quad \forall s && (\text{IR-2}_s)\\
& U(s) \geq 0, \quad \forall s && (\text{IR-1}_s)\\
& u_L(s) \geq \phi(e_H(s)) + u_H(s) \quad \forall s && (\text{IC-2}_s^L)\\
& u_H(s) \geq u_L(s) - \phi(e_L + \Delta\beta), \quad \forall s. && (\text{IC-2}_s^H)
\end{aligned}$$

Two useful benchmarks

Before proceeding with the analysis of problem $\mathcal{P}$, we define two useful benchmarks by removing information frictions. In the first benchmark, we entirely remove any information frictions, while the second assumes the principal observes s but not β.

Let the *first-best allocation* (denoted by e^{FB}) be the effort recommendation the principal would provide if she could directly observe the realized signal and type. Absent any incentive compatibility constraints; the principal would recommend the effort level that maximizes joint surplus. That is, for any pair of signal and type realization, the principal would recommend

$$\psi'(e^{FB}) = 1.$$

[13] Such strong limited liability protection is sometimes referred to as a "financial equilibrium principle" in the context of concession contracts (e.g., Guasch (2004)).

Let the *second-best allocation* be the effort the principal would recommend if she could observe s but not β. In this case, the principal could directly condition the mechanism on s, but would need to elicit β in an incentive-compatible way. When the principal observes s, we are back to the classic Laffont and Tirole (1986) model. Hence, for each s, the second-best allocation is given by the solution to this classic problem and given by

$$\psi'(e_L^{SB}(s)) = 1$$
$$\psi'(e_H^{SB}(s)) = 1 - \frac{1-s}{s}\phi'(e_H^{SB}(s)).$$

That is, the effort of the best type (β_L) is set to the efficient level, while the effort of the worst type (β_H) is distorted downwards to reduce information rents to the highest type. Both benchmarks will be helpful when looking for the problem $\mathcal{P}$'s solution.

Binding Ex-post Incentive Compatibility

We resume the analysis of problem ($\mathcal{P}$), and the next step is to characterize the incentive constraints in a more tractable way. We first take the usual approach of characterizing (IC-$1_{s,\hat{s}}$) as an envelope and a monotonicity constraint.

Lemma 2.2. *(IC-$1_{s,\hat{s}}$) is equivalent to the following two conditions:*

1. $U^{'}(s) = -[u_L(s) - u_H(s)]$ *a.e.;*
2. $u_L(s) - u_H(s)$ *is decreasing.*

The second step is to show that (IC-2_s^L) binds at almost every $s \in [\underline{s}, \overline{s}]$. Such a result is useful because when (IC-2_s^L) binds, we can write $u_L(s)$ as a function of $u_H(s)$ and $e_H(s)$. In particular, as $U(s)$ is absolutely continuous once can write

$$U(s) = su_H(s) + (1-s)u_L(s) = \overline{s}u_H(\overline{s}) + (1-\overline{s})u_L(\overline{s}) + \int_s^{\overline{s}}[u_L(z) - u_H(z)]dz.$$

Proposition 2.1. *Take a mechanism that satisfies all constraints in problem $\mathcal{P}$. Suppose that there exists a positive mass set $\hat{S}$ of signals such that (IC-2_s^L) is slack for all $s \in \hat{S}$. Then, there exists another mechanism that strictly reduces the principal's expected procuring cost.*

The proof takes two steps: first, we show that whenever (IC-2_s^L) is slack, it must be the case that effort is not distorted from the first-best. Second, we show that any mechanism in which (IC-2_s^L) does not bind for almost every $s \in [\underline{s}, \overline{s}]$ can be strictly improved on. For the first step, we start from a mechanism in which for a positive mass set $\hat{S}$ we have (IC-2_s^L) slack and $e_H(s) \neq e^{FB}$. We, then, construct an alternative mechanism that satisfies all constraints and has strictly lower expected procuring costs. Therefore, we can restrict attention to mechanisms that have $e_H(s) = e^{FB}$ whenever (IC-2_s^L) is slack. Similarly, the second step takes a constructive approach. We start from a mechanism in which (IC-2_s^L) is slack for a positive mass set $\tilde{S}$, and then construct a strict improvement.

By Proposition 2.1, we know (IC-2^L_s) binds almost everywhere; hence

$$u_L(s) = \phi(e_H(s)) + u_H(s). \tag{1}$$

Therefore, we can re-write principal's problem as

$$\mathcal{P}' : \quad \min_{(e_L(s),e_H(s),u_H(s))} \int_{\underline{s}}^{\overline{s}} \Big\{(1-s)\Big[\psi(e_L(s)) - e_L(s)\Big] + s\Big[\psi(e_H(s)) - e_H(s)\Big] + \Big[s\beta_H + (1-s)\beta_L\Big] + \Big[(1-s)\phi(e_H(s)) + u_H(s)\Big]\Big\} f(s)ds$$

subject to

$$e_H(s) \text{ decreasing}, \tag{monotonicity}$$

$$su_H(s) + (1-s)u_L(s) = \overline{s}u_H(\overline{s}) + (1-\overline{s})u_L(\overline{s}) + \int_s^{\overline{s}} \phi(e_H(z))dz, \tag{envelope}$$

$$\phi(e_L(s) + \Delta\beta) \geq \phi(e_H(s)), \tag{IC-2^H_s}$$

$$u_H(s) \geq \bar{u}. \tag{IR-2_s}$$

3 High Ex-post Reservation Utility ($\bar{u} = 0$)

Suppose that the firm has a high-valued ex-post outside option ($\bar{u} = 0$), i.e., it cannot be made worse off than her ex-ante outside option in any cost scenario. We will show that it is optimal for the principal to not screen the ex-ante information in such a case.[14]

The first immediate observation is that when the firm has a high-valued ex-post outside option, the ex-post participation constraint implies the ex-ante participation constraint. That is, $u_L(s) \geq u_H(s) \geq 0$ implies

$$U(s) = (1-s)u_L(s) + su_H(s) \geq 0, \text{ for all } s \in [\underline{s}, \overline{s}].$$

The Pooling Mechanism

Before solving problem ($\mathcal{P}'$), we define a mechanism that disregards first-period information and screens the firm's type but not signal.

Definition (optimal pooling mechanism)**.** The *optimal pooling mechanism* is defined by

$$\Big\{\big(\bar{e}_H, \bar{u}\big), \big(\bar{e}_L, \bar{u} + \phi(\bar{e}_H)\big)\Big\},$$

[14]All results immediately hold for $\bar{u} > 0$. We focus on $\bar{u} \leq 0$ because then $-\bar{u}$ has the natural interpretation of the cost of breaking the contract.

where $\bar{e}_H$ and $\bar{e}_L$ are given by

$$\begin{aligned} \psi'(\bar{e}_L) &= 1 \\ \psi'(\bar{e}_H) &= 1 - \frac{1-s^*}{s^*}\phi'(\bar{e}_H) \end{aligned} \tag{2}$$

and $s^* = \int_{\underline{s}}^{\overline{s}} sf(s)ds$ is the expected unconditional probability of an expensive project β_H.

It is important to highlight that the allocation of the pooling mechanism is different from the second-best allocation. When computing the second-best allocation, we allow the principal to observe s, and, hence, directly condition the contracts on s. While in the pooling mechanism, the principal does not observe s and offers the same allocation regardless of s. That is, it pools all s reports.

The pooling mechanism is the solution to the optimal procurement problem with two project types (Laffont and Tirole, 1993, Section 1.3), in which the procurer agency believes the low-cost project β_L happens with probability $1 - s^*$. Hence, the pooling allocation corresponds to *static* one since it does not depends on the first-period information and conditions the mechanism solely on second-period reports. The following theorem establishes that the pooling mechanism is optimal. We first state the theorem, present the proof, and subsequently discuss its intuition which stems directly from the proof.

Theorem 3.1. *The optimal pooling mechanism is the solution to problem $\mathcal{P}'$. That is, it is optimal for the procuring agency to refrain from screening the first-period information and offer a cost-plus contract to the firm.*

Proof of Theorem 3.1. The proof follows four steps:

1. Consider a relaxed problem that drops two constraints: ($envelope$) and (IC-2_s^H).
2. To solve the problem in step 1, we further relax the problem by dropping the ($monotonicity$) constraint and subsequently ironing the solution in the appropriate regions.
3. Show that the solution to the further relaxed problem (step 2) violates ($monotonicity$) throughout the entire support. Hence, the solution to the problem in step 1 requires ironing everywhere.[15]
4. Find the solution to step 1 considering that we must pool all signals (due to step 3), and check that ($envelope$) and (IC-2_s^H) are satisfied.

Step 1: Consider the following relaxed problem:

$$\begin{aligned} \min_{(e_L(s), e_H(s), u_H(s))} \int_{\underline{s}}^{\overline{s}} \Big\{ &(1-s)\Big[\psi(e_L(s)) - e_L(s)\Big] + s\Big[\psi(e_H(s)) - e_H(s)\Big] \\ &+ \Big[s\beta_H + (1-s)\beta_L\Big] + \Big[(1-s)\phi(e_H(s)) + u_H(s)\Big]\Big\} f(s)ds \end{aligned} \tag{3}$$

[15]Such approach was first proposed by Castro-Pires and Moreira (2021) in a static setting with moral hazard and adverse selection.

subject to

$$e_H(s) \text{ decreasing,} \tag{monotonicity}$$

$$u_H(s) \geq 0. \tag{IR-2_s}$$

The first two immediate observations are that, in this relaxed problem, it is optimal to set $u_H(s) = 0$ and $e_L(s) = e^{FB}$ for all $s \in [\underline{s}, \overline{s}]$. Hence, the relaxed problem becomes

$$\min_{e_H(s)} \int_{\underline{s}}^{\overline{s}} \Big\{ s\Big[\psi(e_H(s)) - e_H(s)\Big] + (1-s)\phi(e_H(s))\Big\} f(s)ds \tag{4}$$

subject to

$$e_H(s) \text{ decreasing.} \tag{monotonicity}$$

Step 2: Then, we follow the usual approach for solving problems with a monotonicity constraint. We relax the monotonicity constraint and subsequently iron the solution in the appropriate regions.

Note, however, that when we further drop the monotonicity constraint, we have fully relaxed incentive compatibility in the first period. Hence, not surprisingly, the solution to such further relaxation is given by the second-best allocation:

$$\psi'(e_H^{SB}(s)) = 1 - \frac{1-s}{s}\phi'(e_H^{SB}(s)) \text{ for all } s \in [\underline{s}, \overline{s}].$$

Step 3: Note that $e_H^{SB}(\cdot)$ is strictly increasing, i.e.,

$$\frac{de_H^{SB}(s)}{ds} = \frac{\phi'\big(e_H^{SB}(s)\big)}{s^2\big[\psi''\big(e_H^{SB}(s)\big) + \frac{1-s}{s}\phi'\big(e_H^{SB}(s)\big)\big]} > 0.$$

Hence, introducing back the monotonicity constraint requires ironing over the entire support.[16] Therefore, the solution to the problem in step 1 requires a constant $e_H(\cdot)$.

Step 4: We, then, solve the problem in step 1, considering that $e_H(\cdot)$ must be constant. That is, we solve

$$\min_{e_H \in [0, \beta_H]} \Big\{ s^*\Big[\psi(e_H) - e_H\Big] + (1-s^*)\phi(e_H)\Big\},$$

which delivers the pooling mechanism allocation. Taking stock of everything, the solution to the problem in step 1 becomes $u_H^*(s) = 0$, $u_L^*(s) = \phi(e_H^*(s))$, $e_L^*(s) = e^{FB}$, and $e_H^*(s) = \bar{e}_H$, where

$$\psi'(\bar{e}_H) = 1 - \frac{1-s^*}{s^*}\phi'(\bar{e}_H). \tag{5}$$

It remains to show that the two relaxed constraints are satisfied. For (IC-2_s^H), note that $e_H^*(s) < e_L^*(s) = e^{FB}$. Hence,

$$\phi(e_L^* + \Delta\beta) \geq \phi(e_H^*(s)).$$

Finally, note that all signal reports are pooled together under the pooling mechanism. Therefore, incentive compatibility in the first period is trivially satisfied; hence, it is also the ($envelope$) condition. □

[16] See Toikka (2011)'s separable case for a generalization of Myerson (1981)'s ironing approach.

Intuitively, the more the principal distorts the agent's effort downwards in the high-cost state, the lowest the informational rents left to low-cost agents. Recall that the higher the signal, the less likely the agent will have a low cost. Hence, the less important is to reduce low-cost type's information rents. Setting aside (IC-$1_{s,\hat{s}}$), the principal would like to distort less $e_H(s)$ when s is higher. However, eliciting the ex-ante information requires the principal to distort more the effort choice in case of high costs for agents with higher signals — the opposite of what is optimal absent the ex-ante incentive compatibility constraint. This conflict is precisely what the literature on static screening denominates as non-responsiveness. Given non-responsiveness, the best for the principal is not to screen the ex-ante information.

The critical implication of Theorem 3.1 is that the optimal mechanism does not condition on first-period but only second-period reports. The result implies that in this setting, optimal procurement contracts should not depend on ex-ante self-reported estimated costs but only on realized verifiable ex-post costs. That is, the optimal mechanism has a cost-plus structure in which the firm is reimbursed for its realized costs and receives an additional net transfer to incentivize cost reduction. Such a net transfer is smaller the higher the realized costs. The following subsection presents a simple indirect implementation of such a mechanism.

Indirect Implementation

The previous subsection characterized the optimal direct and truthful revelation mechanism. It is useful to examine an indirect implementation closely resembling contracts observed in practice. In particular, as the optimal contract does not condition on the first-period signals, one can write the optimal mechanism as a transfer function contingent on realized costs. If the realized cost is C, the firm receives $T(C)$.

Given $\bar{e}_L$ and $\bar{e}_H$, denote $\bar{C}_L := \beta_L - \bar{e}_L$ and $\bar{C}_H := \beta_H - \bar{e}_H$. A transfer function that implements the same outcomes as the optimal direct and truthful revelation mechanism is

$$T(C) = \begin{cases} \bar{C}_L + \phi(\bar{e}_H) + \psi(\bar{e}_L) & \text{if} \quad C \leq \bar{C}_L, \\ T(\bar{C}_L) + (C - \bar{C}_L)[1 - \psi'(\bar{e}_H)] & \text{if} \quad C \in (\bar{C}_L, \bar{C}_H], \\ T(\bar{C}_H) + C - \bar{C}_H & \text{if} \quad C > \bar{C}_H. \end{cases}$$

Graphically, the transfer function is represented in figure 1.

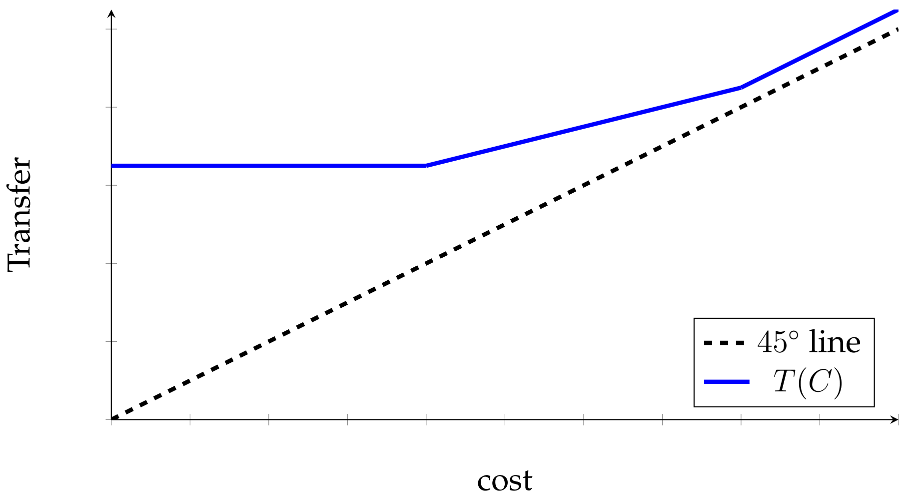

Figure 1: Indirect Implementation.

4 Negative Ex-post Reservation Utility ($\bar{u} < 0$)

The analysis so far addressed the cases in which the ex-ante participation constraint is redundant in the presence of the ex-post participation constraint. We now allow for the ex-post outside option to be worse than the ex-ante ($\bar{u} < 0$). This section shows that the result of not screening the ex-ante information remains valid as long as the ex-post outside option is not substantially lower than the ex-ante. The results are divided into three cases: in the first, the ex-post outside option is sufficiently close to the ex-ante, and our previous results immediately apply, with the ex-ante participation constraint being slack. In the second case, the ex-post outside option takes intermediate values, and both ex-ante and ex-post participation constraints bind at the optimal mechanism. However, even in this case, it is optimal not to screen the first-period information. The third case analyzes the situation in which the ex-post outside option is sufficiently low, which implies that only the ex-ante participation binds. In this last case, it is optimal for the principal to fully screen first-period signals under a mild regularity condition.

4.1 Negative but High Ex-post Outside Option Values

When $\bar{u} < 0$, the ex-post participation (IR-2) holding does not necessarily imply the ex-ante participation (IR-1) also does. The ex-ante expected utility of an agent who reports truthfully under the pooling mechanism is given by

$$\bar{U}(s) = (1-s)u_L(s) + su_H(s) = \bar{u} + (1-s)\phi(\bar{e}_H),$$

which is bigger than zero (the normalized ex-ante outside option value) for any $s \in [\underline{s}, \overline{s}]$ only if $\bar{u}$ is sufficiently large. Note that if $\bar{U}(\overline{s}) \geq 0$, then the optimal pooling mechanism (2) satisfies (IR-1). Hence, it is also optimal when we add the ex-ante participation constraint. Define

$$\bar{u}_1 := -(1-\overline{s})\phi(\bar{e}_H). \tag{6}$$

Therefore, the following result is immediate.

Corollary 1. *The optimal pooling mechanism is optimal for any ex-post outside option* $\bar{u} \geq \bar{u}_1$.

4.2 Intermediate Ex-post Outside Option Values

We now analyze cases in which the ex-post outside option is lower than $\bar{u}_1$. We, then, show that a modified pooling mechanism is optimal as long as $\bar{u}$ is not too low.

Definition (Optimal pooling with binding (IR-1) mechanism)**.** The *optimal pooling with binding (IR-1) mechanism* is

$$\Big\{\big(\tilde{e}_H, \bar{u}\big), \big(\tilde{e}_L, \bar{u} + \phi(\tilde{e}_H)\big)\Big\},$$

where $\tilde{e}_H$ and $\tilde{e}_L$ are given by:

$$\begin{aligned} &\psi'(\tilde{e}_L) = 1, \\ &\phi(\tilde{e}_H) = -\frac{\bar{u}}{1 - \overline{s}}. \end{aligned} \tag{7}$$

The pooling with binding (IR-1) mechanism pools all ex-ante signals but ensures that ex-ante participation is satisfied. Note that when $\bar{u}$ is low (lower than $\bar{u}_1$), the agents with high signals would not be willing to participate in the pooling mechanism. The pooling with binding (IR-1) mechanism limits the effort distortion for high-cost firms to the maximum such that they would be willing to participate even if they have received the worst possible signal. The following theorem shows the pooling with binding (IR-1) mechanism is optimal for a range of ex-post outside option values $\bar{u} \in [\bar{u}_2, \bar{u}_1]$, for some $\bar{u}_2 < \bar{u}_1$.

Theorem 4.1. *There exists* $\bar{u}_2 < \bar{u}_1$ *such that the pooling with binding (IR-1) mechanism is optimal for any* $\bar{u} \in [\bar{u}_2, \bar{u}_1]$.

Theorem 4.1 extends the result that the principal prefers not to screen the first-period information to cases in which the ex-ante participation also binds. For $\bar{u} < \bar{u}_1$, the pooling mechanism does not ensure the participation of firms with all signals. Hence, to ensure ex-ante participation, the principal must increase the ex-ante utility of the firm with the worst signal (and, consequently, the utility of firms with all other signals). One way to increase the firm's ex-ante utility is to reduce distortions in the recommended efforts. In particular, to reduce distortions introduced precisely to limit information rents. That is, by increasing e_H, and consequently increasing u_L, the principal reduces the effort distortion and leaves additional rents to the firm. The pooling with binding (IR-1) mechanism reduces such distortions until even the firm with the worst signal is willing to participate.

Theorem 4.1 highlights that the main force that prevents screening is not the slackness of (IR-1) but (IR-2) be binding. Overall, the central message of Theorem 4.1 is similar to the one of Theorem 3.1: for sufficiently high ex-post outside option, the principal is better off by not screening the ex-ante information. The proof, however, is very different. The approach is to address the design problem with a guess-and-verify algorithm: we first guess that the *pooling with binding (IR-1) mechanism* is a solution to problem $\mathcal{P}$. Then, we construct Lagrange multipliers that sustain such a mechanism as a solution to the procuring agency's problem.

Proof of Theorem 4.1. Suppose that $-(1-\overline{s})\phi(e^{FB}) \leq \bar{u} < \bar{u}_1$. Notice that the optimal pooling with binding (IR-1) mechanism described in (7) satisfies all (IC-$1_{s,\hat{s}}$) constraints with equality. Also, it immediately follows that it satisfies (IC-2_s^L) and (IC-2_s^H). Hence, it suffices to focus on ex-post participation constraints (IR-2_s), (IR-1_s), and optimality conditions.

We present a solution algorithm that proceeds according to the following steps:

1. State a modified version of the principal's problem, denoted by $\mathcal{R}^{\bar{u}}$, which exogenously sets ex-post profits $u_H(s)$ to $\bar{u}$;

2. Guess the set of active (IC-$1_{s,\hat{s}}$) constraints at the pooling with binding (IR-1) solution, and define a further relaxed problem that ignores the remaining inactive constraints, denoted by $\mathcal{R}^{\bar{u}'}$;

3. Construct problem $\mathcal{R}^{\bar{u}'}$'s Lagrangian and present a Karush-Kuhn-Tucker argument to the solution taking the form of the pooling with binding (IR-1) mechanism;

4. Extend the Lagrangian to include the ex-post IR constraints and find multipliers that sustain the pooling with binding (IR-1) mechanism and $u_H^*(s) = \bar{u}$ as an optimal solution.

Step 1. First, we fix $u_H(s) = \bar{u} \geq -(1-\overline{s})\phi(e^{FB})$. Moreover, using Proposition 2.1 we set $u_L(s) = u_H(s) + \phi(e_H(s))$. Then, the modified principal's design problem becomes:

$$\begin{aligned}
\mathcal{R}^{\bar{u}}: \quad & \min_{\{e_L(\cdot), e_H(\cdot)\}} \int_{\underline{s}}^{\overline{s}} \{(1-s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\
& \qquad\qquad + \bar{u} + s\left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s) ds \\
\text{s.t.} \quad & (1-s)\left[\phi(e_H(s)) - \phi(e_H(\hat{s}))\right] \geq 0, \quad \forall \hat{s}, s \\
& \bar{u} + (1-s)\phi(e_H(s)) \geq 0 \quad \forall s.
\end{aligned}$$

Step 2. We then relax problem $\mathcal{R}^{\bar{u}}$ by restricting the set of first-period incentive constraints to consider. We first define a reference signal $\tilde{s}$. Then, we restrict attention to incentive constraints in which a firm with a signal below the reference ("good type") does not envy the allocation assigned to a signal above the reference ("bad type").

Let the reference signal $\tilde{s}$ be such that

$$\frac{1-s''}{s''} \leq \frac{1-\psi'(\tilde{e}_H)}{\phi'(\tilde{e}_H)} \leq \frac{1-s'}{s'} \quad \forall s' < \tilde{s} < s''.$$

Definition. Define the restricted set of incentive constraints as

$$\tilde{IC} = \{\text{IC-1}_{s,\hat{s}} : s < \tilde{s} < \hat{s}\}.$$

Remark 1. If the optimal pooling with binding (IR-1) mechanism solves problem $\mathcal{R}^{\bar{u}}$ restricted to $\tilde{IC}$, then it solves the problem ($\mathcal{R}^{\bar{u}}$) with all the (IC-$1_{s,\hat{s}}$), as (7) trivially satisfies all neglected incentive constraints.

Hence, it suffices to show that the optimal pooling mechanism is a solution to the following version of the relaxed problem:

$$\begin{aligned}
\mathcal{R}^{\bar{u}'}: \quad \min_{\{e_L(\cdot),e_H(\cdot)\}} \int_{\underline{s}}^{\overline{s}} & \{(1-s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\
& + \bar{u} + s\left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s)ds \\
\text{s.t. } & (1-a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] \geq 0, \quad \forall a < \tilde{s} < b \\
& \bar{u} + (1-s)\phi(e_H(s)) \geq 0 \quad \forall s.
\end{aligned}$$

Step 3. Note that the Karush-Kuhn-Tucker theorem for function spaces (Luenberger, 1997, p. 220) applies to concave optimization problems (i.e., convex objectives and feasibility sets for minimization problems). However, the constraints (IC-$1_{s,\hat{s}}$) in problem ($\mathcal{R}^{\bar{u}'}$) are differences of convex functions $\phi(\cdot)$ and are not necessarily convex. The following Lemma 4.1 establishes that the optimization problem $\mathcal{R}^{\bar{u}'}$ satisfies the Karush-Kuhn-Tucker theorem conditions.

Lemma 4.1. *Problem $\mathcal{R}^{\bar{u}'}$'s objective function is concave, its (IC) constraints are quasi-concave functions, and the interior of the feasible set is non-empty.*

By the quasi-concave version of the Karush-Kuhn-Tucker theorem (Arrow and Enthoven, 1961, Theorem 1), a pair $\{e_L(s), e_H(s)\}$ is a solution to problem $\mathcal{R}^{\bar{u}'}$ if there exist positive multipliers $\lambda : [\underline{s}, \tilde{s}] \times [\tilde{s}, \overline{s}] \to \mathbb{R}_+$, $\gamma : [\underline{s}, \overline{s}] \to \mathbb{R}_+$ so that $\{e_L(s), e_H(s)\}$ minimizes the Lagrangian:

$$\begin{aligned}
\mathcal{L} = & \int_{\underline{s}}^{\overline{s}} \{(1-s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] + \bar{u} + s\left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s)ds \\
& - \int_{\underline{s}}^{s^*} \int_{s^*}^{\overline{s}} \lambda(a,b)(1-a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] dbda - \int_{\underline{s}}^{\overline{s}} \gamma(s)\big[\bar{u} + (1-s)\phi\big(\big(e_H(s)\big)\big]ds.
\end{aligned}$$

The minimizer of $\mathcal{L}$ must satisfy the following pointwise first-order conditions (FOC):

$$[e_L(s)]: \quad f(s)(1-s)\left(-1+\psi'(e_L(s))\right) = 0$$

$$[e_H(s)]: \quad \begin{cases} f(s)s\left(-1+\psi'(e_H(s))\right) + f(s)(1-s)\phi'(e_H(s)) - \displaystyle\int_{\tilde{s}}^{\overline{s}} \lambda(s,b)(1-s)\phi'(e_H(s))db \\ \qquad -\gamma(s)(1-s)\phi'(e_H(s)) = 0 \quad \text{if } s \in [\underline{s}, \tilde{s}] \\ f(s)s\left(-1+\psi'(e_H(s))\right) + f(s)(1-s)\phi'(e_H(s)) + \displaystyle\int_{\underline{s}}^{\tilde{s}} \lambda(a,s)(1-a)\phi'(e_H(s))da \\ \qquad -\gamma(s)(1-s)\phi'(e_H(s)) = 0 \quad \text{if } s \in [\tilde{s}, \overline{s}]\,. \end{cases}$$

The existence of a multipliers $\lambda(a,b), \gamma(s) \geq 0$ that satisfy the preceding FOC system is established in the following Proposition 4.1. Therefore, by Remark 1, the optimal pooling mechanism also solves problem $\mathcal{R}^{\bar{u}}$.

Proposition 4.1. *There exist functions $\tilde{\lambda} : [\underline{s}, \tilde{s}] \times [\tilde{s}, \overline{s}] \to \mathbb{R}_{+}$, and $\tilde{\gamma} : [\underline{s}, \overline{s}] \to \mathbb{R}_{+}$ such that the pooling allocation $\{\bar{e}_L, \bar{e}_H\}$ is an optimal solution to problem $\mathcal{R}^0$, with multipliers given by $\tilde{\lambda}(s,\hat{s})$ and $\tilde{\gamma}(s)$.*

Step 4. The last step is to consider the Lagrangian that explicitly incorporates the ex-post participation constraints (IR-2):

$$\begin{aligned} \tilde{\mathcal{L}} = & \int_{\underline{s}}^{\overline{s}} \Big\{(1-s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\ & \qquad + u_H(s) + s\left[\beta_H - e_H(s) + \psi(e_H(s))\right]\Big\} f(s)ds \\ & - \int_{\underline{s}}^{\tilde{s}} \int_{\tilde{s}}^{\overline{s}} \lambda(a,b)\left\{(1-a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] + u_H(a) - u_H(b)\right\} dbda \\ & - \int_{\underline{s}}^{\overline{s}} \gamma(s)\big[u_H(s) + (1-s)\phi(\big(e_H(s)\big)\big]ds - \int_{\underline{s}}^{\overline{s}} \mu(s)[u_H(s) - \bar{u}]ds. \end{aligned}$$

By the same argument in Arrow and Enthoven (1961), if there exist positive multipliers $\lambda(a,b) \geq 0$, $\gamma(s) \geq 0$ and $\mu(s) \geq 0$ such that Lagrangian $\tilde{\mathcal{L}}$ is minimized at $\{\tilde{e}_L, \tilde{e}_H\}$ and $\tilde{u}_H(s) = \bar{u}$, $\forall s$, then the pooling mechanism is a solution to problem $\mathcal{P}$. Proposition 4.2 establishes the existence of such multipliers and concludes the proof of Theorem 4.1.

Proposition 4.2. *There exists $\bar{u}_2 < \bar{u}_1$, such that for any $\bar{u} \in [\bar{u}_2, \bar{u}_1)$, there exist functions $\tilde{\lambda} : [\underline{s}, \tilde{s}] \times [\tilde{s}, \overline{s}] \to \mathbb{R}_{+}$ and $\tilde{\mu} : [\underline{s}, \overline{s}] \to \mathbb{R}_{+}$ such that the pooling with binding IR allocation $\{\tilde{e}_L, \tilde{e}_H\}$ and the ex-post profit level $\tilde{u}_H = \bar{u}$ are a solution to problem $\mathcal{P}$, with multipliers given by $\tilde{\lambda}(s,\hat{s})$, $\tilde{\gamma}(s)$ and $\tilde{\mu}(s)$.*

□

4.3 Low Ex-post Outside Option Values

We now analyze the case in which the ex-post outside option is sufficiently low, such that it becomes slack. We first solve the problem without (IR-2) and then we find $\bar{u}_3$ such that for all $\bar{u} \leq \bar{u}_3$, (IR-2) is slack.

Note that Proposition 2.1 holds even without (IR-2). Hence, we know that (IC-2_s^L) must bind. As usual in mechanism design without an ex-post participation constraint, the ex-ante participation constraint binds for the least efficient firm (in this case, the one with the highest signal $\overline{s}$). Hence, by $U(\overline{s}) = 0$ and Lemma 2.2 we have

$$U(s) = \int_s^{\overline{s}} \phi(e_H(t))dt. \tag{8}$$

We can write the principal's problem without (IR-2) as

$$\mathcal{P}_1 : \quad \min_{(e_L(s), e_H(s))} \int_{\underline{s}}^{\overline{s}} \Big\{(1-s)\Big[\psi(e_L(s)) - e_L(s)\Big] + s\Big[\psi(e_H(s)) - e_H(s)\Big] + \int_s^{\overline{s}} \phi(e_H(t))dt\Big\} f(s)ds$$

subject to

$$e_H(s) \text{ decreasing.}$$

Proposition 4.3. *Suppose $\frac{F(s)}{sf(s)}$ is increasing. The solution $\big(\hat{e}_L(s), \hat{e}_H(s)\big)$ to problem $\mathcal{P}_1$ is characterized by*

$$\psi'\big(\hat{e}_L(s)\big) = 1 \quad \text{and} \quad \frac{1 - \psi'\big(\hat{e}_H(s)\big)}{\phi'\big(\hat{e}_H(s)\big)} = \frac{F(s)}{sf(s)} \quad \text{for all } s \in [\underline{s}, \overline{s}].$$

We take the usual approach of relaxing the monotonicity constraint and then checking it is satisfied. Whenever $\frac{F(s)}{sf(s)}$ is increasing, the solution satisfies the relaxed monotonicity constraint. Moreover, whenever $\frac{F(s)}{sf(s)}$ is strictly increasing, there is full separation of all ex-ante signals.

Note that (IC-$1_{s,\hat{s}}$) implies that $u_H(s)$ must be increasing and by (IC-2_s) we know $u_H(s) < u_L(s)$. Hence, an agent with the lowest signal and highest type realization achieves the lowest ex-post utility. That is,

$$\hat{u}_H(\underline{s}) = \hat{U}(\underline{s}) - (1-\underline{s})\phi\big(\hat{e}_H(\underline{s})\big) = \int_{\underline{s}}^{\overline{s}} \phi\big(\hat{e}_H(t)\big)dt - (1-\underline{s})\phi\big(\hat{e}_H(\underline{s})\big),$$

where the first equality comes from (IC-2_s^L) and the second from (8).

Corollary 2. *Suppose $\frac{F(s)}{sf(s)}$ is increasing and $\bar{u} \leq \hat{u}_H(\underline{s})$. Then, (IR-2) is slack, and the optimal mechanism is characterized by Proposition 4.3.*

In summary, we have established the pattern of optimal contracts for different levels of $\bar{u}$. If the ex-post outside option value is high enough ($\bar{u} \geq \bar{u}_1$), we have full pooling of the first-period signals and a slack ex-ante participation constraint. Suppose the ex-ante outside option takes an intermediary value ($\bar{u}_2 \leq \bar{u} < \bar{u}_1$), then the pooling of the first-period signals is preserved but with a binding ex-ante participation constraint. Finally, for low enough ex-post outside options ($\bar{u} \leq \bar{u}_3$), there is full-screening under a regularity condition. Below, we graphically illustrate the result with an example with two ex-ante signals.[17]

Figure 2 displays the three aforementioned thresholds.[18] At the level of ex-post outside option $\bar{u}_1$, the expected rent of the high-signal firm attains zero, so (IR-1) becomes active, and the pooling with binding (IR-1) mechanism is optimal.

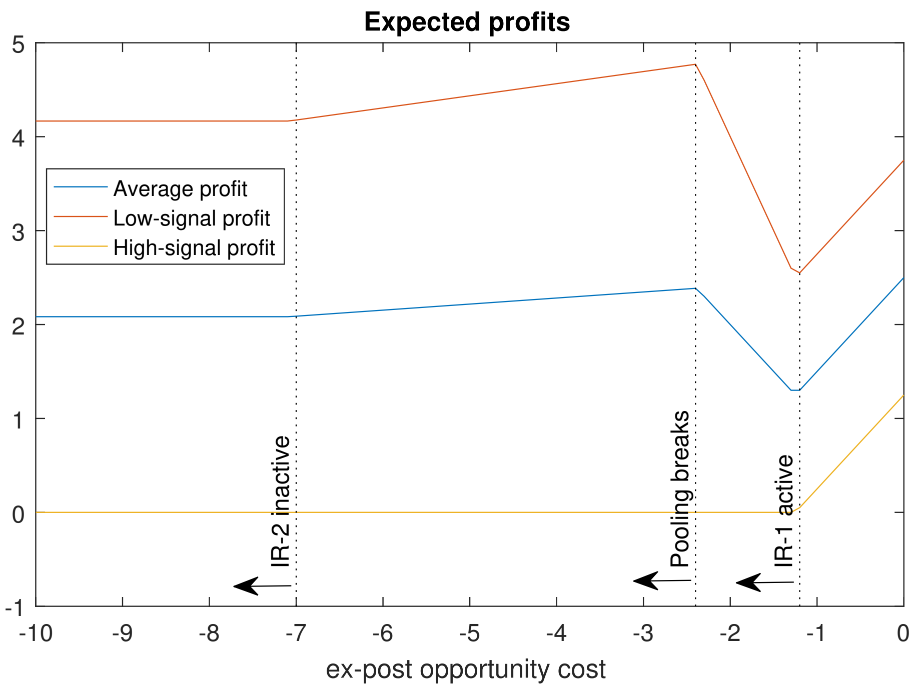

Figure 2: Optimal ex-ante profits in response to $\bar{u}$.

Pooling the first-period signals is no longer optimal for ex-post outside options below $\bar{u}_2$, as the regulator benefits from separating each signal's allocation and reducing the low signal firm information rents. At $\bar{u}_3$, (IR-2) becomes inactive.

[17]In Appendix A, we prove Theorem 3.1 using a constructive approach similar to Theorem 4.1. This approach does not rely on the first-period signal being continuous, and the result holds for any distribution of s regardless of whether it is continuous, discrete, or any combination of the two. We maintain the non-responsiveness approach in the main text because it has the clearest intuition.

[18]The online appendix C describes the functional forms used for the numerical examples and provides the Matlab codes.

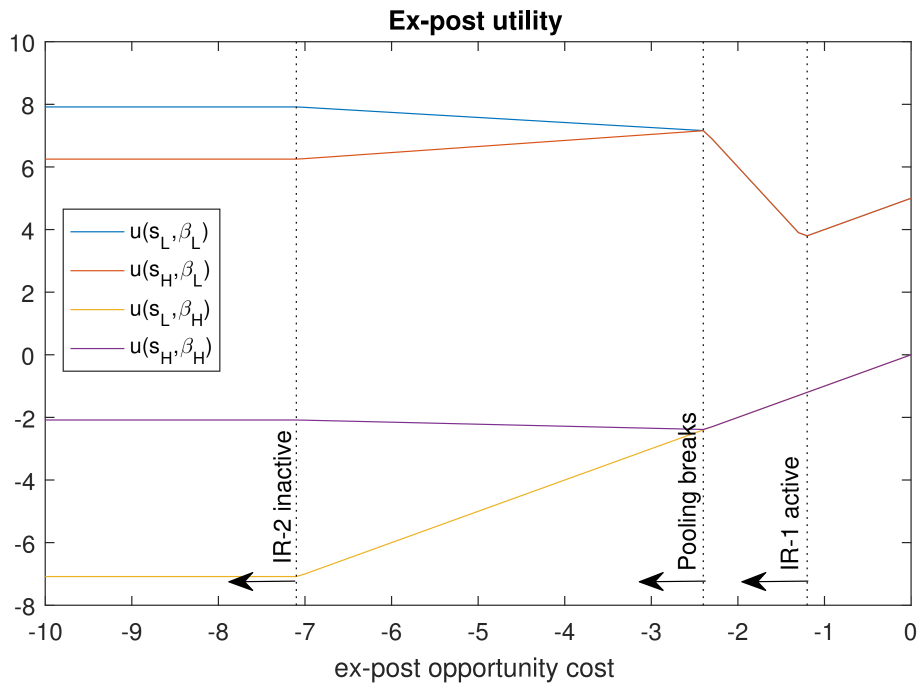

Figure 3: Optimal ex-post profits in response to $\bar{u}$.

Figure 3 shows the optimal ex-post utility level as a function of the ex-post outside option value. For $\bar{u} > \bar{u}_1$, the ex-ante constraint (IR-1) is inactive, the ex-post utility of types β_H and β_L are parallel, as the difference $u_L(s) - u_H(s) = \phi(\bar{e}_H)$ is constant and given by (IC-2_s^L). When (IR-1) becomes active, the two ex-post profit levels begin to distance one from another. The optimal mechanism becomes the optimal pooling with binding (IR-1), which holds type β_H at the ex-post outside option but leaves additional rents to β_L by decreasing the distortions on the effort of the high type firm. The optimal mechanism pools first-period signals while holding the high-cost type's profit at the ex-post outside option and providing additional informational rents to the low-cost type to ensure ex-ante participation. For sufficiently low $\bar{u}$, pooling breaks down, and the principal discriminates profit levels among all report sequences (s, β_i). Finally, when (IR-2) becomes inactive, the levels characterized in Proposition 4.3 describe the ex-post profits.

4.4 Cost Overruns

In the real world, we often observe cost overruns in procurement projects; see Makovšek and Bridge (2021) for a survey of the empirical literature. From a theoretical perspective, the emergence of cost overruns is typically attributed to failures of commitment in the contracting environment. For instance, Arvan and Leite (1990) present an adverse selection model with sequential uncorrelated cost-types, where the principal cannot commit not to cancel the project should it reveal unfavorable. In equilibrium, the more cost-uncertain tasks are left to the tail end of the project, giving rise to possible overruns. A different theoretical strand attributes cost overruns to incomplete contracts, see Bajari and Tadelis (2001).

This section lays a forth potential explanation for cost overruns not from non-commitment or incomplete contracting, but from the procuring agency failing to anticipate the extent of the firm's right to exit the partnership. We show that, in our model, if the procuring agency underestimates the firm's ex-post outside option, then the firm would have an incentive to underreport its cost estimate and exercise its outside option if the project has a high-cost realization.

In Subsection 4.3, we have characterized optimal contracts absent (IR-2). Suppose now that the principal is miscalibrated and believes that $\bar{u} \leq \bar{u}_3$, but the true outside option is $\bar{u} = 0$. By Proposition 4.3, we know the principal offers the mechanism $\Big\{\big(\hat{e}_i(\cdot), \hat{u}_i(\cdot)\big)\Big\}_{i\in\{L,H\}}$. Note that

$$\begin{aligned}\hat{u}_H(s) &= \int_s^{\overline{s}} \phi\big(\hat{e}_H(t)\big)dt - (1-s)\phi\big(\hat{e}_H(s)\big) \\ &= \int_s^{\overline{s}} \Big[\phi\big(\hat{e}_H(t)\big) - \phi\big(\hat{e}_H(s)\big)\frac{1-s}{\overline{s}-s}\Big]dt < 0,\end{aligned}$$

where the inequality comes from the fact that $\hat{e}_H(s)$ is decreasing and $\overline{s} < 1$.

As $\hat{u}_H(s) < 0$, the firm — regardless of what signal it has reported — prefers to take its ex-post outside option in case the type realization is β_H. Hence, regardless of its signal report, the firm gets a payoff of 0 when the project's cost is high. Note, however, that $\hat{u}_L(s) > 0$ is decreasing in s. Hence, reporting $\underline{s}$ is strictly better than any other report for any realization of the signal s. That is, when the procuring agency underestimates the firm's ex-post outside option, the firm has the incentive to under-report estimated costs (report the lowest signal $\underline{s}$) and take the ex-post outside option in case of a high-cost realization.

5 Multiple Firms

We have assumed so far that only one firm is available for the project. Such an assumption is appropriate in many scenarios but not in others.[19] In this section, we explore whether sufficient competition might allow the procurement agency to screen firms based on their ex-ante information.

We extend our model to accommodate multiple firms competing for the same project. We assume that any one of n firms can execute the project. As before, each firm has an ex-post project cost of $C^i = \beta^i - e^i$, where $\beta^i \in \{\beta_L, \beta_H\}$ and $e^i \in [0, \beta_H]$, and exerting effort e^i generates a disutility $\psi(e^i)$ for the firm. In period 1, each firm receives a private signal $s^i \in [\underline{s}, \overline{s}]$ about its own type, where $Pr(\beta^i = \beta_H | s^i) = s^i$. The principal, then, must choose one firm to execute the project. In period 2, the selected firm learns β^i and decides whether to complete the project or take an ex-post outside option $\bar{u} = 0$.

[19]For instance, the Buy American Act requires the United States government to exclusively buy domestic products when available. In some industries, such as naval or aerospace, often such restriction implies that a single firm qualifies as a candidate for the project. In other less concentrated sectors, one can observe more competition.

Firms' signals and types are assumed to be independent and identically distributed across firms.

An important remark is that only the selected firm learns its intrinsic cost. The interpretation is that the firm only fully learns a project's cost when it is about to execute it. In this setting, a direct and truthful revelation mechanism is a set of functions $x_i(\overrightarrow{s})$, $e^i_L(\overrightarrow{s})$, $e^i_H(\overrightarrow{s})$, $u^i_L(\overrightarrow{s})$, $u^i_H(\overrightarrow{s})$, $\ell_i(\overrightarrow{s})$ inducing truth-telling. $\overrightarrow{s}$ is the vector of announced signals, $x_i(\overrightarrow{s})$ is the probability firm i is selected, $e^i_j(\overrightarrow{s})$ is the recommended effort if firm i is selected and it is of type $\beta_j \in \{\beta_L, \beta_H\}$, $u^i_j(\overrightarrow{s})$ is the ex-post utility of firm i if selected when it is of type β_j, and $\ell_i(\overrightarrow{s})$'s are transfers for the firms conditional on their reported signals regardless of the project assignment. Note that as the selected firm only observes its β^i after being selected, the selection x_i cannot depend on β's, but it can on the entire profile of signal announcements. Also, as only the selected firm observes its β^i, the recommended efforts and implemented ex-post utilities cannot depend on the non-selected firms' β's, but they can depend on the entire profile of reported signals.

The expected utility (in period 1) of a firm that is of type s_i but reports type $\hat{s}_i$ while the other firms report truthfully is

$$U_i(\hat{s}_i|s_i) = \mathbb{E}_{s_{-i}}\Big[\ell_i(\hat{s}_i, \mathbf{s_{-i}}) + x_i(\hat{s}_i, \mathbf{s_{-i}})\Big(s_i u^i_H(\hat{s}_i, \mathbf{s_{-i}}) + (1-s_i)u^i_L(\hat{s}_i, \mathbf{s_{-i}})\Big)\Big].$$

By standard arguments (as in Lemma 2.2), one can characterize incentive compatibility in period one as an envelope and a monotonicity condition.

Lemma 5.1. *Incentive compatibility in period 1 is equivalent to the following two conditions:*

1. *$U_i(s)$ is absolutely continuous with*

$$U_i'(s) = -\mathbb{E}_{s_{-i}}\Big[x_i(s_i, \mathbf{s_{-i}})\big(u^i_L(s_i, \mathbf{s_{-i}}) - u^i_H(s_i, \mathbf{s_{-i}})\big)\Big];$$

2. $q_i(s_i) := \mathbb{E}_{s_{-i}}\Big[x_i(s_i, \mathbf{s_{-i}})\big(u^i_L(s_i, \mathbf{s_{-i}}) - u^i_H(s_i, \mathbf{s_{-i}})\big)\Big]$ *is decreasing in s_i.*

Note that when there was only one firm, the first-period incentive compatibility constraints required $u_L(s) - u_H(s)$ to be decreasing, which conflicted with what the principal would prefer if she directly observed the signal. When multiple firms are competing for the project, the principal has an additional tool for screening: the project assignment probability $x_i(\overrightarrow{s})$.

We then redefine our benchmark *second-best allocation*, which must also include the project assignment decision. In this setting, the second-best allocation is defined as the principal's optimal allocation (project assignment and effort recommendations) when she can directly observe signals (s) but not types (β). That is, the allocation that solves the principal's problem when we drop the incentive compatibility constraint of the first period. It is direct that the second-best allocation is given by selecting the firm with the lowest signal and then using the Laffont and Tirole (1986) mechanism. That is,

$$x_i^{SB}(\overrightarrow{s}) = \begin{cases} 1 \text{ if } s_i < s_j \quad \forall j \neq i \\ 0 \text{ otherwise,} \end{cases}$$

and[20]

$$\psi'\big(e_L^{SB}(\overrightarrow{s})\big) = 1 \quad \text{and} \quad \frac{1-\psi'\big(e_H^{SB}(\overrightarrow{s})\big)}{\phi'\big(e_H^{SB}(\overrightarrow{s})\big)} = \frac{1-s_i}{s_i} \quad \text{for all } \overrightarrow{s} \in [\underline{s}, \overline{s}]^n.$$

We now show that if the number of competing firms is sufficiently large, the principal can implement the second-best allocation at approximately no additional cost when there is no ex-post participation constraint. Competition among multiple firms allows the principal to elicit the ex-ante information almost for free. However, when there is an ex-post participation constraint, the cost of implementing the second-best allocation explodes to infinity as we increase the number of firms.

First, we show that when the number of competing firms is sufficiently large, there exist transfers such that the second-best allocation can be implemented even if the principal does not directly observe signals. We say that an allocation $\{x_i(\cdot), e_L^i(\cdot), e_H^i(\cdot)\}$ is *implementable* if there exist transfers and ex-post utilities $\{\ell_i(\cdot), u_L^i(\cdot), u_H^i(\cdot)\}$ such that jointly they satisfy incentive compatibility in both periods.

Lemma 5.2. *There exists $N \in \mathbb{N}$ such that for any number of firms larger than N, the second-best allocation is implementable.*

It remains to construct the transfers that implement the second-best allocation and to compute the principal's payoff. We divide the analysis into two cases: first, when there are no exit rights, that is, there is no ex-post participation constraint; second, when the firm's ex-post outside option is as high as the ex-ante.

5.1 Without Ex-post Participation Constraints

The goal is to implement the second-best allocation at the lowest possible cost to the principal. When there are no exit rights, and the principal can directly observe the first-period signal, she can set up transfers to assure incentive compatibility in the second period while holding all firms at their ex-ante outside option value. That is, she can set $\ell_i(\overrightarrow{s}) = 0$, and the payoffs of the selected firm at

$$\tilde{u}_L(s_i) = s_i\phi(e_H^{SB}(s_i)) \text{ and } \tilde{u}_H(s_i) = -(1-s_i)\phi(e_H^{SB}(s_i)).$$

One can trivially check that incentive compatibility in the second period and participation in the first are satisfied. Hence, the principal's cost when a firm with signal s_i is selected is

$$\pi(s_i) = (1-s_i)\Big[\beta_L + \psi\big(e_L^{SB}(s_i)\big) - e_L^{SB}(s_i)\Big] + s_i\Big[\beta_H + \psi\big(e_H^{SB}(s_i)\big) - e_H^{SB}(s_i)\Big].$$

Moreover, as the principal selects the firm with the lowest signal, her ex-ante expected cost when she directly observes signals is

$$\Pi = \int_{\underline{s}}^{\overline{s}} \pi(s_i)n[1-F(s_i)]^{n-1}f(s_i)ds_i, \tag{9}$$

[20]Ties have measure zero and can be broken arbitrarily.

which converges to $\pi(\underline{s})$ as n goes to infinity.

When the principal does not directly observe the signals, she must also incentivize the firms to reveal their first-period private information. To implement the second-best allocation, the principal must set transfers so that firms have the incentive to reveal their signals truthfully. Naturally, it becomes costlier to the principal to elicit such information given that she must leave information rents to lower signals. However, we show that for a sufficiently large number of firms, the principal's implementation cost is approximately the same as when she directly observes signals. When there is sufficient competition (namely a large number of firms), and there is no ex-post participation constraint, the procuring agency can elicit the ex-ante information essentially for free.

Proposition 5.1. *If there is no ex-post participation constraint, the principal's expected cost of implementing the second-best allocation converges to $\pi(\underline{s})$ as the number of firms increase.*

The proof of Proposition 5.1 consists of constructing transfers $\ell_i(\cdot)$ that implement the second-best allocation. Then, we show that the expected value of such transfers converges to zero when the number of firms increases.

The proposition above implies that as competition increases, the principal can elicit the ex-ante information and implement the second-best allocation at no additional cost when firms are locked in the relationship. However, the result relies on the absence of an ex-post participation constraint. In the next subsection, we show that the cost of implementing the second-best allocation explodes to infinity when the number of firms increases and firms have sufficiently high exit rights.

5.2 With Ex-post Participation Constraints

As before, we look for the cheapest way to implement the second-best allocation. However, now the selected firm has an ex-post outside option which assures a payoff of 0. As in the previous subsection, we start with the benchmark where the principal can observe the firms' signals directly. Hence, the cheapest transfers that implement the second-best allocation when the principal observes signals are given by $\ell_i(\overrightarrow{s}) = 0$, $\hat{u}_L(s_i) = \phi(e_H^{SB}(s_i))$, and $\hat{u}_H(s_i) = 0$.

However, note that such a mechanism does not incentivize firms to truthfully reveal their signals when the principal cannot observe them directly. The payoff of a firm who reports signal $\hat{s}$, but has received signal s while the others report truthfully is given by

$$U_i(\hat{s}|s) = (1-s)\phi(e_H^{SB}(\hat{s}))[1-F(\hat{s})]^{n-1}.$$

As $\phi(e_H^{SB}(\hat{s}))[1-F(\hat{s})]^{n-1}$ is decreasing for sufficiently large n, the firm would have the incentive to under-report its signal.

The principal must modify the mechanism by choosing transfers conditional on reported signals ($\hat{\ell}_i(\overrightarrow{s})$) such that to assure incentive compatibility in the first period and ex-post participation. We show that the principal's cost of implementing the second-best allocation explodes when the number of firms increases.

Proposition 5.2. *Suppose the firms have an ex-post outside option of $\bar{u} = 0$. Then, the principal's expected cost of implementing the second-best allocation diverges to infinity when the number of firms increases.*

The proof consists of constructing the transfers that implement the second-best allocation and showing that expected payments diverge to infinity. Absent the participation constraint, the principal can generate incentives for truthful revelation by making firms that report better signals receive less favorable transfers in case their intrinsic cost turns out to be high. However, when there is an ex-post participation constraint, the principal cannot pay below a certain level. That is, the principal is limited on how much she can punish a firm that has reported a low signal but ends up with a high cost. Hence, firms benefit from under-reporting their signals to increase their probability of being selected. Then, if the chosen firm has a high-cost realization, it takes its ex-post outside option. To avoid such under-reporting and implement the second-best allocation (which requires selecting the firm with the lowest signal), the principal must leave strictly positive rents to all other firms. In particular, she must pay firms who are not selected to not under-report. Proposition 5.2 shows that the payments that avoid under-reporting are bounded away from zero even when competition increases. Hence, as the number of firms increases, the cost of such payments explodes.

6 Discussion

In this section, we briefly discuss two features in our modeling approach: the first one regarding the principal's commitment power and the second about the cardinality of the type space.

6.1 Weakening Principal's Commitment

We have assumed that the principal could fully commit to a mechanism. Regardless if it was not sequentially optimal or renegotiation-proof. However, in many of the results discussed, the optimal contract entails full pooling in the first period and offers the optimal static contract conditional on the principal's belief in the second period. Therefore, the optimal mechanism under full commitment is renegotiation-proof, and intra-period commitment is enough for those results.

This observation contrasts with other dynamic procurement models, which find distinct optimal contracts under full versus partial commitment. For instance, Laffont and Tirole (1990) contrast the equilibria with commitment and renegotiation in a two-period procurement model. However, in their setting, the production occurs in both periods, and the agent's type is persistent over time. Instead, we analyze a framework in which production occurs only at the final period, but the information arrives gradually.

Finally, our results of pooling of the first-period information resemble the literature on ratchet effects, e.g., Freixas et al. (1985), Laffont and Tirole (1987b, 1988), and Gerardi and Maestri (2020). In such papers, the principal adjusts its offer whenever the information is revealed at the beginning of the interaction. Then, the agent does not have the incentive to reveal his type, which generates pooling. If the principal could commit to a long-term contract, pooling would not always arise. However, pooling persists in our setting even when the principal can fully commit to the mechanism. Our source of pooling is non-responsiveness and not the lack of commitment.

6.2 Beyond Binary Types

The finding of full-pooling as optimal is novel for a sequential mechanism design model with ex-post participation constraints and payoff curvature. It stands in sharp contrast to the result in Krähmer and Strausz (2016), namely, that sequential screening is optimal with multiple units for sale (i.e., nonlinear payoffs) under a cross-hazard rate condition on the distribution of types. They argue that multiple units represent an additional screening instrument available to the principal, differently from the indivisible-unit case, where linear payoffs lead to an optimal threshold mechanism.[21]

Krähmer and Strausz (2016) assume binary signals and continuous types together with conditions over the distribution of signals and types such that screening the first-period information is optimal. Their approach relies on finding sufficient conditions under which the monotonicity constraints regarding incentive compatibility do not bind. They impose conditions over the hazard rates and cross-hazard rates between signals and types that prevent non-responsiveness from arising. Meanwhile, our paper examines a setting in which the monotonicity constraint binds throughout the entire support and, consequently, full-pooling of the first-period information arises. Hence, we see our analysis as complementary to theirs. A key distinction between the two settings is that we assume binary types and a continuum of signals while they assume the opposite.

In Appendix B, we show that our results hold beyond the binary-type case. We extend the analysis to a setting with 2-signals and 3-types. Full-pooling remains optimal for a large class of distributions provided a regularity condition over the cross-hazard rates is satisfied. In a related paper but with linear payoffs, Bergemann et al. (2020) argue that the optimality of ex-ante screening depends on how "different" forecasting technologies are from each other, as measured by a transformation of the conditional distributions' cross-hazard rates. A general treatment for arbitrary signal-type distributions remains an open question and is left for future work.

7 Conclusion

This paper develops a model of optimal procurement with dynamic cost information and ex-post exit rights. The motivation for ex-post individual rationality is that procurement relationships often entail limited liability to the contractor. In our model with

[21] Krähmer and Strausz (2015) have shown that, under ex-post participation constraints, the optimal selling mechanism is a threshold mechanism (posted price), which pools first-period information.

two ex-post types, the associated optimal policy exhibits the pooling of ex-ante signals. Hence, the procuring agency prefers to ignore self-reported cost forecasts in contract design, as exit rights increase the burden of ex-ante information rents.

The paper also explores how optimal contracts change when the firm's ex-post outside option value decreases. For high values, the optimal mechanism does not screen the first-period signals, and the ex-ante participation constraint is slack. For moderately lower values, the ex-ante participation constraint binds, but the pooling of the first-period signals remains optimal. For sufficiently low values, the ex-post participation constraint is slack, and it is optimal to screen the first-period information if the distribution of signals satisfies a regularity condition.

Moreover, we show that competition solves the asymmetric information problem only when the ex-post outside option is not high enough. On the one hand, when firms have a weak limited liability protection, the principal can leverage competition to reduce the information rents. On the other hand, if the ex-post outside option is sufficiently high, it becomes increasingly costly to prevent misreporting from all firms. Note, however, that our results with multiple firms restrict attention to how costly it is to implement the second-best allocation. A complete characterization of the optimal mechanism with multiple firms remains an open question.

Appendix A Omitted Proofs

Proof of Lemma 2.1. Manipulating the first equation of (IC-2_s), we get:

$$\begin{aligned} u_L(s) &\geq t_H(s) - \psi(\beta_L - C_H(s)) + u_H(s) - u_H(s) \\ &= u_H(s) + \psi(\beta_H - C_H(s)) - \psi(\beta_L - C_H(s) + \beta_H - \beta_H), \end{aligned}$$

which results in (IC-2_s^L). Similarly,

$$\begin{aligned} u_H(s) &\geq t_L(s) - \psi(\beta_H - C_L(s)) + u_L(s) - u_L(s) \\ &= u_L(s) + \psi(\beta_H - C_L(s)) - \psi(\beta_L - C_L(s) + \beta_H - \beta_H), \end{aligned}$$

which results in (IC-2_s^H).

Moreover, adding up the two equations in (IC-2_s) yields

$$\begin{aligned} &\psi(\beta_L - C_H(s)) + \psi(\beta_H - C_L(s)) - \psi(\beta_H - C_H(s)) - \psi(\beta_L - C_L(s)) \geq 0, \\ \Leftrightarrow \quad &\int_{C_L(s)}^{C_H(s)} \int_{\beta_L}^{\beta_H} \psi''(\beta - C) d\beta dC \geq 0 \end{aligned}$$

which, together with $\psi'' > 0$ and $\beta_H > \beta_L$, yield $C_H(s) \geq C_L(s)$. □

Proof of Lemma 2.2. The proof is a standard argument characterizing incentive compatibility. We present it here for completeness.

Note that $U(\hat{s}|s) = u_L(\hat{s}) - s\big[u_L(\hat{s}) - u_H(\hat{s})\big]$. As effort is bounded by β_H, we have that $|u_L(\hat{s}) - u_H(\hat{s})| \leq \phi(\beta_H)$. Therefore, $U(\hat{s}|\cdot)$ is $\phi(\beta_H)$-Lipschitz continuous and differentiable in s.

If (IC-$1_{s,\hat{s}}$) is satisfied for all $s, \hat{s}$, then, $U(s) = U(s|s) = \underset{\hat{s}}{sup}\, U(\hat{s}|s)$ for all $s \in [\underline{s}, \overline{s}]$. Then, $U(s)$ must be absolutely continuous and by Milgrom and Segal (2002)'s Envelope Theorem

$$U'(s) = \frac{\partial U(s|s)}{\partial s} = -\big[u_L(s) - u_H(s)\big],$$

which proves the necessity of the envelope condition. Now we show that together with monotonicity, they are necessary and sufficient.

Define $g(\hat{s}, s) := U(s) - U(\hat{s}|s)$. Note that g describes the firm's losses from misreporting its signal. Hence, a mechanism satisfies the ex-ante incentive compatibility if and only if g is weakly positive for all $s, \hat{s}$. Note that, as $U(\cdot)$ and $U(\hat{s}|\cdot)$ are absolutely continuous, $g(\hat{s}, \cdot)$ is absolutely continuous and $g(\hat{s}, \hat{s}) = 0$. Hence, the mechanism satisfies ex-ante incentive compatibility if and only if

$$\begin{aligned} g(\hat{s}, s) &= g(\hat{s}, s) - g(\hat{s}, \hat{s}) \\ &= \int_{\hat{s}}^{s} \frac{\partial g(\hat{s}, x)}{\partial s} dx \\ &= \int_{\hat{s}}^{s} \Big[U'(x) - \frac{\partial U(\hat{s}|x)}{\partial s}\Big] dx \\ &= \int_{\hat{s}}^{s} \Big\{\big[u_L(\hat{s}) - u_H(\hat{s})\big] - \big[u_L(x) - u_H(x)\big]\Big\} dx \geq 0 \quad \forall s, \hat{s}, \end{aligned}$$

which holds if and only if the monotonicity constraint is satisfied. $\square$

Proof of Proposition 2.1. We start with a lemma showing that whenever (IC-2^L_s) is slack, then $e_H(s) = e^{FB}$.

Lemma A.1. *Take a mechanism $\{(e_i(\cdot), u_i(\cdot)))\}_{i\in\{L,H\}}$ that satisfies all constraints in problem $\mathcal{P}$. Suppose that there exists a positive mass set $\hat{S}$ such such that (IC-2^L_s) is slack and $e_H(s) \neq e^{FB}$. Then, there exists an alternative mechanism that strictly reduces the principal's expected procuring cost.*

Proof of Lemma A.1. Let an alternative mechanism $\{(\hat{e}_H(\cdot), \hat{u}_H(\cdot), (\hat{e}_L(\cdot), \hat{u}_L(\cdot))\}$ be such that $\hat{u}_H(\cdot) = u_H(\cdot)$, $\hat{u}_L(\cdot) = u_L(\cdot)$, $\hat{e}_L(s) = e_L(s)$ and

$$\hat{e}_H(s) = \begin{cases} min\{e^{FB}, \phi^{-1}(u_L(s) - u_H(s))\} \text{ if } s \in \hat{S} \\ e_H(s) \text{ otherwise.} \end{cases}$$

Note that (IC-$1_{s,\hat{s}}$), (IR-1), (IR-2) and (IC-2^H_s) are directly satisfied by the alternative mechanism since we did not change u_L, u_H, and e_L. For (IC-2^L_s), note that

$$\phi\big(\hat{e}_H(s)\big) \leq min\big\{\phi(e^{FB}), u_L(s) - u_H(s)\big\} \leq \hat{u}_L(s) - \hat{u}_H(s).$$

Therefore, the alternative mechanism is ex-ante and ex-post incentive compatible and individually rational. It remains to show that it reduces the principal's expected procuring cost. Note that $\hat{e}_H(s)$ is closer to e^{FB} than $e_H(s)$ was (strictly so for any $s \in \hat{S}$). Hence, as $\psi(e) - e$ is strictly convex and minimized at e^{FB}, the alternative mechanism is strictly cheaper than the original. $\square$

By Lemma A.1, we can restrict attention to mechanisms such that $e_H(s) = e^{FB}$ whenever (IC-2^L_s) is slack. The next step is to show that (IC-2^L_s) binds almost everywhere. Take a mechanism $\{(e_L(\cdot), u_L(\cdot)), (e_H(\cdot), u_H(\cdot))\}$ that satisfies all constraints in problem $\mathcal{P}$ and has (IC-2^L_s) slack for a positive mass set $\tilde{S}$. As $[u_L(s) - u_H(s)]$ must be decreasing and $e_H(s) = e^{FB}$ whenever (IC-2^L_s) is slack, there exists $\tilde{s} \in (\underline{s}, \overline{s})$ such that for all $s < \tilde{s}$

1. $e_H(s) = e^{FB}$ and
2. $[u_L(s) - u_H(s)] > \phi(e^{FB})$,

while for all $s > \tilde{s}$ we have $u_L(s) - u_H(s) = \phi(e_H(s))$.

Note that $e_H(s) \leq e^{FB}$ for all $s > \tilde{s}$. Otherwise, the principal could decrease $e_H(s)$ for all $s > \tilde{s}$ and still satisfy all constraints while strictly reducing the objective function.

Now we construct an alternative mechanism that is strictly better than $\{(e_L(\cdot), u_L(\cdot)), (e_H(\cdot), u_H(\cdot))\}$. Let $\{(\hat{e}_L(\cdot), \hat{u}_L(\cdot)), (\hat{e}_H(\cdot), \hat{u}_H(\cdot))\}$ be such that $\hat{e}_L(\cdot) = e_L(\cdot)$, $\hat{e}_H(\cdot) = e_H(\cdot)$,

$$\hat{u}_H(s) = \begin{cases} U(\tilde{s}) - (1 - \tilde{s})\phi(e^{FB}) \text{ if } s \leq \tilde{s} \\ u_H(s) \text{ otherwise,} \end{cases}$$

and

$$\hat{u}_L(s) = \begin{cases} U(\tilde{s}) + \tilde{s}\phi(e^{FB}) \text{ if } s \leq \tilde{s} \\ u_L(s) \text{ otherwise.} \end{cases}$$

We need to check that $\{(\hat{e}_L(\cdot), \hat{u}_L(\cdot)), (\hat{e}_H(\cdot), \hat{u}_H(\cdot))\}$ satisfies all constraints.

Regarding (IC-$1_{s,\hat{s}}$), note that $\hat{u}_L(s) - \hat{u}_H(s)$ is decreasing, and that

$$\hat{U}(\hat{s}|s) := \hat{u}_L(\hat{s}) - s[\hat{u}_L(\hat{s}) - \hat{u}_H(\hat{s})].$$

Also, note that $\hat{U}(s) := \hat{U}(s|s)$ is such that

$$\hat{U}(s) = \hat{U}(\tilde{s}) + \int_s^{\tilde{s}} [\hat{u}_L(t) - \hat{u}_H(t)]dt.$$

Therefore, (IC-$1_{s,\hat{s}}$) is satisfied.

For (IR-2) note that for all $s \geq \tilde{s}$ nothing has changed. For $s < \tilde{s}$ note that

$$\hat{u}_H(s) = U(\tilde{s}) - (1-\tilde{s})\phi(e^{FB}) \geq U(\tilde{s}) - (1-\tilde{s})[u_L(\tilde{s}) - u_H(\tilde{s})] = u_H(\tilde{s}) \geq \bar{u}.$$

For (IR-1) note that $\hat{U}(\cdot)$ is decreasing and $\hat{U}(\overline{s}) = U(\overline{s}) \geq 0$. Hence, ex-ante participation is satisfied.

For (IC-2_s^L), note that ex-post utilities for $s \geq \tilde{s}$ have not changed, while for $s < \tilde{s}$

$$\hat{u}_L(s) - \hat{u}_H(s) = \phi(e^{FB}) = \phi(\hat{e}_H(s)).$$

Finally, for (IC-2_s^H), note that ex-post utilities for $s \geq \tilde{s}$ have not changed, while for $s < \tilde{s}$

$$\hat{u}_L(s) - \hat{u}_H(s) = \phi(e^{FB}) \leq u_L(s) - u_H(s) \leq \phi(e_L(s) + \Delta\beta) = \phi(\hat{e}_L(s) + \Delta\beta).$$

The last step is to show that costs have strictly decreased. First, note that all efforts implemented in the new mechanism are the same as in the original. Hence, the total surplus has not changed, and to show the principal is better off, it suffices to show that the expected utility of the firms strictly decreases. Moreover, as nothing has changed for types above $\tilde{s}$ we only need to show that the expected utility of types $s < \tilde{s}$ has decreased. Note that

$$\begin{aligned}
\int_{\underline{s}}^{\tilde{s}} U(s)f(s)ds &= \int_{\underline{s}}^{\tilde{s}} \Big[U(\tilde{s}) - \int_s^{\tilde{s}} U'(t)dt\Big] f(s)ds \\
&= \int_{\underline{s}}^{\tilde{s}} \Big[\hat{U}(\tilde{s}) + \int_s^{\tilde{s}} [u_L(t) - u_H(t)]dt\Big] f(s)ds \\
&> \int_{\underline{s}}^{\tilde{s}} \Big[\hat{U}(\tilde{s}) + \int_s^{\tilde{s}} \phi(e^{FB})dt\Big] f(s)ds \\
&= \int_{\underline{s}}^{\tilde{s}} \hat{U}(s)f(s)ds.
\end{aligned}$$

□

Proof of Lemma 4.1. Showing that the constraints are quasi-concave functions: Let $x > x'$, and thus $\phi(x) > \phi(x')$. Take $\theta \in [0,1]$ and suppose that $\phi(\theta x + (1-\theta)x') < \phi(x')$. The contradiction follows immediately from the assumptions $\psi', \psi'' > 0$ and the definition of $\phi(\cdot)$. The result comes from the fact that any affine transformation of a quasi-concave function is quasi-concave.

Showing non-empty interior of the feasible set: Define $\hat{e}$ such that $\phi(\hat{e}) := -\bar{u}/(1-\overline{s})$. Then,

$$e_H(s) := (1-s)e^{FB} + s\hat{e}$$

belongs to the interior of the feasible set. □

Proof of Proposition 4.1. Substituting the pooling with binding IR-1 mechanism $\{\tilde{e}_L, \tilde{e}_H\}$ defined in (7) in the FOC equations, the equation for $e_L(s)$ becomes immediately satisfied (first-best). In turn, all terms $\phi'(\tilde{e}_H)$ cancel out in the equations for $e_H(s)$, so they become:

$$\text{for } s \in [\underline{s}, \tilde{s}]: \quad f(s)s\left[\frac{1-s}{s} - \frac{1-\tilde{s}}{\tilde{s}}\right] - \gamma(s)(1-s) = \int_{\tilde{s}}^{\overline{s}} \lambda(s,b)(1-s)db \tag{10a}$$

$$\text{for } s \in [\tilde{s}, \overline{s}]: \quad f(s)s\left[\frac{1-\tilde{s}}{\tilde{s}} - \frac{1-s}{s}\right] + \gamma(s)(1-s) = \int_{\underline{s}}^{\tilde{s}} \lambda(a,s)(1-a)da. \tag{10b}$$

We now define functions $\tilde{\gamma}$ and $\tilde{\lambda}$ such that the first-order conditions are satisfied. First, define $\tilde{\gamma}(s) = 0$ for $s \in [\underline{s}, \tilde{s}]$, and

$$\tilde{\gamma}(s) = \frac{2s^*}{(\overline{s}-\tilde{s})^2}\left[\frac{1-s^*}{s^*} - \frac{1-\tilde{s}}{\tilde{s}}\right]\frac{s-\tilde{s}}{1-s} \quad \text{if } s \in (\tilde{s}, \overline{s}].$$

As $s^* \leq \tilde{s}$, we have that $\tilde{\gamma}(s) \geq 0$ for all $s \in [\underline{s}, \overline{s}]$. Second, define $\tilde{\alpha} : [\underline{s}, \tilde{s}] \to \mathbb{R}_+$ and $\tilde{\beta} : [\tilde{s}, \overline{s}] \to \mathbb{R}_+$ as

$$\tilde{\alpha}(a) = f(a)a\left[\frac{1-a}{a} - \frac{1-\tilde{s}}{\tilde{s}}\right] - \tilde{\gamma}(a)(1-a)$$

and

$$\tilde{\beta}(b) = f(b)b\left[\frac{1-\tilde{s}}{\tilde{s}} - \frac{1-b}{b}\right] + \tilde{\gamma}(b)(1-b).$$

Note that $\tilde{\alpha}(a) \geq 0$ and $\tilde{\beta}(b) \geq 0$ for all $(a,b) \in [\underline{s}, \tilde{s}] \times [\tilde{s}, \overline{s}]$. Moreover,

$$\int_{\underline{s}}^{\tilde{s}} \tilde{\alpha}(a)da - \int_{\tilde{s}}^{\overline{s}} \tilde{\beta}(b)db = s^*\left[\frac{1-s^*}{s^*} - \frac{1-\tilde{s}}{\tilde{s}}\right] - \int_{\underline{s}}^{\overline{s}} \tilde{\gamma}(s)(1-s)ds = 0.$$

Then, define the Lagrange multipliers as

$$\tilde{\lambda}(a,b) = \frac{\tilde{\beta}(b)\tilde{\alpha}(a)}{(1-a)\int_{\tilde{s}}^{\overline{s}} \tilde{\beta}(z)dz} \geq 0.$$

We, then, have the first-order conditions (10) satisfied. □

Proof of Proposition 4.2. Reorganize $\tilde{\mathcal{L}}$ to obtain:

$$\begin{aligned}\tilde{\mathcal{L}} = \quad & \mathcal{L} + \int_{\underline{s}}^{\tilde{s}} u_H(s)\left[f(s) - \int_{\tilde{s}}^{\overline{s}} \lambda(s,b)db - \gamma^*(s) - \mu(s)\right] ds \\ & + \int_{\tilde{s}}^{\overline{s}} u_H(s)\left[f(s) + \int_{\underline{s}}^{\tilde{s}} \lambda(a,s)da - \tilde{\gamma}(s) - \mu(s)\right] ds + \bar{u}\Big[1 + \int_{\underline{s}}^{\overline{s}} \mu(s)ds\Big].\end{aligned}$$

Note that, for given multipliers, the minimizer of $\tilde{\mathcal{L}}$ coincides with the minimizer of $\mathcal{L}$ if we define:

$$\tilde{\mu}(s) = \begin{cases} f(s) - \int_{\tilde{s}}^{\overline{s}} \tilde{\lambda}(s,b)db - \tilde{\gamma}(s) & \text{if } s \in [\underline{s}, \tilde{s}] \\ f(s) + \int_{\underline{s}}^{\tilde{s}} \tilde{\lambda}(a,s)da - \tilde{\gamma}(s) & \text{if } s \in [\tilde{s}, \overline{s}]. \end{cases}$$

It remains to show when the proposed multiplier $\tilde{\mu}(\cdot)$ is positive for all $s \in [\underline{s}, \overline{s}]$. Note that, for $s < \tilde{s}$:

$$\tilde{\mu}(s) = f(s) - \int_{\tilde{s}}^{\overline{s}} \tilde{\lambda}(s,b)db = f(s)\frac{s(1-\tilde{s})}{\tilde{s}(1-s)} \geq 0.$$

For $s > \tilde{s}$:

$$\tilde{\mu}(s) = f(s) + \int_{\underline{s}}^{\tilde{s}} \tilde{\lambda}(a,s)da - \tilde{\gamma}(s),$$

which might be negative if $\tilde{\gamma}(s)$ is large enough. Note, however, that if $\bar{u} = \bar{u}_1$, then $\tilde{s} = s^*$, and $\tilde{\gamma}(s) = 0$ for all $s \in [\underline{s}, \overline{s}]$. Hence, $\tilde{\mu}(s) \geq inf\{f(s) : s \in [\underline{s}, \overline{s}]\} > 0$. Moreover, for each s, $\tilde{\mu}(s)$ varies continuously with $\bar{u}$, which implies that there exists $\bar{u}_2 < \bar{u}_1$ such that for all $\bar{u} \in [\bar{u}_2, \bar{u}_1]$ we have $\tilde{\mu}(s) \geq 0$ for all $s \in [\underline{s}, \overline{s}]$. □

Proof of Proposition 4.3.

Integrating by parts, we can re-write $\mathcal{P}_1$ as

$$\min_{(e_L(s), e_H(s))} \int_{\underline{s}}^{\overline{s}} \Big\{(1-s)\Big[\psi(e_L(s)) - e_L(s)\Big] + s\Big[\psi(e_H(s)) - e_H(s)\Big] + \phi(e_H(s))\frac{F(s)}{f(s)}\Big\} f(s)ds$$

subject to

$$e_H(s) \text{ decreasing.}$$

Then, we relax the monotonicity constraint and minimize pointwise.

$$\psi'\big(\hat{e}_L(s)\big) = 1 \quad \text{and} \quad \frac{1 - \psi'\big(\hat{e}_H(s)\big)}{\phi'\big(\hat{e}_H(s)\big)} = \frac{F(s)}{sf(s)} \quad \text{for all } s \in [\underline{s}, \overline{s}].$$

As $F(s)/sf(s)$ is increasing, we have that $\hat{e}_H(s)$ is decreasing, and we found the solution. □

Proof of Lemma 5.2. An allocation is implementable if and only if its associated $q_i(s_i)$ is decreasing in s_i, and $u_L^i(s_i, s_{-i}) - u_H^i(s_i, s_{-i}) \geq \phi(e_H(s_i))$. Hence, an allocation is implementable if and only if

$$\hat{q}_i(s_i) := \mathbb{E}_{s_{-i}}\Big[x_i(s_i, s_{-i})\phi(e_H(s_i))\Big] \text{ is decreasing in } s_i.$$

By the definition of the second-best allocation, we get

$$\hat{q}_i(s_i) = \phi(e_H^{SB}(s_i))[1 - F(s_i)]^{n-1}.$$

Taking the derivative with respect to s_i we get

$$\hat{q}_i'(s_i) = (1 - F(s_i))^{n-2}\Bigg[-(n-1)f(s_i)\phi(e_H^{SB}(s_i)) + (1 - F(s_i))\phi'(e_H^{SB}(s_i))\frac{de_H^{SB}(s_i)}{ds_i}\Bigg],$$

which is negative if and only if the square brackets term is negative. Note that if the derivative of $\bar{e}_H(s_i)$ is bounded, then the term inside the square brackets is negative for a sufficiently high n. By the definition of $e_H^{SB}(s_i)$

$$\frac{de_H^{SB}(s_i)}{ds_i} = \frac{\phi'(e_H^{SB}(s_i))}{s_i^2\psi''(e_H^{SB}(s_i)) + s_i(1 - s_i)\phi''(e_H^{SB}(s_i))} \leq \frac{\phi'(e_H^{SB}(\overline{s}))}{\underline{s}^2\psi''(e_H^{SB}(\underline{s})) + \underline{s}(1 - \overline{s})\phi''(e_H^{SB}(\underline{s}))}.$$

Hence, for n sufficiently large, $\hat{q}_i'(s_i) \leq 0$ for all $s_i \in [\underline{s}, \overline{s}]$. □

Proof of Proposition 5.1. For each number of firms n, we define a mechanism that implements the second-best allocation so that the expected principal's cost converges in probability to $\pi(\underline{s})$. For each $n \in \mathbb{N}$, let the project assignment and recommended efforts be the second-best allocation. Moreover, let the u_H and u_L be given by $\tilde{u}_H$ and $\tilde{u}_L$. It remains to construct transfers ℓ that assure first-period incentive compatibility and ex-ante participation.

By Lemma 5.2 we know that for large enough n, the second-best allocation satisfies the monotonicity requirement for incentive compatibility. For the envelope condition to hold, it must be that

$$\begin{aligned}\tilde{U}_i(s_i) &= \mathbb{E}_{s_{-i}}\Big[\tilde{\ell}_i(\hat{s}_i, s_{-i}) + x_i^{SB}(\hat{s}_i, s_{-i})\Big(s_i\tilde{u}_H^i(\hat{s}_i, s_{-i}) + (1 - s_i)\tilde{u}_L^i(\hat{s}_i, s_{-i})\Big)\Big] \\ &= \tilde{U}_i(\overline{s}) + \int_{s_i}^{\overline{s}} \phi(e_H^{SB}(s))[1 - F(s)]^{n-1}ds.\end{aligned}$$

By letting $\tilde{U}(\overline{s}) = 0$, we can write

$$\tilde{\ell}_i(\overrightarrow{s}) = \int_{s_i}^{\overline{s}} \phi(e_H^{SB}(s))[1 - F(s)]^{n-1}ds. \tag{11}$$

By construction, the new mechanism satisfies incentive compatibility in both periods and assures ex-ante participation since $\tilde{U}_i(\overline{s}) = 0$ and $\tilde{U}_i'(s_i) \leq 0$. It remains to show that the principal's expected cost (denoted by $\tilde{\Pi}$) converges in probability to $\pi(\underline{s})$ as n goes to infinity.

Note that this mechanism is the same as the case in which the principal observes the first-period signal but with the addition of the transfers $\tilde{\ell}_i$'s. Hence,

$$\begin{aligned}\tilde{\Pi} - \Pi =& \sum_{i=1}^{n} \mathbb{E}_{s_i}[\tilde{\ell}_i(\mathbf{s_i})] \\ =& n \int_{\underline{s}}^{\overline{s}} \int_{s_i}^{\overline{s}} \phi(e_H^{SB}(s))[1-F(s)]^{n-1} ds f(s_i) ds_i \\ =& \int_{\underline{s}}^{\overline{s}} \frac{F(s)}{f(s)} \phi(e_H^{SB}(s)) n[1-F(s)]^{n-1} f(s) ds.\end{aligned} \tag{12}$$

Note that $n[1-F(s)]^{n-1}f(s)$ is the density of the minimum of n independent random variables distributed according to F, which converges in probability to $\underline{s}$. Therefore, $\tilde{\Pi} - \Pi$ converges to $\phi(e_H^{SB}(\underline{s}))F(\underline{s})/f(\underline{s}) = 0$. As Π converges to $\pi(\underline{s})$, so it does $\tilde{\Pi}$. □

Proof of Proposition 5.2. The mechanism that implements the second-best allocation in the cheapest possible way minimizes expected transfers subject to ex-ante and ex-post participation, and incentive compatibility for truthful reporting in both periods.

When $\bar{u} = 0$ ex-post participation directly implies on ex-ante participation. Moreover, ex-post participation requires the payoffs of all firms to be ex-post weakly positive. Note that such requirement imposes that transfers to non-selected firms must be weakly positive ($\ell^i(\overrightarrow{s}) \geq 0$), and that the ex-post utility of a selected firm who has a high type must also be positive. That is, ex-post participation is given by

$$\ell_i(\overrightarrow{s}) \geq 0 \text{ and } [\ell_i(\overrightarrow{s}) + u_H(\overrightarrow{s})] \geq 0 \quad \text{for all } \overrightarrow{s}, i. \tag{$IR-2$}$$

After being selected, the firm's incentives to report its type remains as before. That is, ex-post incentive compatibility is given by

$$u_L^i(\overrightarrow{s}) - u_H^i(\overrightarrow{s}) \geq \phi(e_H^{SB}(\overrightarrow{s})) \quad \text{for all } \overrightarrow{s}, i. \tag{$IC-2$}$$

The ex-ante incentive compatibility is characterized by Lemma 5.1 as an envelope and a monotonicity condition. Finally, given the second-best allocation rule, the principal's expected transfers are given by

$$\begin{aligned}\sum_{i=1}^{n} \mathbb{E}_{\overrightarrow{s}} \Bigg[\ell_i(\overrightarrow{\mathbf{s}}) + x_i^{SB}(\overrightarrow{\mathbf{s}}) \Big[(1-\mathbf{s_i})[\beta_L - e_L^{SB}(\overrightarrow{\mathbf{s}}) + \psi(e_L^{SB}(\overrightarrow{\mathbf{s}})) \\ + u_L(\overrightarrow{\mathbf{s}})] + \mathbf{s_i}[\beta_H - e_H^{SB}(\overrightarrow{\mathbf{s}}) + \psi(e_H^{SB}(\overrightarrow{\mathbf{s}})) + u_H(\overrightarrow{\mathbf{s}})] \Big] \Bigg]\end{aligned} \tag{13}$$

where the sum of ℓ_i's denote the payments unconditional on project assignment and the term multiplied by $x_i^{SB}(\overrightarrow{\mathbf{s}})$ denotes the payments to the selected firm[22].

Given that we fixed the allocation as the second-best, the only choices are the transfers unconditional to project assignment (ℓ_i's) and ex-post utilities (u_L^i and u_H^i). Then,

[22]The transfers to the selected firm are written as a function of recommended efforts and ex-post utilities as done in previous sections.

finding the cheapest way possible to implement the second-best requires minimizing (13) subject to $(IR-2)$, $(IC-2)$, (*envelope*) and (*monotonicity*). That is, the cheapest implementation of the second-best allocation can be found by solving

$$\min_{\{u_L^i(\cdot),u_H^i(\cdot),\ell_i(\cdot)\}} \sum_{i=1}^{n} \mathbb{E}_{\overrightarrow{s}}\Bigg[\ell_i(\overrightarrow{\mathbf{s}}) + x_i^{SB}(\overrightarrow{\mathbf{s}})\Big[(1-\mathbf{s_i})[\beta_L - e_L^{SB}(\overrightarrow{\mathbf{s}}) + \psi(e_L^{SB}(\overrightarrow{\mathbf{s}}))$$
$$+ u_L(\overrightarrow{\mathbf{s}})] + \mathbf{s_i}[\beta_H - e_H^{SB}(\overrightarrow{\mathbf{s}}) + \psi(e_H^{SB}(\overrightarrow{\mathbf{s}})) + u_H(\overrightarrow{\mathbf{s}})]\Big]\Bigg]$$

subject to

$$u_L^i(\overrightarrow{s}) - u_H^i(\overrightarrow{s}) \geq \phi(e_H^{SB}(\overrightarrow{s})) \quad \text{for all } \overrightarrow{s}, i \tag{$IC-2$}$$

$$\ell_i(\overrightarrow{s}) \geq 0 \text{ and } [\ell_i(\overrightarrow{s}) + u_H(\overrightarrow{s})] \geq 0 \quad \text{for all } \overrightarrow{s}, i \tag{$IR-2$}$$

$$q_i(s_i) = \mathbb{E}_{s_{-i}}[x_i^{SB}(s_i, \mathbf{s_{-i}})(u_L^i(s_i, \mathbf{s_{-i}}) - u_H^i(s_i, \mathbf{s_{-i}}))] \quad \text{decreasing,} \tag{$monotonicity$}$$

$$U_i(s_i) = U(\overline{s}_i) + \int_{s_i}^{\overline{s}} \mathbb{E}_{s_{-i}}[x_i^{SB}(z, \mathbf{s_{-i}})(u_L^i(z, \mathbf{s_{-i}}) - u_H^i(z, \mathbf{s_{-i}}))]dz \quad \text{for all } \overrightarrow{s}, i, \tag{$envelope$}$$

where

$$U_i(s_i) := \mathbb{E}_{s_{-i}}\left[\ell_i(s_i, \mathbf{s_{-i}}) + x_i^{SB}(s_i, \mathbf{s_{-i}})(s_i u_H^i(s_i, \mathbf{s_{-i}}) + (1-s_i)u_L^i(s_i, \mathbf{s_{-i}})\right] \quad \text{for all } i, s_i.$$

We proceed with two auxiliary lemmas that simplify the problem.

Lemma A.2. *There is no loss in setting $u_H^i(\overrightarrow{s}) = 0$ for all $i, \overrightarrow{s}$.*

Proof. Take any vector of triples $(u_L^i(\cdot), u_H^i(\cdot), \ell_i(\cdot))$ that satisfy all four constraints. Define an alternative mechanism in which

$$\tilde{u}_H^i(\overrightarrow{s}) := 0, \quad \tilde{u}_L^i(\overrightarrow{s}) := u_L^i(\overrightarrow{s}) - u_H^i(\overrightarrow{s}) \text{ and } \tilde{\ell}_i(\overrightarrow{s}) = \ell_i(\overrightarrow{s}) + x_i^{SB}(\overrightarrow{s})(1-s_i)u_H^i(\overrightarrow{s}).$$

By construction, the objective function has stayed the same. Also, as $\tilde{u}_L^i(\overrightarrow{s}) - \tilde{u}_H^i(\overrightarrow{s}) = u_L^i(\overrightarrow{s}) - u_H^i(\overrightarrow{s})$ the first two constraints are trivially satisfied. Finally, note that by definition $\tilde{\ell}_i(\overrightarrow{s}) \geq 0$. Hence, all constraints were satisfied, and the objective function was unchanged. □

Lemma A.3. *It is optimal to set $u_L^i(\overrightarrow{s}) = \phi(e_H^{SB}(\overrightarrow{s}))$ for all $i, \overrightarrow{s}$.*

Proof. If $(IC-2)$ is slack for any pair $(\overrightarrow{s}, i)$, then the principal can reduce $u_L^i(\overrightarrow{s})$ and save on information rents. □

It only remains to solve for the optimal report-conditional transfers (ℓ's) that assure incentive compatibility in the first-period and ex-post participation.

First-period incentive compatibility is assured by monotonicity (which is satisfied under the second-best allocation and $u_L^i(\overrightarrow{s}) - u_h^i(\overrightarrow{s}) = \phi(e_H^{SB}(\overrightarrow{s})$, and an envelope condition, as described in Lemma 5.1. Ex-post participation holds if and only if $\hat{\ell}_i(\overrightarrow{s}) \geq 0$

for all vectors of reports. As the best allocation is given to the lowest type $\underline{s}$, the best is to set it to the lowest value consistent with ex-post participation, which is 0. Then, using the envelope characterization of incentive compatibility, we can write the transfers of the cheapest mechanism to implement the second-best allocation as

$$\begin{aligned}\mathbb{E}_{s_{-i}}\left[\hat{\ell}_i(s_i, s_{-i})\right] =& U_i(s_i) - \mathbb{E}_{s_{-i}}\left[x_i^{SB}(\overrightarrow{s})[(1-s_i)u_L^i(\overrightarrow{s}) + s_i u_H^i(\overrightarrow{s})]\right] \\ =& U_i(\underline{s}) + \int_{\underline{s}}^{s_i} U_i'(z)dz - [1-F(s_i)]^{n-1}\phi(e_H^{SB}(s_i))(1-s_i) \\ =& \Big\{ \Big[\phi(e_H^{SB}(\underline{s}))(1-\underline{s}) - [1-F(s_i)]^{n-1}\phi(e_H^{SB}(s_i))(1-s_i)\Big] \\ & - \int_{\underline{s}}^{s_i} \phi(e_H^{SB}(z))[1-F(z)]^{n-1}dz\Big\}.\end{aligned}$$

Where the first equality comes from the definition of U_i, the second from replacing x_i^{SB}, u_L^i, u_H^i, and $(envelope)$. The third comes from integration by parts and setting $\hat{\ell}_i(\underline{s}, s_{-i}) = 0$.

Note that only the expectation of $\hat{\ell}_i(s_i, s_{-i})]$ with respect to s_{-i} matters for incentive compatibility. Hence, there is no loss in setting ℓ_i to depend only on s_i. With a slight abuse of notation, we write the report-conditional transfers simply as $\hat{\ell}(s_i)$. By construction, such a mechanism satisfies incentive compatibility in the first and the second periods; it remains to verify that $\hat{\ell}(s_i) \geq 0$ for all $s_i \in [\underline{s}, \overline{s}]$. Note that $\hat{\ell}(\underline{s}) = 0$. Moreover, for sufficiently high n

$$\hat{\ell}'(s_i) = -(1-s_i)\frac{d[1-F(s_i)]^{n-1}\phi(e_H^{SB}(s_i))}{ds_i} \geq 0.$$

Hence, for sufficiently high n, the mechanism described here satisfies the ex-post participation constraint. We now show that its cost diverges to infinity as we increase the number of firms.

Note that the total expected report-conditional transfers paid by the principal is

$$n\int_{\underline{s}}^{\overline{s}} \hat{\ell}(s_i)f(s_i)ds_i =$$

$$n\phi(e_H^{SB}(\underline{s}))(1-\underline{s}) - \int_{\underline{s}}^{\overline{s}} \phi(e_H^{SB}(z)\frac{[1-F(z)]}{f(z)}nf(z)[1-F(z)]^{n-1}dz.$$

Note that $\lim_{n\to\infty}\{n\phi(e_H^{SB}(\underline{s}))(1-\underline{s})\} = +\infty$, while

$$\lim_{n\to\infty}\int_{\underline{s}}^{\overline{s}} \phi(e_H^{SB}(z))\frac{[1-F(z)]}{f(z)}nf(z)[1-F(z)]^{n-1}dz = \phi(e_H^{SB}(\underline{s}))\frac{[1-F(\underline{s})]}{f(\underline{s})} < +\infty.$$

Therefore, the cost of implementing the second-best allocation diverges to infinity as we increase the number of firms. □

The constructive approach

This subsection presents an alternative proof for Theorem 3.1. It takes a constructive approach similar to the one taken for $\bar{u} < 0$ in the main text. This alternative proof generalizes the result beyond a continuum of first-period signals, allowing us to accommodate any distribution with or without atoms. We kept the non-responsiveness approach in the body of the text as it is the approach that conveys intuition more clearly, despite losing generality.

Let $q : S \to (0, 1)$ be a right-continuous and weakly increasing function that denotes the probability of being a high-cost type conditional on having received signal s. That is, $Pr(\beta = \beta_H|s) = q_s$. Note that in our original formulation $q_s = s$. By introducing the function $q : S \to (0, 1)$, which can be constant on sub-intervals of S, we allow for atoms in the ex-ante probability of being a high-cost firm[23].

The constructive approach addresses the design problem with a guess-and-verify algorithm: we first guess that the *pooling mechanism* is an optimal solution to problem $\mathcal{P}$. Then, we construct Lagrange multipliers that sustain such a mechanism as a solution to the procuring agency's problem.

Notice the pooling mechanism described in (2) satisfies all (IC-$1_{s,\hat{s}}$) constraints with equality. Also, it immediately follows that it satisfies (IC-2_s^L) and (IC-2_s^H). Hence, it suffices to focus on ex-post participation constraints (IR$_s$) and optimality conditions. Our solution algorithm proceeds according to the following steps:

1. State a relaxed version of the principal's problem, denoted by $\mathcal{R}$, which exogenously sets ex-post profits $u_H(s)$ to zero;
2. Guess-and-verify the set of active (IC-$1_{s,\hat{s}}$) constraints at the pooling solution. Define a further relaxed problem that ignores the remaining inactive constraints, denoted by $\mathcal{R}^0$;
3. Construct problem $\mathcal{R}^0$'s Lagrangian and present a Karush-Kuhn-Tucker argument to the solution taking the form of the pooling mechanism (2); and
4. Extend the Lagrangian to include the ex-post IR constraints and find multipliers that sustain the pooling mechanism (2) and $u_H^*(s) = 0$ as an optimal solution.

Step 1. First, we fix $u_H(s) = 0$. Then, the modified principal's design problem becomes:

$$\begin{aligned}\mathcal{R}: \quad \min_{\{e_L(\cdot), e_H(\cdot)\}} \int_{\underline{s}}^{\overline{s}} & \{(1-q_s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\ & + q_s \left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s) ds \\ \text{s.t.} \quad & (1-q_s)\left[\phi(e_H(s)) - \phi(e_H(\hat{s}))\right] \geq 0, \quad \forall \hat{s}, s.\end{aligned}$$

Step 2. We then relax problem $\mathcal{R}$ by restricting the set of first-period incentive constraints to pairs of signals in which a firm with a below-average signal ("good type") does not envy the allocation assigned to an above-average signal ("bad type"). If the pooling mechanism (2) solves this relaxed problem, then it also solves the problem $\mathcal{R}$, as it trivially satisfies all neglected (IC-$1_{s,\hat{s}}$) constraints.[24]

[23]Note that any atom in the ex-ante probability of being a high-cost firm can be represented by q_s being constant in a given sub-interval.

[24] This approach was first proposed by Krähmer and Strausz (2015).

Definition. Define IC* as the set of incentive constraints in which a firm with a below-average probability of having a high cost ("good type") does not envy the allocation assigned to an above-average probability of having a high cost ("bad type"). That is, let $q^* = \int_{\underline{s}}^{\overline{s}} q_s f(s)ds$, and $s^* := inf\{s \in S : q_s \geq q^*\}$, then define

$$\text{IC}^* = \{\text{IC-1}_{s,\hat{s}} : (1-s) > (1-s^*) > (1-\hat{s})\}$$

Remark 2. If the pooling mechanism $\Big\{\big(\bar{e}_H, \bar{u}\big), \big(\bar{e}_L, \phi(\bar{e}_H)\big)\Big\}$, as defined in (2), solves problem $\mathcal{R}$ restricted to IC*, then it solves the global problem $\mathcal{R}$.

Hence, it suffices to show that the pooling mechanism is a solution to the following version of the relaxed problem:

$$\begin{aligned}\mathcal{R}^0 : \quad \min_{\{e_L(\cdot), e_H(\cdot)\}} \quad & \int_{\underline{s}}^{\overline{s}} \{(1-q_s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\ & \qquad + q_s \left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s)ds \\ & \text{s.t. } (1-q_a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] \geq 0, \quad \forall a < s^* < b.\end{aligned}$$

Step 3. Note that the Karush-Kuhn-Tucker theorem for function spaces (Luenberger, 1997, p. 220) applies to concave optimization problems (i.e., convex objectives and feasibility sets for minimization problems). However, the constraints (IC-$1_{s,\hat{s}}$) in problem $\mathcal{R}^0$ are differences of convex functions $\phi(\cdot)$ and are not necessarily convex. Lemma 4.1 establishes that the optimization problem $\mathcal{R}^0$ satisfies the Karush-Kuhn-Tucker theorem conditions.

By the quasi-concave version of the Karush-Kuhn-Tucker theorem (Arrow and Enthoven, 1961, Theorem 1), a pair $\{e_L(s), e_H(s)\}$ is a solution to problem $\mathcal{R}$ if there exist positive multipliers $\lambda : [\underline{s}, s^*] \times [s^*, \overline{s}]^2 \to \mathbb{R}_+$ so that $\{e_L(s), e_H(s)\}$ minimizes the Lagrangian:

$$\begin{aligned}\mathcal{L} = & \int_{\underline{s}}^{\overline{s}} \{(1-q_s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] + q_s\left[\beta_H - e_H(s) + \psi(e_H(s))\right]\} f(s)ds \\ & - \int_{\underline{s}}^{s^*} \int_{s^*}^{\overline{s}} \lambda(a,b)(1-q_a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] dbda.\end{aligned}$$

The minimizer of $\mathcal{L}$ must satisfy the following pointwise first-order conditions (FOC):

$$[e_L(s)] : \quad f(s)(1-q_s)\left(-1 + \psi'(e_L(s))\right) = 0$$

$$[e_H(s)] : \begin{cases} f(s)q_s\left(-1+\psi'(e_H(s))\right) + f(s)(1-q_s)\phi'(e_H(s)) - \displaystyle\int_{s^*}^{\overline{s}} \lambda(s,b)(1-q_s)\phi'(e_H(s))db = 0 \\ \hfill \text{if } s \in [\underline{s}, s^*] \\ f(s)q_s\left(-1+\psi'(e_H(s))\right) + f(s)(1-q_s)\phi'(e_H(s)) + \displaystyle\int_{\underline{s}}^{s^*} \lambda(a,s)(1-q_a)\phi'(e_H(s))da = 0 \\ \hfill \text{if } s \in [s^*, \overline{s}]\,. \end{cases}$$

The following result establishes the existence of multipliers $\lambda(a,b) \geq 0$ satisfying the preceding FOC system. Therefore, by Remark 2, the pooling mechanism also solves problem $\mathcal{R}$.

Proposition A.1. *There exists a function* $\lambda^* : [\underline{s}, s^*] \times [s^*, \overline{s}] \to \mathbb{R}_+$ *such that the pooling allocation* $\{\bar{e}_L, \bar{e}_H\}$ *is an optimal solution to problem* $\mathcal{R}^0$*, with multipliers given by* $\lambda^*(s,\hat{s})$*.*

Proof of Proposition A.1. Substituting the pooling mechanism $\{\bar{e}_L, \bar{e}_H\}$ defined in (2) in the FOC equations, the equation for $e_L(s)$ becomes immediately satisfied (first-best). In turn, all terms $\phi'(\bar{e}_H)$ cancel out in the equations for $e_H(s)$, so they become:

for $s \in [\underline{s}, s^*]$:

$$f(s)q_s\left[\frac{1-q_s}{q_s} - \frac{1-q^*}{q^*}\right] = \int_{s^*}^{\overline{s}} \lambda(s,b)(1-s)db \tag{14a}$$

for $s \in [s^*, \overline{s}]$:

$$f(s)q_s\left[\frac{1-q^*}{q^*} - \frac{1-q_s}{q_s}\right] = \int_{\underline{s}}^{s^*} \lambda(a,s)(1-q_a)da. \tag{14b}$$

Let the left-hand side of equations (14a) and (14b) be defined as $\underline{r}(s)$ and $\overline{r}(s)$, respectively. Reorganizing the expressions, one obtains

$$\underline{r}(s) = \frac{f(s)}{q^*}(q^* - q_s) > 0, \quad \forall s < s^*$$

$$\overline{r}(s) = \frac{f(s)}{q^*}(q_s - q^*) > 0, \quad \forall s > s^*$$

Notice the terms $\int_{s^*}^{\overline{s}} \overline{r}(b)db$ and $\int_{\underline{s}}^{s^*} \underline{r}(a)da$ are equal, and denote them by r^*. Defining the Lagrange multipliers as

$$\lambda(a,b) = \frac{\overline{r}(b)\underline{r}(a)}{(1-a)r^*} \geq 0$$

we have first-order conditions (14) satisfied. □

Step 4. The last step is to consider the Lagrangian that explicitly incorporates the ex-post participation constraints (IR-2):

$$\begin{aligned}\mathcal{L}^* = \int_{\underline{s}}^{\overline{s}} \Big\{&(1-q_s)\left[\beta_L - e_L(s) + \psi(e_L(s)) + \phi(e_H(s))\right] \\ &+ q_s\left[\beta_H - e_H(s) + \psi(e_H(s)) + u_H(s)\right]\Big\} f(s)ds \\ - \int_{\underline{s}}^{s^*}\int_{s^*}^{\overline{s}} &\lambda(a,b)\left\{(1-q_a)\left[\phi(e_H(a)) - \phi(e_H(b))\right] + u_H(a) - u_H(b)\right\} dbda \\ &- \int_{\underline{s}}^{\overline{s}} \mu(s)u_H(s)ds.\end{aligned}$$

By the same argument in Arrow and Enthoven (1961), if there exist positive multipliers $\lambda(a,b) \geq 0$ and $\mu(s) \geq 0$ such that Lagrangian $\mathcal{L}^*$ is minimized at $\{\bar{e}_L, \bar{e}_H\}$ and $u_H^*(s) = 0, \forall s$, then the pooling mechanism is a solution to problem $\mathcal{P}$. Theorem A.1 establishes the existence of such multipliers.

Theorem A.1. *There exist functions $\lambda^* : [\underline{s}, s^*] \times [s^*, \overline{s}] \to \mathbb{R}_+$ and $\mu^* : [\underline{s}, \overline{s}] \to \mathbb{R}_+$ such that the pooling allocation $\{\bar{e}_L, \bar{e}_H\}$ and the ex-post profit level $u_H^* = 0$ are a solution to problem $\mathcal{P}$, with multipliers given by $\lambda^*(s, \hat{s})$ and $\mu^*(s)$.*

Therefore, the pooling mechanism is optimal. In other words, the principal does not charge different franchise fees to different signals, offering the same ex-post menu of incentive contracts irrespective of the firm's cost forecast. This result stands in contrast to the optimal policy without an ex-post participation constraint, where the regulator achieves an improved rent-efficiency trade-off from screening ex-ante signals.

Proof of Theorem A.1. Reorganize $\mathcal{L}^*$ to obtain:

$$\begin{aligned}\mathcal{L}^* = \quad & \mathcal{L} + \int_{\underline{s}}^{s^*} u_H(s) \left[f(s) - \int_{s^*}^{\overline{s}} \lambda(s,b) db - \mu(s) \right] ds. \\ & + \int_{s^*}^{\overline{s}} u_H(s) \left[f(s) + \int_{\underline{s}}^{s^*} \lambda(a,s) da - \mu(s) \right] ds.\end{aligned}$$

Note that $\mathcal{L}^*$ coincides with $\mathcal{L}$ if we define:

$$\mu^*(s) = \begin{cases} f(s) - \int_{s^*}^{\overline{s}} \lambda^*(s,b) db & \text{if } s \in [\underline{s}, s^*] \\ f(s) + \int_{\underline{s}}^{s^*} \lambda^*(a,s) da & \text{if } s \in [s^*, \overline{s}] \end{cases}$$

To see that the proposed multiplier $\mu^*(\cdot)$ is always positive, note that, for $s < s^*$:

$$\mu^*(s) = f(s) - \int_{s^*}^{\overline{s}} \frac{\overline{r}(b)\underline{r}(s)}{(1-q_s)r^*} db = \frac{f(s)}{q^*} \left[q^* - \frac{q^* - q_s}{1 - q_s} \right] = f(s) \frac{q_s(1-q^*)}{q^*(1-q_s)} \geq 0.$$

Moreover, for $s > s^*$:

$$\mu^*(s) = f(s) + \int_{\underline{s}}^{s^*} \lambda^*(a,s) da \geq 0.$$

□

Appendix B Beyond binary types

This appendix extends the discussion presented in Section 3. We present a simple version of the model, featuring 2 signals and 3 ex-post types, illustrating that pooling might persist beyind the binary types case.

2 signals and 3 ex-post types

The case of 2×2-dimensional type-space is a special case of the model presented in Section 2, so it results in optimal signal-pooling regardless of the information structure. Now, consider the case of 2 signals s_i, $i \in \{L, H\}$ and a 3 project types $\beta \in \{\beta_0, \beta_1, \beta_2\}$. Assume $\Pr(s_i) = \nu_i$, $\Pr(\beta_k|i) = q_{i,k}$ and $\beta_2 - \beta_1 = \beta_1 - \beta_0 = \Delta\beta$.

Let $e_{i,k}$ denote the level of cost-reducing effort recommended to a firm who reports signal s_i and project type β_k, and $\phi(e) \equiv \psi(e) - \psi(e - \Delta\beta)$ as in Section 2. In the discrete case, the (binding) ex-post individual rationality constraints are given by:

$$\begin{aligned} u_i(\beta_0) &= \phi(e_{i,1}) + \phi(e_{i,0}) \\ u_i(\beta_1) &= \phi(e_{i,0}) \\ u_i(\beta_2) &= 0, \qquad i = L, H. \end{aligned}$$

Plugging into the ex-ante individual rationality constraint for signal i, it obtains:

$$\begin{aligned} & \sum_k q_{i,k} \left[u_i(\beta_k) - u_j(\beta_k) \right] \geq 0, \qquad j \neq i \\ \Leftrightarrow \quad & q_{i,0} \left[\phi(e_{i,1}) - \phi(e_{i,1}) \right] + (q_{i,0} + q_{i,1}) \left[\phi(e_{i,2}) - \phi(e_{i,2}) \right] \geq 0. \end{aligned} \tag{IC$_{i,j}$}$$

Denote the unconditional β-distribution as $\bar{q}_k = \nu_L q_{L,k} + \nu_H q_{H,k}$, and the pooling mechanism as $\bar{u}(\beta_2) = 0$ and $\bar{e}_k$, $k \in \{0, 1, 2\}$, such that:

$$\begin{aligned} \psi'(\bar{e}_0) &= 1 \\ \psi'(\bar{e}_1) &= 1 - \frac{\bar{q}_0}{\bar{q}_1} \phi'(e_{1,k}) \\ \psi'(\bar{e}_2) &= 1 - \frac{\bar{q}_0 + \bar{q}_1}{\bar{q}_2} \phi'(e_{2,k}). \end{aligned} \tag{15}$$

Note the pooling mechanism in (15) satisfies monotonicity constraints $\bar{e}_k - \bar{e}_{k-1} \leq \Delta\beta$, for $k = 1, 2$, under the conventional monotone hazard rate assumption:[25]

$$\frac{\bar{q}_0}{\bar{q}_1} \leq \frac{\bar{q}_0 + \bar{q}_1}{\bar{q}_2}.$$

[25] This is the discrete version of $d\left[G(\beta)/g(\beta)\right]/d\beta \geq 0$, required for the monotonicity of optimal mechanisms with more than two types (see Laffont and Tirole, 1993, section 1.4).

The problem's Lagrangian under ex-post participation constraints becomes as follows:

$$\begin{aligned}\mathcal{L}^* = \quad & \sum_{i\in\{L,H\}} \nu_i \left\{ \sum_{k\in\{0,1,2\}} q_{i,k}\left[\beta_k - e_{i,k} + \psi(e_{i,k})\right] + q_{i,0}\phi(e_{i,1}) + (q_{i,0}+q_{i,1})\phi(e_{i,2}) \right\} \\ & - \nu_L\lambda_{LH}\left\{ q_{L,0}\left[\phi(e_{L,1}) - \phi(e_{H,1})\right] + (q_{L,0}+q_{L,1})\left[\phi(e_{L,2}) - \phi(e_{H,2})\right]\right\} \\ & \quad - \nu_H\lambda_{HL}\left\{ q_{H,0}\left[\phi(e_{H,1}) - \phi(e_{L,1})\right] + (q_{H,0}+q_{H,1})\left[\phi(e_{H,2}) - \phi(e_{L,2})\right]\right\} \\ & + \nu_L u_L(\beta_2)\left[1 - \lambda_{LH} + \lambda_{HL} - \mu_L\right] + \nu_H u_H(\beta_2)\left[1 - \lambda_{HL} + \lambda_{LH} - \mu_H\right]\end{aligned}$$

with associated first-order conditions, for $i \in \{L,H\}$:

$$q_{i,0}\left[1 - \psi'(e_{i,0})\right] = 0 \tag{16}$$

$$q_{i,1}\left[1 - \psi'(e_{i,1})\right] + q_{i,0}\phi'(e_{i,1}) - \lambda_{ij}q_{i,1}\phi'(e_{i,1}) + \lambda_{ji}q_{j,1}\phi'(e_{i,1}) = 0 \tag{17}$$

$$q_{i,2}\left[1 - \psi'(e_{i,2})\right] + (q_{i,0}+q_{i,1})\phi'(e_{i,2}) - \lambda_{ij}(q_{i,0}+q_{i,1})\phi'(e_{i,2}) + \lambda_{ji}(q_{j,0}+q_{j,1})\phi'(e_{i,2}) = 0 \tag{18}$$

$$u_i(\beta_2)\left[1 - \lambda_{ij} + \lambda_{ji} - \mu_i\right] \geq 0. \tag{19}$$

Notice from summing (19) for $i \in \{L,H\}$ that there exist $\mu_L, \mu_H \geq 0$ that support active ex-post participation constraints if, and only if

$$\lambda_{LH}, \lambda_{HL} \in (0,1)\,.$$

Moreover, the pooling mechanism (15) always satisfies the FOC (16), as it recommends first-best effort to the least expensive project type. Evaluating (17) and (18) at the pooling allocations $\bar{e}_1$ and $\bar{e}_2$, and canceling out the $\phi'(.)$ terms, it obtains, for $i \in \{L,H\}$:

$$\begin{aligned}\frac{q_{i,0}}{q_{i,1}}\lambda_{ij} - \frac{q_{j,0}}{q_{i,1}}\lambda_{ji} &= \frac{q_{i,0}}{q_{i,1}} - \frac{\bar{q}_0}{\bar{q}_1} \\ \frac{q_{i,0}+q_{i,1}}{q_{i,2}}\lambda_{ij} - \frac{q_{j,0}+q_{j,1}}{q_{i,2}}\lambda_{ji} &= \frac{q_{i,0}+q_{i,1}}{q_{i,2}} - \frac{\bar{q}_0+\bar{q}_1}{\bar{q}_2}.\end{aligned}$$

We may restrict attention to the previous system for $i = L$, as there are only two linearly independent equations. Denote the hazard rates $\bar{h}(\beta_k) = \sum_{l=0}^{k-1}\bar{q}_l/\bar{q}_k$ and $h_{i,j}(\beta_k) = \sum_{l=0}^{k-1} q_{j,l}/q_{i,k}$, with $h_i(\beta) \equiv h_{i,i}(\beta)$.[26] Stating the FOC system in matrix form, we get:

$$\begin{bmatrix} h_L(\beta_1) & -h_{L,H}(\beta_1) \\ h_L(\beta_2) & -h_{L,H}(\beta_2)\end{bmatrix} \cdot \begin{bmatrix}\lambda_{LH} \\ \lambda_{HL}\end{bmatrix} = \begin{bmatrix} h_L(\beta_1) - \bar{h}(\beta_1) \\ h_L(\beta_2) - \bar{h}(\beta_2)\end{bmatrix}$$

Operating the linear system above using Cramer's rule, one may check that

$$\frac{h_L(\beta_2)}{h_L(\beta_1)} \geq \frac{\bar{h}(\beta_2)}{\bar{h}(\beta_1)} \geq \frac{h_{L,H}(\beta_2)}{h_{L,H}(\beta_1)} \tag{20}$$

is a sufficient condition for the solution $\lambda^*_{LH}, \lambda^*_{LH} \in (0,1)$. Therefore, if the information structure satisfies (20), the pooling mechanism (15) is an optimal solution to the principal's design problem under exit rights.[27]

[26] The term $h_{i,j}(.)$ corresponds to Krähmer and Strausz (2015)'s definition of "cross-hazard rate".

[27] We conjecture (20) is also necessary for the optimality of pooling.

References

Arrow, K. J. and A. C. Enthoven (1961). Quasi-concave programming. *Econometrica 29*(4), 779–800.

Arvan, L. and A. P. N. Leite (1990, September). Cost overruns in long term projects. *International Journal of Industrial Organization 8*(3), 443–467.

Ashlagi, I., C. Daskalakis, and N. Haghpanah (2022). Sequential mechanisms with ex post individual rationality. *Operations Research 0*(0), null.

Bajari, P. and S. Tadelis (2001). Incentives versus transaction costs: A theory of procurement contracts. *RAND Journal of Economics 32*(3), 387–407.

Baron, D. P. and D. Besanko (1984). Regulation and information in a continuing relationship. *Information Economics and Policy 1*(3), 267–302.

Bergemann, D., F. Castro, and G. Y. Weintraub (2020). The scope of sequential screening with ex post participation constraints. *Journal of economic theory 188*, 105055–.

Bergemann, D. and J. Välimäki (2019, June). Dynamic mechanism design: An introduction. *Journal of Economic Literature 57*(2), 235–74.

Calveras, A., J.-J. Ganuza, and E. Hauk (2004). Wild bids. gambling for resurrection in procurement contracts. *Journal of Regulatory Economics 26*(1), 41–68.

Castro-Pires, H. and H. Moreira (2021). Limited liability and non-responsiveness in agency models. *Games and Economic Behavior 128*, 73–103.

Chakraborty, I., F. Khalil, and J. Lawarree (2021). Competitive procurement with ex post moral hazard. *The RAND Journal of Economics 52*(1), 179–206.

Courty, P. and H. Li (2000). Sequential screening. *Review of Economic Studies 67*(4), 697–717.

Flyvbjerg, B., M. K. S. holm, and S. L. Buhl (2003). How common and how large are cost overruns in transport infrastructure projects? *Transport Reviews 23*(1), 71–88.

Freixas, X., R. Guesnerie, and J. Tirole (1985). Planning under incomplete information and the ratchet effect. *The review of economic studies 52*(2), 173–191.

Gerardi, D. and L. Maestri (2020). Dynamic contracting with limited commitment and the ratchet effect. *Theoretical Economics 15*(2), 583–623.

Gottlieb, D. and H. Moreira (2021, May). Simple contracts with adverse selection and moral hazard. *Theoretical Economics* (forthcoming).

Guasch, J. (2004). *Granting and Renegotiating Infrastructure Concessions: Doing It Right*. WBI Development Studies. World Bank.

Krähmer, D. and R. Strausz (2015). Optimal sales contracts with withdrawal rights. *Review of Economic Studies 82*(2), 762–790.

Krähmer, D. and R. Strausz (2016). Optimality of sequential screening with multiple units and ex post participation constraints. *Economics Letters 142*(C), 64–68.

Krähmer, D. and R. Strausz (2022). Dynamic screening with verifiable bankruptcy. *Working paper*.

Krasikov, I. and R. Lamba (2021). A theory of dynamic contracting with financial constraints. *Journal of Economic Theory 193*, 105196.

Krishna, R. V., G. Lopomo, and C. R. Taylor (2013). Stairway to heaven or highway to hell: Liquidity, sweat equity, and the uncertain path to ownership. *The RAND Journal of Economics 44*(1), 104–127.

Krähmer, D. and R. Strausz (2011, 02). Optimal Procurement Contracts with Pre-Project Planning. *The Review of Economic Studies 78*(3), 1015–1041.

Laffont, J.-J. and J. Tirole (1986). Using cost observation to regulate firms. *Journal of Political Economy 94*(3), 614–41.

Laffont, J.-J. and J. Tirole (1987a). Auctioning incentive contracts. *Journal of Political Economy 95*(5), 921–37.

Laffont, J.-J. and J. Tirole (1987b). Comparative statics of the optimal dynamic incentive contract. *European Economic Review 31*(4), 901–926.

Laffont, J.-J. and J. Tirole (1988). The dynamics of incentive contracts. *Econometrica 56*(5), 1153–75.

Laffont, J.-J. and J. Tirole (1990). Adverse selection and renegotiation in procurement. *The Review of economic studies 57*(4), 597–625.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. The MIT Press.

Luenberger, D. (1997). *Optimization by Vector Space Methods*. Professional Series. Wiley.

Makovšek, D. and A. Bridge (2021). Procurement choices and infrastructure costs. In *Economic Analysis and Infrastructure Investment*, pp. 277–327. National Bureau of Economic Research, Inc.

McAfee, R. P. and J. McMillan (1987, Summer). Competition for Agency Contracts. *RAND Journal of Economics 18*(2), 296–307.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Myerson, R. (1986). Multistage games with communication. *Econometrica 54*(2), 323–58.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Ollier, S. and L. Thomas (2013). Ex post participation constraint in a principal–agent model with adverse selection and moral hazard. *Journal of Economic Theory 148*(6), 2383–2403.

Pavan, A., I. Segal, and J. Toikka (2014). Dynamic mechanism design: A myersonian approach. *Econometrica 82*(2), 601–653.

Riordan, M. H. and D. E. M. Sappington (1987, June). Awarding Monopoly Franchises. *American Economic Review 77*(3), 375–387.

Siemiatycki, M. (2015). *Cost overruns on infrastructure projects: Patterns, causes, and cures*. Institute on Municipal Finance and Governance.

Toikka, J. (2011). Ironing without control. *Journal of Economic Theory 146*(6), 2510–2526.