Theoretical Economics 17 (2022), 1357–1401 1555-7561/20221357

Simple contracts with adverse selection and moral hazard

Daniel Gottlieb
Department of Management, London School of Economics

Humberto Moreira
Brazilian School of Economics and Finance, FGV EPGE

We study a principal–agent model with moral hazard and adverse selection. Risk-neutral agents with limited liability have arbitrary private information about the distribution of outputs and the cost of effort. We show that under a multiplicative separability condition, the optimal mechanism offers a single contract. This condition holds, for example, when output is binary. If the principal's payoff must also satisfy free disposal and the distribution of outputs has the monotone likelihood ratio property, the mechanism offers a single debt contract. Our results generalize if the output distribution is "close" to multiplicatively separable. Our model suggests that offering a single contract may be optimal in environments with adverse selection and moral hazard when agents are risk-neutral and have limited liability.

Keywords. Principal–agent problem, contract theory, mechanism design.

JEL classification. D82, D86.

1. Introduction

Real-world contracts are often simpler than theory predicts. Unlike standard adverse selection models, contracting parties offer a limited number of contracts, usually a single one. Unlike in standard moral hazard models, similar contracts are offered in fundamentally different environments. As Hart and Holmstrom (1987) and Chiappori and Salanie (2003) argue in their surveys of the literature:

> The extreme sensitivity to informational variables that comes across from this type of modeling is at odds with reality. Real world schemes are simpler than the theory would dictate and surprisingly uniform across a wide range of circumstances. *(Hart and Holmstrom, 1987, p. 105)*

Daniel Gottlieb: d.gottlieb@lse.ac.uk
Humberto Moreira: humberto.moreira@fgv.br

We thank Eduardo Azevedo, Dirk Bergemann, Vinicius Carrasco, Sylvain Chassang, Gonzalo Cisternas, Alex Edmans, Mehmet Ekmekci, Eduardo Faingold, Leandro Gorno, Faruk Gul, Jason Hartline, Tibor Heumann, Bengt Holmström, Johannes Hörner, Ohad Kadan, Lucas Maestri, George Mailath, David Martimort, Stephen Morris, Roger Myerson, Larry Samuelson, Yuliy Sannikov, Jean Tirole, Rakesh Vohra, John Zhu, Alessandro Pavan, and two anonymous referees for comments and suggestions. Rafael Mourão provided outstanding research assistance. Gottlieb gratefully acknowledges financial support from the Dorinda and Mark Winkelman Distinguished Scholar Award. 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.

 https://doi.org/10.3982/TE2992

> The recent literature (...) provides very strong evidence that contractual forms have large effects on behavior. As the notion that "incentives matter" is one of the central tenets of economists of every persuasion, this should be comforting to the community. On the other hand, it raises an old puzzle: if contractual form matters so much, why do we observe such a prevalence of fairly simple contracts? *(Chiappori and Salanie, 2003, p. 34)*

More recently, a literature has studied how simplicity may arise from a desire to offer a contract that is robust to informational assumptions.[1] In this paper, we follow the more traditional mechanism design approach and study the optimality of offering a single contract in a general model with adverse selection, moral hazard, and limited liability.

We consider a principal–agent relationship with bilateral risk neutrality and limited liability. In reality, virtually all contracting parties have limited liability. Entrepreneurs raising capital from investors enjoy limited liability as the value of their equity cannot fall below zero, and minimum wage laws enforce limited liability in employment contracts. Moreover, adverse selection and moral hazard are jointly present in many contracting situations. Managers, for example, take actions that affect the firm's profitability. Usually, managers also have better knowledge about the efficacy of each action.

The agent selects an unobserved "effort" from a very general effort space. The agent also has private information, in an arbitrary way, about the distribution of outputs and effort costs, resulting in a model where types and efforts are multidimensional (possibly infinite dimensional) and unordered. The interaction between adverse selection, moral hazard, and limited liability implies that, while it is generally possible to screen agents, it may not be cost-effective to do so.

Under a multiplicative separability condition, the optimal mechanism offers a single contract to all agents regardless of the type space or the distribution of types. This condition—which always holds if output is binary or under the spanning condition of Grossman and Hart (1983)—is equivalent to assuming that agents rank the "power" of any contracts equally. For example, with binary outputs, the power of a contract is determined solely by its bonus.

The intuition for our main result is as follows. Starting with an arbitrary menu of contracts, suppose the principal removes all but the contract with the highest power. Limited liability ensures that all agent types still participate in the mechanism. There are two effects. First, efficiency increases since agents receive a contract with a higher power. Second, because agents are risk-neutral, they always pick the contract that maximizes their expected payment conditional on their effort. Thus, holding effort fixed, limiting the choice of contracts offered to the agent reduces their expected payments. Since both effects increase the principal's profits, offering multiple contracts is suboptimal. If the output distribution also satisfies the monotone likelihood ratio property and the principal's payoff must be monotone ("free disposal"), the optimal mechanism consists of the principal taking a single debt contract.

Our results are not knife-edge in the sense that it is still generically optimal to offer a single contract if the distribution is close to multiplicatively separable. More broadly,

[1] See Carroll (2019) and Bergemann and Morris (2005) for a survey.

our paper shows that offering flexibility to agents through menus of contracts can hurt the principal. Because the agent is risk-neutral, he always chooses the contract with the highest expected payment conditional on his effort, which is precisely the contract with the highest cost to the principal. That is, holding the agent's effort fixed, reducing the number of contracts always increases the principal's profits. In particular, when the principal can identify the contract with the highest power (i.e., when multiplicative separability holds), she can simultaneously reduce informational rents and increase efficiency by removing all other contracts.

Although the framework we study has been widely applied to financial contracting, it has many other applications. One such application is procurement and regulation. Despite the central role that menus of contracts play in the theory of procurement and regulation, they are rarely observed in practice.[2] Accordingly, many papers try to identify conditions for simple procurement contracts to be close to optimal.[3] We generalize the classic model of Laffont and Tirole (1986, 1993) by allowing effort to affect the regulated firm's costs stochastically and assuming that the firm has limited liability. We then obtain conditions for the optimal mechanism to offer a single contract and for the optimal contract to be a price cap. Since limited liability constraints are a key aspect of most procurement contracts (see, e.g., Burguet, Ganuza, and Hauk, 2012), our model provides an explanation for the absence of menus of contracts in procurement.

Our results also contribute to an applied literature by identifying assumptions under which researchers can focus on a single contract when solving their models, simplifying the task of obtaining comparative statics results. For example, we show that under multiplicative separability, there is no loss of generality in restricting attention to a single debt contract when one introduces adverse selection in the pure moral hazard model of Innes (1990).

Related literature

We consider a principal–agent relationship with bilateral risk neutrality and limited liability, as commonly studied in corporate finance (see Tirole, 2005). We build on this standard environment by adding adverse selection in an arbitrary way and allowing effort to be multidimensional. Our work is related to a literature that identifies conditions for contracts to take the form of debt and for equilibria to have complete pooling.

[2]For example, Bajari and Tadelis (2001) argues that "the descriptive engineering and construction management literature (...) suggests that menus of contracts are not used. Instead, the vast majority of contracts are variants of simple fixed-price (FP) and cost-plus (C+) contracts." In her survey of the literature, Netz (2000) argues that "price-cap regulation is the most commonly discussed and used form of incentive regulation."

[3]Using the Laffont–Tirole framework, Rogerson (2003) and Chu and Sappington (2007) show that a pair of simple contracts can achieve a large fraction of the surplus under a certain range of parametric settings—75 or 73 percent when costs follow either uniform or power distributions, respectively—for quadratic costs. Bajari and Tadelis (2001) assumes that there is a fixed cost of specifying each state of nature in the contract to rationalize the simplicity of observed contracts.

In a single-task setting with pure moral hazard, limited liability, and free disposal, Innes (1990) shows that contracts take the form of debt if the distribution of output satisfies the monotonicity of the likelihood ratio property.[4] Our main focus is on the lack of menus of contracts, which, of course, can only be addressed by introducing adverse selection. Nevertheless, our Theorem 2, which obtains conditions for a single debt contract to be optimal, extends their result to settings that also have adverse selection under multiplicative separability.[5]

In a signaling model of financial contracting (pure adverse selection), Nachman and Noe (1994) show that there is complete pooling if and only if firm types are strictly ordered by conditional stochastic dominance. When firms are ordered by conditional stochastic dominance, investors face a lemons problem: while they would like to offer better terms to healthier firms, less healthy firms are more likely to accept each contract. Our paper emphasizes different forces that may lead to pooling. In their model, pooling occurs when the distribution of types induces a market breakdown for all but the worst contract. In our model, pooling happens because of moral hazard and limited liability: giving flexibility to agents requires the principal to leave excessive rents. For example, when output is binary, complete pooling occurs in our model for any parameters of the model (i.e., regardless of whether types are ordered). Similarly, Demarzo and Duffie (1999) consider a signaling model of security design and show that under a uniform worst-case condition and a free disposal constraint, equilibrium contracts take the form of debt. Using this model, several authors studied whether intermediaries pool different assets in equilibrium. Biais and Mariotti (2005) embed the framework in a mechanism design setting and show that there is some screening, though a coarse one (all or nothing). Attar, Marriotti, and Salanié (2011) studies competition with nonexclusive contracts among sellers of perfectly divisible goods. In equilibrium, sellers sell either zero or all of their assets and the equilibrium resembles that in the canonical market for lemons.[6]

In a one-dimensional screening setting (pure adverse selection), Guesnerie and Laffont (1984) showed that optimal mechanisms are "nonresponsive" when the first-best allocation is decreasing. This occurs because optimality clashes with incentive compatibility, which requires allocations to be nondecreasing. The reason for pooling in our model is different from nonresponsiveness. For example, if the agent only has private

[4]Poblete and Spulber (2012) generalizes the analysis of Innes (1990) by introducing a critical ratio notion that captures the returns to providing incentives for effort. They show that it is optimal to offer a debt contract if this ratio is nondecreasing.

[5]See Jewitt, Kadan, and Swinkels (2008) for a general analysis of moral hazard models with limited liability. Moral hazard models with bilateral risk neutrality, limited liability, and free disposal include, for example, Matthews (2001), Dewatripont, Legros, and Matthews (2003), Poblete and Spulber (2012), and Chaigneau, Edmans, and Gottlieb (2018). Adverse selection models in this setting include Nachman and Noe (1994), Demarzo and Duffie (1999), DeMarzo (2005), and DeMarzo, Kremer, and Skrzypacz (2005).

[6]There is also a literature that obtains the optimality of debt contracts in settings with costly state verification (cf. Townsend, 1979) or with nonpecuniary penalties for default (Diamond, 1984; Bolton and Scharfstein, 1990). As Faure-Grimaud and Mariotti (1999) show, the optimality of debt in these settings is intrinsically related to the failure of the single-crossing condition. Riley and Zeckhauser (1983) show that a seller of a nondivisible good with commitment prefers to set a single price rather than engage in mechanisms that allow for dynamic screening of the consumer's willingness to pay.

information about the distribution of output, the first-best is increasing and therefore implementable in a pure adverse selection environment. Nevertheless, with multiplicative separability, the principal offers at most one contract. More related to our work, Ollier and Thomas (2013) substitute the traditional (interim) participation constraint by an ex post constraint in a one-dimensional model with binary outcomes. They show that under conditions that ensure that the first-order approach holds, there is no benefit from screening.[7]

As argued previously, our application to procurement and regulation builds on Laffont and Tirole (1986, 1993). In their model, there is both adverse selection and moral hazard. However, because the link between effort, types, and output is deterministic, the model can be reduced to a pure adverse selection model. For this reason, it is often referred to as a model with "false moral hazard" (cf. Laffont and Martimort, 2002). We allow effort to affect the regulated firm's costs stochastically so the problem cannot be reduced to a pure adverse selection model. Picard (1987), Melumad and Reichelstein (1989), and Caillaud, Guesnerie, and Rey (1992) also introduce noise in the relationship between output and effort, and show that under certain conditions, the principal can achieve the same utility as in the absence of noise. Therefore, unlike in our model, they find that there is no cost from moral hazard. Our model differs from theirs in two ways. First, they assume that all private information concerns the cost of effort, whereas we also allow the agent to have private information about the distribution of output. Introducing this dimension of private information makes moral hazard costly. Second, they do not assume that agents have limited liability. Limited liability also prevents the principal from eliminating moral hazard at no cost.

In Section 2, we present the general model and show our main result. In Section 3, we introduce free disposal and obtain conditions for the optimality of debt. Then Section 4 concludes. All proofs are in the Appendix.

2. Model

There is a risk-neutral principal and a risk-neutral agent with limited liability. The agent has private information about the environment, captured by a type $\theta \in \Theta$. From the principal's perspective, types are distributed according to a probability measure μ. Our formulation allows types to be finite or infinite dimensional, and their distribution may be discrete or continuous. We will discuss the assumptions on Θ and μ below.

The agent chooses an effort e from the compact metric space E. Exerting effort e costs c_e^θ to type θ. We assume that the least-costly effort has a nonpositive cost:

[7]Their results rely on the presence of countervailing incentives, whereas we show that each feasible mechanism is weakly dominated by a mechanism that offers a single contract. With our approach, we are able to generalize their result to settings with no conditions on costs or distributions and with arbitrary type and effort spaces. Chade and Swinkels (2019) also analyze a principal–agent problem with both moral hazard and adverse selection. They consider a risk-averse agent without limited liability and propose a decoupling approach. Chade and Schlee (2020) show that when a monopolistic firm has additional cost of providing insurance, coverage denials to the worst risk and full pooling can be optimal.

$\min_{e\in E} c_e^\theta \leq 0$ for all θ. This condition is satisfied in standard frameworks where the lowest effort costs zero, as well as in more general multitask frameworks that allow the agent to derive private benefits from certain actions.[8]

The principal does not directly observe the effort chosen by the agent. Instead, she observes the output from the partnership, which is stochastically affected by the agent's effort. Let $X := \{x_1, \dots, x_N\}$ be the set of possible (real-valued) outputs with $x_1 < \cdots < x_N$. A type-θ agent who exerts effort e produces output x_i with probability $p_e^\theta(x_i) := \Pr(x = x_i|\theta, e)$.[9] Our model allows—but does not require—the agent to have private information about both the distribution of outputs and effort costs. The case where the cost of effort e is common knowledge, for example, is accommodated by letting c_e^θ be constant in θ. Moreover, we do not impose any particular structure on the space of types and efforts (besides the technical requirement of compactness). There is no need to impose an order structure on them. Moreover, types and effort may be complements, substitutes, or neither in terms of output distributions and costs.

To simplify notation, we consider deterministic mechanisms. We generalize our results to random mechanisms in the Supplemental Material, available in a supplementary file on the journal website, http://econtheory.org/supp/2292/supplement.pdf. By the revelation principle, we can focus on direct mechanisms. A direct mechanism consists of a contract and an effort recommendation for each type. A contract specifies a payment to the agent contingent on each output. Formally, let $\mathcal{B}(\Theta)$ denote the Borel σ-field of Θ. A direct mechanism is a pair of $\mathcal{B}(\Theta)$-measurable functions $w : \Theta \times X \to \mathbb{R}$ and $e : \Theta \to E$, so that a type-θ agent is recommended effort $e(\theta)$ and gets paid $w^\theta(x)$ in case of output x.

Given a mechanism (w, e), a type-θ agent gets expected payoff

$$U(\theta) := \sum_{i=1}^{N} p_{e(\theta)}^\theta(x_i) w^\theta(x_i) - c_{e(\theta)}^\theta.$$

The mechanism has to satisfy the *incentive compatibility (IC)*, *individual rationality (IR)*, and *limited liability (LL)* constraints:

$$U(\theta) \geq \sum_{i=1}^{N} p_{\hat{e}}^\theta(x_i) w^{\hat{\theta}}(x_i) - c_{\hat{e}}^\theta \quad \forall \theta, \hat{\boldsymbol{\theta}}, \hat{e} \tag{IC}$$

$$U(\theta) \geq 0 \quad \forall \theta \tag{IR}$$

$$w^\theta(x_i) \geq 0 \quad \forall \theta, i. \tag{LL}$$

[8]Allowing the cost of the lowest effort to be positive makes participation random as in Rochet and Stole (2002). In Appendix D, we present a stronger sufficient condition under which our main result generalizes to this case. However, in the Supplemental Material, we also show that our result may fail when this condition does not hold even with binary output.

[9]The assumption of finitely many outputs is helpful to ensure the existence of an optimal mechanism. It is straightforward to generalize our results if contract payments are required to be bounded (such as, for example, if both parties have limited liability).

An *optimal* mechanism maximizes the principal's expected profit,

$$\int_\Theta \sum_{i=1}^{N} p^\theta_{e(\theta)}(x_i)\big[x_i - w^\theta(x_i)\big]\, d\mu(\theta), \tag{1}$$

among mechanisms that satisfy IC, IR, and LL. We say that an optimal mechanism is *essentially unique* if, for almost all types, any other optimal mechanism gives the same payoff to both the principal and the agent.

To ensure the existence of an optimal mechanism, we make the following technical assumptions, which are satisfied by all standard agency models.

ASSUMPTION 1. *The space Θ is measurable, μ is a probability measure on $\mathcal{B}(\Theta)$, $(e,\theta) \longmapsto c^\theta_e$, and $(e,\theta) \longmapsto p^\theta_e(x_i)$ are continuous functions for all x_i, and $p^\theta_e(x_i) \geq \underline{p}$ for some $\underline{p} > 0$.*

The example below describes the case of binary outputs, which will be useful in illustrating our results.

EXAMPLE 1. Let $N = 2$. We refer to x_2 as *success*, x_1 as *failure*, and $\Delta x := x_2 - x_1$ as the *incremental output*. A contract consists of a *fixed payment* $w^\theta(x_1)$ and a *bonus* $w^\theta(x_2) - w^\theta(x_1)$. If effort is binary, each type can be described by the four-dimensional vector $(p^\theta_0(x_2), p^\theta_1(x_2), c^\theta_0, c^\theta_1)$ that specifies the probability of success conditional on each effort and the cost of each effort. If effort is continuous, each type is described by the infinite-dimensional function $(p^\theta_\cdot(x_2), c^\theta_\cdot) : E \to \mathbb{R}^2$ specifying the probability of success and the cost of each effort. ♢

Multiplicative separability We now introduce a class of economies that will play a key role in our analysis.

DEFINITION 1. A distribution p^θ_e satisfies *multiplicative separability* (MS) if there exist functions $h : X \to \mathbb{R}$ and $I : E \times \Theta \to \mathbb{R}$ such that

$$p^\theta_e(x) + I(e,\theta)h(x) = p^\theta_{\tilde{e}}(x) + I(\tilde{e},\theta)h(x) \quad \forall e, \tilde{e}, \theta, x. \tag{2}$$

Multiplicative separability always holds when there are only two outputs. It also holds under the linearity of the distribution function condition (Grossman and Hart (1983) and Hart and Holmstrom (1987)), which is obtained by taking $I(e,\theta) \in [0,1]$ and $h(x) = p_0(x) - p_1(x)$ for some distributions p_0 and p_1:

$$p^\theta_e(x) = I(e,\theta)p_1(x) + \big[1 - I(e,\theta)\big]p_0(x). \tag{3}$$

This condition is commonly used in pure moral hazard models, along with the convexity of costs, to justify the first-order approach. Our results, however, do not rely on the validity of the first-order approach.[10]

[10]There is no relationship between MS and Holmstrom's (1979) sufficient statistic result. For example, MS always holds with binary outputs. However, as long as the likelihood ratio $\frac{p^\theta_{e_H}(x)}{p^\theta_{e_L}(x)}$ is not constant in θ for

Multiplicative separability can be characterized in terms of rankings between different efforts. In Appendix C, we show that MS is equivalent to the following condition. For any two contracts satisfying LL, w, and $\tilde{w}$, if there exist $\theta_0 \in \Theta$, $e_0, \tilde{e}_0 \in E$ for which $p_{e_0}^{\theta_0} \neq p_{\tilde{e}_0}^{\theta_0}$ and

$$\sum_{i=1}^{N}\big[p_{e_0}^{\theta_0}(x_i) - p_{\tilde{e}_0}^{\theta_0}(x_i)\big]w(x_i) = \sum_{i=1}^{N}\big[p_{e_0}^{\theta_0}(x_i) - p_{\tilde{e}_0}^{\theta_0}(x_i)\big]\tilde{w}(x_i),$$

then

$$\sum_{i=1}^{N}\big[p_{e}^{\theta}(x_i) - p_{\tilde{e}}^{\theta}(x_i)\big]w(x_i) = \sum_{i=1}^{N}\big[p_{e}^{\theta}(x_i) - p_{\tilde{e}}^{\theta}(x_i)\big]\tilde{w}(x_i)$$

for all $\theta \in \Theta$ and all $e, \tilde{e} \in E$. In words: if one type has the same incentive to exert two efforts with different distributions under contracts w and $\tilde{w}$, so do all other types and all efforts. That is, all types have the same ordering of contracts in terms of incentives. Of course, they may still pick different efforts depending on their distributions and effort costs.[11]

For example, with two outputs, a contract's incentives are uniquely determined by the bonus it pays in case of success, which is a real number. All types agree that a contract with a higher bonus gives more incentives than one with a lower bonus. Therefore, with two outputs, incentives can be ordered according to the bonus, so that MS always holds.

MS also has a simple geometric interpretation (see Appendix C). Suppose there are three outputs, so the probabilities conditional on each type and effort lie on the two-dimensional simplex. MS holds if and only if

$$p_{e}^{\theta}(x_2) - p_{\tilde{e}}^{\theta}(x_2) = \phi\big[p_{e}^{\theta}(x_3) - p_{\tilde{e}}^{\theta}(x_3)\big] \quad \forall \theta, e, \tilde{e}$$

for some constant ϕ. Figure 1 illustrates this condition graphically. Each point in the graph corresponds to the probability distribution conditional on a type and an effort. We draw two points with the same color if they correspond to the same type but different efforts. MS requires the lines connecting these points to have the same slope (ϕ). More generally, MS requires all efforts to have the same proportional effect on the probability of each outcome for all types. With two outputs, all distributions must lie on the same horizontal line, so MS automatically holds.

some efforts e_L and e_H, x will not be a sufficient statistic for e given (x, θ). Moreover, because types also affect the cost of effort, the optimal pure moral hazard contract is also a function of θ even if types do not affect the likelihood ratio.

[11]MS does not imply that types cannot be screened. In general, when faced with a menu of contracts, different types may pick different contracts. However, as we will show below, when MS holds, it is not cost-effective to offer menus of contracts.

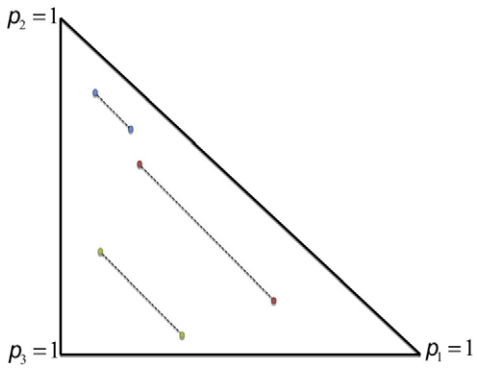

FIGURE 1. MS with three outputs.

2.1 *Single contract*

We now show that, whenever MS holds, the optimal mechanism offers the same contract to all types.[12] Different types may still choose different efforts depending on their output distributions and effort costs.

THEOREM 1. *Suppose MS holds. There exists an essentially unique optimal mechanism that offers a single contract to all types.*

To understand the theorem, it is easier to start with the case of two outputs. Note first that IC and LL imply that IR never binds (since the agent can always ensure a nonnegative payoff by exerting the least costly effort and payments are nonnegative by LL). Moreover, with two outputs, any mechanism that pays a bonus greater than the incremental output to some type gives the principal a payoff that is lower than the contract that offers all types a constant payment of zero. Therefore, any such mechanism cannot be optimal.

The key observation is that for any mechanism with bonuses lower than the incremental output, if multiple contracts are being offered, the principal can improve by removing all but the contract with the highest bonus from the mechanism. Since IR does not bind, all agents pick the remaining contract once the other ones are removed. There are two effects from this migration to the contract with the highest bonus: a reduction in rents and an increase in efficiency.

By incentive compatibility, the risk-neutral agents pick the contract that maximizes their expected payments conditional on their effort choice. Thus, holding effort constant, making agents switch to a contract that was also available to them reduces their expected payments (rent reduction). But switching their contract does not leave efforts constant. Instead, offering only the contract with the highest bonus induces agents to pick efforts with a higher probability of success, which increases the total surplus (efficiency gain). More precisely, efficiency gain by migrating each type to the contract with

[12]MS can be slightly weakened since it does not need to hold for all efforts, only the one that principal wants to implement. For example, Theorem 1 remains unchanged if (2) fails at points where $p_\cdot^\theta(x)$ and $c_\cdot^\theta$ are locally concave (since such efforts are not implementable).

the highest bonus equals the increased probability of success times the incremental output net of the bonus paid to the agent,

$$(p_{\tilde{e}} - p_e)(\Delta x - b),$$

where e is the agent's old effort, $\tilde{e}$ is the agent's new effort, and $p_{\tilde{e}} \geq p_e$ because the new bonus exceeds the old one. Since the incremental output exceeds the bonus, both terms are positive, showing that efficiency increases for each type with the substitution.

The rent reduction effect—removing contracts reduces the expected payment to all types—remains unchanged with multiple outputs. The main difficulty with multiple outputs is with the efficiency effect. Without MS, there may not be a single contract that provides the highest incentives to all agent types. Moreover, even when MS holds, so that contracts can be ordered based on their incentives, it is not always optimal to offer the greatest incentives.

High-powered contracts may be unprofitable for two reasons. First, they may incentivize an inefficient effort. Second, because some bonuses may exceed the incremental output (i.e., $x_{i+1} - x_i < w(x_{i+1}) - w(x_i)$ for some i), improving the distribution of outputs may decrease the principal's profits.[13] There are three possible cases. If we start with a mechanism in which the principal would like to encourage effort for all types, replacing contracts by the one with the highest power (which exists by MS) increases profits. If we start with a mechanism in which the principal would like to discourage effort for all types, it is profitable to replace all contracts by the one with the lowest power (which also exists by MS). Last, if the principal would like to encourage effort by some types and discourage effort by some other types, it is possible to identify a single contract that leaves all effort incentives constant. Therefore, the second effect is also nonnegative.

Note that limited liability and risk neutrality play an important role in Theorem 1. Limited liability ensures that agents do not leave the mechanism if their contract is removed. Without it, the participation constraint would bind for some type. Then, removing their contracts would induce them to prefer not to participate, and the principal would face a trade-off between extracting rents from types who participate and excluding some types. Risk neutrality implies that, holding effort fixed, the principal and the agent split a pie of a fixed size. Since the agent always picks the contract with the highest expected payment, providing more freedom of choice to the agent can only hurt the principal (holding effort fixed). With risk aversion, different bonuses also affect the size of the pie since lower bonuses insure the agent better. Then, removing low-powered contracts improves efficiency but worsens risk-sharing.

The next proposition determines the states in which LL binds:

PROPOSITION 1. *Suppose MS holds. Let w^* be an optimal contract and let $e(\theta)$ be an optimal effort for type θ when offered contract w^*. Then, either (i) $w^*(x_i) = 0$ for all $i \notin \arg\max_{\tilde{i}}\{\frac{h(x_{\tilde{i}})}{\int_\Theta p^\theta_{e(\theta)}(x_{\tilde{i}})\,d\mu(\theta)}\}$ or (ii) $w^*(x_i) = 0$ for all $i \notin \arg\min_{\tilde{i}}\{\frac{h(x_{\tilde{i}})}{\int_\Theta p^\theta_{e(\theta)}(x_{\tilde{i}})\,d\mu(\theta)}\}$.*

[13] If the principal has access to a free disposal technology, as we discuss in Section 3, bonuses cannot exceed the incremental output, so the second concern discussed above is not an issue.

To understand Proposition 1, recall that in models of pure moral hazard, payments depend on the likelihood ratio between the effort being implemented and the effort to which the agent is most tempted to deviate (see Holmstrom (1979)). Since agents are risk-neutral in our model, with pure moral hazard, each agent's optimal contract would pay zero in all outcome realizations except for the one with the highest likelihood ratio between the effort that the principal wants to implement and the one with a binding IC. Under MS, the distribution of outputs satisfies

$$\frac{p^\theta_{\tilde{e}}(x_i)}{p^\theta_e(x_i)} - 1 = -\big[I(\tilde{e},\theta) - I(e,\theta)\big]\frac{h(x_i)}{p^\theta_e(x_i)},$$

so the likelihood ratio between efforts e and $\tilde{e}$ for type θ is maximized either at the outcome with the highest or with the lowest ratio $\frac{h(x_i)}{p^\theta_e(x_i)}$ (depending on whether $I(\tilde{e},\theta) < I(e,\theta)$ or $I(\tilde{e},\theta) > I(e,\theta)$, respectively). With adverse selection and MS, Theorem 1 shows that the principal offers a single contract to all types. Then the relevant ratio replaces each type's individual probability distribution $p^\theta_{e(\theta)}(x_i)$ by the average probability $\int_\Theta p^\theta_{e(\theta)}(x_i)\,d\mu(\theta)$. In particular, an optimal contract can only pay a positive amount in multiple outcomes if they have the same ratios $\frac{h(x_i)}{\int_\Theta p^\theta_{e(\theta)}(x_i)\,d\mu(\theta)}$. Therefore, for generic distributions, optimal contracts pay zero in all but one outcome realization.

In particular, since MS always holds when there are two outputs, Proposition 1 implies that the optimal mechanism offers a single contract, which pays zero in case of failure and a positive bonus in case of success.

COROLLARY 1. *Let $N = 2$. There exists an essentially unique optimal mechanism that offers a single contract to all types, with a zero fixed payment $(w(x_1) = 0)$ and a bonus lower than the incremental output $(w(x_2) < \Delta x)$.*

We now show that the moral hazard component is important for the optimality of offering the same contract to all types.

EXAMPLE 2 (Observable Effort). There are two outputs, two types (A and B), and two efforts (0 and 1, or "low" and "high"). Effort costs are $c^A_1 = 1$, $c^B_1 = \frac{2}{3}$, and $c^A_0 = c^B_0 = 0$. The probability of success conditional on a high effort equals $p^A_1 = \frac{2}{3}$ for type A and $p^B_1 = \frac{1}{3}$ for type B. We assume that a project fails with a sufficiently high probability if they exert low effort and take $x_H - x_L$ to be large enough for the principal to prefer to implement high effort from both types. In Appendix B, we show that when effort is observable, the optimal mechanism offers the payments $w^A_0 = 0$, $w^A_1 = \frac{3}{2}$, and $w^B_0 = w^B_1 = \frac{2}{3}$. ◊

In the example, the principal screens types by offering a contract with zero fixed payment and a high bonus to the type with the highest probability of success (A) and a higher fixed payment with no bonus to the type with the lowest success probability (B). Crucially, this mechanism is not feasible when effort is not observable, since both types would choose zero effort. In fact, by Theorem 1, the optimal mechanism must offer the same contract to all types in that case.

Next, we show that offering multiple contracts can be optimal when MS does not hold.

EXAMPLE 3 (General Distributions). There are $N \geq 3$ outputs and $N-1$ types: $\Theta = \{1, \ldots, N-1\}$. There are two efforts (0 and 1), and all types have the same effort costs: $c_0^\theta = 0$ and $c_1^\theta = (N-2)/(N-1)$ for all θ. The conditional probability of each output is

$$p_1^\theta(x) = \begin{cases} \dfrac{1}{2} & \text{if } x = x_{\theta+1} \\ \dfrac{1}{2(N-1)} & \text{if } x \neq x_{\theta+1} \end{cases} \quad \text{and} \quad p_0^\theta(x) = \begin{cases} \dfrac{1}{2} & \text{if } x = x_1 \\ \dfrac{1}{2(N-1)} & \text{if } x \neq x_1. \end{cases}$$

For each type, the lowest output (x_1) is the most likely outcome with low effort, whereas $x_{\theta+1}$ is the most likely outcome with high effort. Suppose that x_1 is low enough, so that it is optimal for the principal to implement high effort from all types.

In Appendix B, we show that the optimal mechanism offers the contracts

$$w^\theta(x) = \begin{cases} 2 & \text{if } x = x_{\theta+1} \\ 0 & \text{if } x \neq x_{\theta+1}. \end{cases}$$

Each type's contract pays 2 if output matches that agent's type and 0 otherwise. Therefore, the optimal mechanism consists of $N-1$ different contracts, one for each type. ◊

As the previous example illustrates, in general, the ranking of incentives is only a partial order. A contract that convinces one type to exert effort may be ineffective in incentivizing another type. In the example, the cheapest way to incentivize type θ to exert effort was to pay a bonus in case of output $x_{\theta+1}$. Multiplicative separability rules out cases like these, ensuring that all types order the power of different contracts in the same way.

2.2 *Robustness*

Our previous results rely on MS to characterize the optimal mechanism. Since MS is a strong assumption, it is natural to ask whether they hold for distributions that are close to MS. We now show that this is generically true, so that the single-contract result is not knife-edge in the sense that it is robust to perturbations to the distribution.

To avoid the complexity associated with genericity in infinite-dimensional spaces, we assume that the effort space E and the type space Θ are both finite. Consider a set of economies indexed by a vector of probabilities and costs:

$$\left\{p_{e,i}^\theta, c_e^\theta : e \in E,\, \theta \in \Theta,\, x_i \in X\right\}.$$

Since these are finite-dimensional vectors, we take any of the equivalent norms of Euclidean space. We say that a property is generic if it holds in an open and dense set.[14]

Proposition 2 shows that, generically, the results from Theorem 1 and Proposition 1 also hold if a distribution is "close" to multiplicatively separable.

[14]Since this subsection assumes that Θ and E are finite, economies are embedded in an Euclidean space with the topology derived from any of its equivalent norms.

PROPOSITION 2. *Let E and Θ be finite. For generic economies satisfying MS, there is a neighborhood around it for which the optimal mechanism is unique and offers a single contract to all types. Moreover, this contract pays zero in all but one outcome realization.*

Recall that when the distribution satisfies MS and there are no ties in the likelihood ratio, the optimal contract pays zero in all but one outcome realization. The proof of Proposition 2 establishes that if we apply a small perturbation to the output distribution or the agent's cost, the set of binding constraints in the principal's problem remains unchanged. Then it is optimal to pay zero in all but the same outcome realization as before the perturbation, although the amount paid in that outcome realization must be adjusted to incorporate the change in the incentive constraints.

Proposition 2 shows that, generically, the principal does not lose any profit by offering a single contract if the distribution is close to MS. The proof follows the logic of the maximum theorem. As we showed in Example 3, this result does not generalize for arbitrary distributions (far from MS).[15]

In the Supplemental Material, we study how much profit the principal loses by offering a single contract in economies with finitely many types and two efforts. We obtain an explicit lower bound on the principal's profit as a function of the distance between the distribution of outputs and a distribution that satisfies MS. Castro-Pires and Moreira (2021) generalize Proposition 2 to the case of risk-averse agents with a sufficiently convex cost of effort.

3. FREE DISPOSAL

We now introduce a *free disposal* (FD) constraint, which requires the principal's profit to be nondecreasing:

$$y - w^\theta(y) \geq x - w^\theta(x) \tag{FD}$$

for all $\theta \in \Theta$ and all $x, y \in X$ with $y \geq x$.[16] As Innes (1990) argues, FD can be seen as an additional incentive constraint if the principal can costlessly reduce output or if the agent can borrow from outside lenders to inflate output. An *optimal FD mechanism* maximizes the principal's expected profit (1) among mechanisms that satisfy IC, IR, LL and FD.

In this section, we consider the optimality of debt contracts. The principal gets a *debt contract* if his payments $x - w(x)$ equal $\min\{x, \bar{x}\}$ for some face value $\bar{x}$ or, equivalently, if the agent is given a call option $w(x) = \max\{x - \bar{x}, 0\}$. Incentive-compatible mechanisms cannot offer a menu of debt contracts, since agents would always pick the debt contract with the lowest face value. Therefore, for a mechanism to offer debt contracts only, it

[15]The genericity condition is also required for Proposition 2. As we show in the Supplemental Material, it is possible to construct (nongeneric) distributions arbitrarily close to MS in which offering a single contract is suboptimal.

[16]FD still allows the agent's payment $w^\theta(x)$ to be non-monotonic. The result from this section continues to hold if we also impose free disposal on the agent's side or limited liability on the principal, since debt contracts, which will be shown to be optimal, automatically satisfy these constraints.

needs to offer the same contract to all types. Accordingly, we assume that MS holds, so the principal offers a single contract.

The existing literature established that in the version of the model with one-dimensional effort ($E \subset \mathbb{R}$) with pure moral hazard, the monotone likelihood ratio property (MLRP) is sufficient for the optimality of debt. With one-dimensional efforts, a probability mass function p_e^θ satisfies MLRP if, for any e_L, e_H with $e_L < e_H$, the ratio $\frac{p_{e_H}^\theta(x)}{p_{e_L}^\theta(x)}$ is nondecreasing in x. Intuitively, MLRP means that the "evidence" in favor of higher efforts increases with output. MLRP plays an important role in the monotonicity of contracts (Holmstrom (1979), Grossman and Hart (1983)). Accordingly, we work with the following notion of MLRP.

DEFINITION 2. A distribution satisfying MS has the monotone likelihood ratio property if it can be written as in (2) and $\frac{p_{e_H}^\theta(x)}{p_{e_L}^\theta(x)}$ is nondecreasing in x for any e_L, e_H, and θ with $I(e_H, \theta) < I(e_L, \theta)$.

Our next result establishes that when the distributions satisfy MLRP, not only is it optimal to offer only one contract, but this contract takes the form of debt.

THEOREM 2. *Suppose MS holds and the distributions satisfy MLRP. There exists an optimal FD mechanism that gives the principal a single debt contract.*

The argument for the optimality of debt builds on Innes (1990) and Matthews (2001). With MLRP, higher outputs are more "indicative" of higher effort. Therefore, transferring payments from lower to higher outputs relaxes each type's moral hazard IC constraint. In general, one also has to take into account the ICs from adverse selection. However, because of multiplicative separability, the optimal mechanism offers a single contract so that the adverse selection ICs automatically hold.

Finally, it is also straightforward to adapt our proofs of Theorems 1 and 2 for the case of bilateral limited liability (i.e., both ex post principal and agent's payments are nonnegative). In this case, an optimal mechanism exists even with infinitely many outputs. Moreover, the optimal mechanism will generically offer a single live-or-die contract, in which either the principal or the agent's limited liability constraint will bind.

Bilateral free disposal

Next, we show that MS can be substantially weakened if one is willing to impose bilateral free disposal. For simplicity, we focus on the case of binary effort. Let $E = \{0, 1\}$ be the space of possible efforts and let the type space Θ be a compact topological space. Suppose there exists a continuous weak order $\succcurlyeq$ in Θ.[17]

Let $\Delta p^\theta = p_1^\theta - p_0^\theta$ denote type θ's *incremental distribution*, that is, the change in the probability of each output from exerting high rather than low effort. We assume that the

[17]Compactness can be substantially relaxed. All we need is that Θ is path connected, and that it has a minimum and a maximum value with respect to $\succcurlyeq$. As usual, $\succcurlyeq$ is a weak order if it is complete and transitive, and it is continuous if its upper and lower contour sets are closed.

incremental distributions are ordered according to the first-order stochastic dominance (FOSD), i.e., for all $(w(x_i))_{i=1}^N$ nondecreasing in x_i,

$$\theta_H \succcurlyeq \theta_L \quad \Longrightarrow \quad \sum_{i=1}^N \Delta p_i^{\theta_H} w(x_i) \geq \sum_{i=1}^N \Delta p_i^{\theta_L} w(x_i).$$

We also assume that incremental costs are nondecreasing,

$$\theta_H \succcurlyeq \theta_L \quad \Longrightarrow \quad \Delta c^{\theta_H} \leq \Delta c^{\theta_L},$$

where $\Delta c^\theta = c_1^\theta - c_0^\theta$. In particular, if $c_0^\theta = c_0 \leq 0$ and $p_0^\theta = p_0$ (both constant in θ), these conditions are equivalent to p_1^θ nondecreasing in terms of FOSD and c_1^θ nonincreasing in θ.

We say that a mechanism (w^θ) satisfies bilateral free disposal (BFD) if

$$0 \leq w^\theta(x_{i+1}) - w^\theta(x_i) \leq x_{i+1} - x_i, \quad i = 1, \ldots, N-1,$$

for all θ. BFD requires payments of both principal and agent to be nondecreasing. As Innes (1990) argues, BFD can be seen as an additional incentive constraint if the principal and the agent can costlessly reduce output or if they can borrow from outside lenders in order to inflate output. An *optimal BFD mechanism* maximizes the principal's expected profit (1) among mechanisms that satisfy IC, IR, LL, and BFD.

The proposition below shows that the optimal optimal BFD mechanism offers the same contract to all types.

PROPOSITION 3. *Let $E = \{0, 1\}$ and let Θ be a compact topological space with a continuous weak order $\succcurlyeq$. Suppose the incremental distributions satisfy FOSD and incremental costs are nondecreasing. There exists an essentially unique optimal BFD mechanism that offers a single contract to all types.*

As in Theorem 1, the optimal contract with BFD may not be a debt contract. However, as in Theorem 2, if the output distribution for each type also satisfies MLRP, the optimal contract takes the form of a debt contract.

EXAMPLE: PROCUREMENT AND REGULATION

We now describe how our framework can be adapted to a setting that builds on the classic model of Laffont and Tirole (1986, 1993).[18] Their model can be reduced to a pure adverse selection model because effort affects the regulated firm's cost deterministically. Our model incorporates moral hazard by allowing effort to affect the firm's cost stochastically.

A regulated firm produces an indivisible good that generates a consumer surplus of $S > 0$ at a random monetary cost $C \in \{C_1, \ldots, C_N\}$ with $C_1 < \cdots < C_N$. The firm's manager chooses a cost-reducing effort $e \in E$, which is not observed by the regulator.

[18] See the Supplemental Material for formal definitions and statements.

Let p_e^θ denote the probability that the firm's cost equals C conditional on type θ and effort e.

The manager is protected by limited liability and has cost of effort c_e^θ, with $\min_{e\in E} c_e^\theta \leq 0$. This is satisfied if the lowest effort costs zero or if the manager gets private benefits out of some activities. The firm's manager has private information about both his ability to cut costs (the conditional distribution of costs p_e^θ) and the cost of effort (c_e^θ). The regulator observes the monetary cost C incurred by the firm but not the manager's effort e. As an accounting convention, assume that the regulator reimburses the firm's monetary costs in addition to paying the firm an amount conditional on the observed cost C. A contract is a function that specifies a transfer to the firm conditional on each possible cost C.

Because the government uses distortionary taxation to raise public funds, the regulator faces a shadow cost of public funds $\lambda > 0$. We consider the goal of a utilitarian regulator who maximizes the sum of the expected utility of the firm's manager and the surplus of consumers. A contract is a function that specifies a transfer to the firm conditional on each possible cost C.

As in Theorem 1, we show in the Supplemental Material that under MS, there exists an essentially unique optimal mechanism that offers a single contract to all types. It may also be desirable to include a free disposal constraint (FD), requiring the firm's compensation for cutting costs not to exceed the amount cut. FD must hold, for example, if the firm's manager can secretly inflate firm costs. As in Theorem 2, we show in the Supplemental Material that under MS and MLRP, there exists an essentially unique optimal FD mechanism that offers all types the contract $w(C) = \max\{\bar{C} - C; 0\}$, for some $\bar{C}$. This optimal contract specifies a "reasonable cost" $\bar{C}$ and reimburses the regulated firm for any cost cuts beyond $\bar{C}$. As, by our accounting convention, the regulator pays the firm's cost directly, the firm's revenues equal

$$w(C) + C = \max\{\bar{C}, C\}.$$

This reimbursement rule consists of a standard price cap except that, because of limited liability, the regulator must bail out firms with cost realizations above the cap.[19] Price caps are the most common form of incentive regulation. For example, the U.S. Federal Communications Commission uses them to regulate the telephone industry. Price caps are often used in procurement as well. Prospective reimbursement systems commonly used in health care specify an amount $\bar{C}$ based on what a service should cost, and let providers keep cost savings $\bar{C} - C$ to themselves. They are used, for example, in Medicare.

4. Conclusion

The observation that many contracts are simple and relatively uniform across different sectors is an old puzzle in contract theory. While standard adverse selection models with

[19]If the realized cost is below the price cap $\bar{C}$, the firm is profitable and the reimbursement rule is the same as with a standard price cap. However, because the firm is protected by limited liability, the regulator cannot force it to remain active if its profits are negative. Therefore, the reimbursement rule above is a price cap with the additional feature that the regulator must bail out firms that incur losses.

the usual regularity conditions on preferences and technology predict that agents will be offered large menus of contracts, contracting parties typically offer a limited number of contracts, often a single one. While standard moral hazard models predict that contracts should be fine tuned to the likelihood ratio of output, similar contracts are offered in different environments.

We argue that these two features endogenously emerge in a general model of moral hazard and adverse selection if contracts must satisfy limited liability. With binary outcomes, the principal always offers a single contract regardless of any parameters of the model. The joint presence of moral hazard and adverse selection is key for this result. When either types or effort are observed, the principal typically prefers to offer different contracts to different types. With multiple outputs, it is also optimal to offer a single contract if the distribution of output satisfies a separability condition (which is a strong condition, but always satisfied with binary outcomes). Moreover, if the marginal distribution satisfies MLRP, this optimal FD contract is a debt contract for the principal.

Our paper shows that offering menus of contracts can hurt the principal. This is particularly stark with bilateral risk neutrality where, holding effort fixed, the agent always selects the most expensive contract to the principal. Then reducing the number of contracts offered to the agent always increases the principal's profits (for a fixed effort). If, in addition, we can identify a most efficient contract from the menu (such as when output is binary or when the distribution is multiplicatively separable), the principal can always improve by eliminating other contracts.

Our results are separate and complementary to the literature that studies contracts that are robust to assumptions about the informational environment. Insofar as environments in practice are approximately multiplicatively separable, our results imply that offering a single contract is optimal. Our results are also useful for applied researchers, who may wish to study environments with both adverse selection and moral hazard but may not want to introduce menus of contracts in their model. We show that to do so, it suffices to assume multiplicative separability, which always holds when there are two outputs or under the linearity of the distribution function condition.

The simplicity results rely on the presence of limited liability and bilateral risk neutrality. Limited liability ensures that increasing the power of a contract will not force the agent to abandon the mechanism. Bilateral risk neutrality means that, holding effort fixed, principal and agent are perfectly misaligned so that the principal always benefits by reducing the agent's flexibility. With risk aversion, their misalignment is no longer perfect because there are potential gains from risk-sharing. While risk neutrality is a reasonable assumption in many settings (such as the optimal compensation of wealthy managers, procurement contracts, or the regulation of large companies), there are many other settings where they are not (for example, insurance contracts or sharecropping). When the agent is sufficiently risk-averse and the technology is not close to MS, our results no longer hold. In our companion paper (Gottlieb and Moreira 2014), we study optimal mechanisms in the binary-outcome model when agents are risk-averse and when there are no limited liability constraints. Although we obtain some simplicity results, optimal mechanisms are considerably more complex than they are here.

Appendix A: Proofs

Before presenting the proofs of Theorems 1 and 2 and Proposition 1, it is helpful to introduce some notation. We will write $p^\theta_{e,i}$ for $p^\theta_e(x_i)$, h_i for $h(x_i)$ and w_i for $w(x_i)$. It will also be helpful to write variables in terms of increments: $\Delta x_i \equiv x_i - x_{i-1}$ and $\Delta w_i \equiv w_i - w_{i-1}$ for $i \geq 1$, where we set $x_0 \equiv 0$ and $w_0 \equiv 0$ so that Δx_1 and Δw_1 are well defined. With a slight abuse of notation, let $\Delta x \equiv x_N - x_1$ denote the difference between the highest and lowest outputs.

Proving Theorem 1

The following result will be important in order to establish existence of an optimal mechanism, by allowing us to restrict the set of possible contracts to a compact set.

Lemma 1. *Let (w, e) be a mechanism satisfying IC and LL. Suppose that $w^\theta_i > \frac{\Delta x}{\underline{p}^2}$ for some θ and i. Then (w, e) is not optimal.*

Proof. The proof has two steps. The first step shows that if an incentive-compatible mechanism offers a high enough payment in one outcome realization, then every other contract must also have a high enough payment in some outcome realization (otherwise, everyone would prefer the former contract). The second step shows that any contract that makes a sufficiently high payment in some outcome realization is dominated by the null contract.

Step 1. Suppose the mechanism offers a contract $\tilde{w} = (\tilde{w}_1, \ldots, \tilde{w}_N)$ with $\tilde{w}_i > \frac{\Delta x}{\underline{p}^2}$ for some output i. Let θ be a type that picks w^θ and exerts effort e. Then this type's incentive compatibility constraint gives

$$\sum_{i=1}^N p^\theta_{e,i} w_i - c^\theta_e \geq \sum_{i=1}^N p^\theta_{e,i} \tilde{w}_i - c^\theta_e.$$

Since $\max\{w^\theta_1, \ldots, w^\theta_N\} \geq \sum_{i=1}^N w_i p^\theta_{e,i}$ and $\tilde{w}_i \geq 0$ for all i, this inequality implies

$$\max_{i \in \{1,\ldots,N\}} \{w^\theta_i\} \geq \underline{p}\tilde{w}_i > \frac{\Delta x}{\underline{p}},$$

where the last inequality uses $\tilde{w}_i > \frac{\Delta x}{\underline{p}^2}$.

Step 2. We now show that this mechanism gives the principal a lower payoff than offering the contract that always pays zero to all types. Since the probability is bounded below by $\underline{p}$ and payments are nonnegative, the principal's payoff from offering $w^\theta = (w^\theta_1, \ldots, w^\theta_N)$ to type θ is

$$\sum_{i=1}^N p^\theta_{e(\theta),i}(x_i - w_i) \leq x_N - \underline{p} w_i \quad \forall i.$$

Let $e_0 \in \arg\max c_e^\theta$. The principal's payoff from offering type θ a zero payment in all outcome realizations is $\sum_{i=1}^N p_{e^0,i}^\theta x_i \geq x_1$. Combining these two inequalities, we obtain the following necessary condition for w to give a weakly higher payoff to the principal than the null contract:

$$w_i \leq \frac{\Delta x}{\underline{p}} \quad \forall i.$$

Thus, if

$$w_i > \frac{\Delta x}{\underline{p}}$$

for some i, then the principal is strictly better offering the null contract. □

By Lemma 1 and our Assumption 1, Assumptions 5.1–5.9 of Kadan, Reny, and Swinkels (2017) hold. Their Theorem 5.11 implies that there exists an optimal random mechanism.[20]

The proof will use the fact that any contract satisfying IC and LL also satisfies IR (because $\min_e c_e^\theta \leq 0$). The payoff of a type-θ agent who exerts effort e and gets contract w equals

$$v_e^\theta(w) := \sum_{i=1}^N p_{e,i}^\theta w_i - c_e^\theta. \tag{4}$$

The principal's payoff from such a type is

$$u_e^\theta(w) := \sum_{i=1}^N p_{e,i}^\theta (x_i - w_i). \tag{5}$$

By MS, a type-θ agent who switches from effort $\tilde{e}$ to e while keeping the same contract w gains

$$v_e^\theta(w) - v_{\tilde{e}}^\theta(w) = -\big[I(e,\theta) - I(\tilde{e},\theta)\big] \sum_{i=1}^N h_i w_i + c_{\tilde{e}}^\theta - c_e^\theta. \tag{6}$$

In turn, this switch changes the principal's payoff by

$$u_e^\theta(w) - u_{\tilde{e}}^\theta(w) = -\big[I(e,\theta) - I(\tilde{e},\theta)\big] \sum_{i=1}^N h_i (x_i - w_i). \tag{7}$$

If $I(e,\theta) \geq I(\tilde{e},\theta)$, the principal (weakly) gains from shifting effort from $\tilde{e}$ to e if and only if

$$\sum_{i=1}^N h_i (x_i - w_i) \leq 0. \tag{8}$$

[20] In the Supplemental Material, we extend the proof that follows to random mechanisms. Therefore, the restriction to deterministic mechanisms is without loss of generality.

Similarly, the agent's expected payment (weakly) increases from this shift in efforts if and only if $\sum_{i=1}^{N} h_i w_i \leq 0$.

The proof shows that the principal can (weakly) profit by removing all but one contract from any feasible menu of contracts.

PROOF OF THEOREM 1. Let (w, e) be a feasible mechanism. Let $\mathcal{M} := \{w^\theta : \theta \in \Theta\} \subset \mathbb{R}^N$ denote the set of all contracts in this mechanism. By Lemma 1, there is no loss of generality in assuming that $\mathcal{M}$ is bounded. Its closure, $\bar{\mathcal{M}}$, is compact in $\mathbb{R}^N$. There are three cases to consider.

Case 1: $\sum_{i=1}^{N} h_i(x_i - w_i^\theta) \geq 0$ *for all* $\theta \in \Theta$. Let $w^+ \in \arg\max_{w \in \bar{\mathcal{M}}} \sum_{i=1}^{N} h_i w_i$, which exists because $\bar{\mathcal{M}}$ is compact and the objective function is a continuous linear functional. Let $e^+(\theta)$ be an effort that maximizes the agent's payoff under contract w^+. Then the agent's payoff with the effort chosen in the original mechanism, e, cannot exceed the agent's payoff with effort $e^+(\theta)$,

$$v^\theta_{e^+(\theta)}\big(w^+\big) \geq v^\theta_{e(\theta)}\big(w^+\big),$$

which, by MS, can be written as

$$\big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i w_i^+ \geq c^\theta_{e^+(\theta)} - c^\theta_{e(\theta)}, \tag{9}$$

with strict inequality in case $e(\theta)$ is a suboptimal effort choice for type θ under contract w^+. Similarly, because $e(\theta)$ is the agent's effort choice with contract w^θ,

$$\big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i w_i^\theta \leq c^\theta_{e^+(\theta)} - c^\theta_{e(\theta)}. \tag{10}$$

Combining (9) and (10), we obtain

$$\begin{aligned}\big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i w_i^\theta &\leq c^\theta_{e^+(\theta)} - c^\theta_{e(\theta)} \\ &\leq \big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i w_i^+, \end{aligned} \tag{11}$$

where the last inequality is strict in case $e(\theta)$ is a suboptimal effort choice for type θ under contract w^+. By the definition of w^+,

$$\sum_{i=1}^{N} h_i w_i^+ \geq \sum_{i=1}^{N} h_i w_i^\theta.$$

Therefore, if $\sum_{i=1}^{N} h_i w_i^+ > \sum_{i=1}^{N} h_i w_i^\theta$, it follows from (11) that $I(e(\theta), \theta) \geq I(e^+(\theta), \theta)$. If $\sum_{i=1}^{N} h_i w_i^+ = \sum_{i=1}^{N} h_i w_i^\theta$, then $e(\theta)$ cannot be a suboptimal effort choice for type θ under contract w^+ since in this case the second inequality being strict would lead to contradiction. Therefore, we can take $e^+(\theta) = e(\theta)$, which, again, gives $I(e(\theta), \theta) \geq I(e^+(\theta), \theta)$.

We now establish that replacing contract w^θ by w^+ increases the principal's payoff from type θ. We first show that, holding effort fixed, the principal is better off with the substitution of contracts. Since w^+ is the limit of sequence in $\mathcal{M}$, the agent's utility is continuous, and the original mechanism is incentive compatible, it follows that

$$v^\theta_{e(\theta)}\big(w^\theta\big) \geq v^\theta_{e(\theta)}\big(w^+\big). \tag{12}$$

Substitute the expression for the agent's payoff, multiply both sides by -1, and add $\sum_{i=1}^N p^\theta_{e(\theta),i} x_i$ to both sides to write

$$\sum_{i=1}^N p^\theta_{e(\theta),i}\big(x_i - w_i^+\big) \geq \sum_{i=1}^N p^\theta_{e(\theta),i}\big(x_i - w_i^\theta\big), \tag{13}$$

which states that, holding effort $e(\theta)$ fixed, the principal gets a higher profit with contract w^+ than with w^θ.

To show that the change in effort also benefits the principal, notice that since $I(e(\theta), \theta) \geq I(e^+(\theta), \theta)$ and $w^+ \in \bar{\mathcal{M}}$ (so that $\sum_{i=1}^N h_i(x_i - w_i^+) \geq 0$), the following inequality holds:

$$\big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^N h_i\big(x_i - w_i^+\big) \geq 0.$$

Use MS to rewrite this inequality as

$$\sum_{i=1}^N p^\theta_{e^+(\theta),i}\big(x_i - w_i^+\big) \geq \sum_{i=1}^N p^\theta_{e(\theta),i}\big(x_i - w_i^+\big), \tag{14}$$

which states that the principal gains from the change in effort.

Combining (13) and (14) establishes that the principal's profit from θ with the new contract exceeds her profit with the original contract:

$$u^\theta_{e^+(\theta)}\big(w^+\big) \geq u^\theta_{e(\theta)}\big(w^\theta\big).$$

By construction, the mechanism (w^+, e^+) is incentive compatible and satisfies LL. Moreover, from the previous argument, it raises the principal's payoff pointwise (i.e., it raises the principal's payoff conditional on each type).

Case 2: $\sum_{i=1}^N (x_i - w_i^\theta) h_i \leq 0$ *for all* $\theta \in \Theta$. The proof of Case 2 is similar to the one of Case 1, except that, instead of substituting all contracts by the one that maximizes $\sum_{i=1}^N h_i w_i$, we substitute them by the one that minimizes this expression. Formally, let $w^- \in \arg\min_{w \in \bar{\mathcal{M}}} \sum_{i=1}^N h_i w_i$ and let $e^-(\theta)$ be an effort that maximizes the agent's payoff under contract w^-.

As in Case 1, incentive compatibility gives

$$\big[I\big(e(\theta), \theta\big) - I\big(e^-(\theta), \theta\big)\big] \sum_{i=1}^N h_i w_i^\theta \leq c^\theta_{e^-(\theta)} - c^\theta_{e(\theta)}$$

$$\leq \big[I\big(e(\theta),\theta\big) - I\big(e^-(\theta),\theta\big)\big]\sum_{i=1}^{N} h_i w_i^-. \tag{15}$$

Since $w^- \in \bar{\mathcal{M}}$, it satisfies

$$\sum_{i=1}^{N} h_i w_i^- \leq \sum_{i=1}^{N} h_i w_i^\theta,$$

so that, by inequality (15), it follows from the same argument as in Case 1 that $I(e^-(\theta),\theta) \geq I(e(\theta),\theta)$. As in Case 1, incentive compatibility implies that, holding effort $e(\theta)$ fixed, the principal's profit is higher with contract w^- than with w_θ:

$$\sum_{i=1}^{N} p^\theta_{e(\theta),i}\big(x_i - w_i^-\big) \geq \sum_{i=1}^{N} p^\theta_{e(\theta),i}\big(x_i - w_i^\theta\big). \tag{16}$$

Next, we show that the change in effort also benefits the principal. Because $I(e^-(\theta),\theta) \geq I(e(\theta),\theta)$ and because $w^- \in \bar{\mathcal{M}}$, the following inequality holds:

$$\big[I\big(e(\theta),\theta\big) - I\big(e^-(\theta),\theta\big)\big]\sum_{i=1}^{N} h_i\big(x_i - w_i^-\big) \geq 0.$$

Use MS to rewrite this inequality as

$$\sum_{i=1}^{N} p^\theta_{e^-(\theta),i}\big(x_i - w_i^-\big) \geq \sum_{i=1}^{N} p^\theta_{e(\theta),i}\big(x_i - w_i^-\big), \tag{17}$$

which shows that the principal gains from the change in effort. Combining (16) and (17) establishes that the principal's profit from θ with the new contract exceeds her profit with the original contract: $u^\theta_{e^-(\theta)}(w^-) \geq u^\theta_{e(\theta)}(w^\theta)$. Therefore, the mechanism (w^-, e^-) is incentive compatible, satisfies LL, and increases the principal's payoff pointwise relative to the original mechanism.

Case 3: There exist $\theta_+, \theta_- \in \Theta$ for which $\sum_{i=1}^N h_i(x_i - w_i^{\theta_+}) \geq 0 \geq \sum_{i=1}^N h_i(x_i - w_i^{\theta_-})$. First, we establish that, because of the risk neutrality and limited liability, introducing "scaled down versions" of the contracts from the original mechanism preserves incentive compatibility, meaning that no type would benefit from deviating to such a contract. More precisely, let

$$\mathcal{N} = \big\{\alpha w^\theta;\ \theta \in \Theta \text{ and } \alpha \in [0,1]\big\}$$

denote the menu of contracts obtained by introducing scaled down versions of all contracts in $\mathcal{M}$. Then, for all $\theta, \hat{\theta} \in \Theta$, $e \in E$, and $\alpha \in [0,1]$,

$$v^\theta_{e(\theta)}\big(w^\theta\big) \geq \sum_{i=1}^{N} p^\theta_{e,i} w_i^{\hat{\theta}} - c^\theta_e \geq \sum_{i=1}^{N} p^\theta_{e,i}\big(\alpha w_i^{\hat{\theta}}\big) - c^\theta_e,$$

where the first inequality follows from incentive compatibility of the original mechanism and the second inequality follows from $w_i^{\hat{\theta}} \geq \alpha w_i^{\hat{\theta}} \geq 0$ (by LL and the fact that $\alpha \leq 1$).

Therefore, there is no loss of generality in assuming that the principal offers the menu of contracts $\mathcal{N}$ rather than $\mathcal{M}$.

Let $w^0 \in \mathcal{N}$ be a contract that satisfies

$$\sum_{i=1}^{N} h_i\big(x_i - w_i^0\big) = 0. \tag{18}$$

We claim that w^0 exists. Indeed, suppose first that $\sum_{i=1}^N h_i x_i \geq 0$. Then, because $\sum_{i=1}^N h_i x_i \leq \sum_{i=1}^N h_i w_i^{\theta-}$, there exists $\alpha^0 \in [0, 1]$ such that

$$\sum_{i=1}^{N} h_i x_i = \alpha^0 \sum_{i=1}^{N} h_i w_i^{\theta-}.$$

Similarly, suppose that $\sum_{i=1}^N h_i x_i \leq 0$. Then, because $\sum_{i=1}^N h_i x_i \geq \sum_{i=1}^N h_i w_i^{\theta+}$, there exists $\alpha^0 \in [0, 1]$ such that

$$\sum_{i=1}^{N} h_i x_i = \alpha^0 \sum_{i=1}^{N} h_i w_i^{\theta+}.$$

As in Cases 1 and 2, incentive compatibility implies that, holding effort fixed, the principal's profit is higher with contract w^0 than with w^θ:

$$\sum_{i=1}^{N} p^{\theta}_{e(\theta),i}\big(x_i - w_i^0\big) \geq \sum_{i=1}^{N} p^{\theta}_{e(\theta),i}\big(x_i - w_i^\theta\big). \tag{19}$$

Let $e^0(\theta)$ be an effort that maximizes the type θ's payoff under contract w^0. We claim that changing efforts from $e(\theta)$ to $e^0(\theta)$ does not affect the principal's profit. To see this, multiply both sides of (18) by $I(e(\theta), \theta) - I(e^0(\theta), \theta)$ to write

$$\big[I\big(e(\theta), \theta\big) - I\big(e^0(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i\big(x_i - w_i^0\big) = 0.$$

Using MS, we can rewrite this equality as

$$\sum_{i=1}^{N} p^{\theta}_{e^0(\theta),i}\big(x_i - w_i^0\big) = \sum_{i=1}^{N} p^{\theta}_{e(\theta),i}\big(x_i - w_i^0\big), \tag{20}$$

which shows that the principal gets the same payoff with both effort profiles.

Combining (19) and (20) establishes that the mechanism (w^0, e^0) is incentive compatible, satisfies LL, and raises the principal's payoff pointwise relative to the original mechanism.

Next, we establish the essential uniqueness claim. Let (w, e) be any optimal mechanism and construct $(\bar{w}, \bar{e})$ as done previously. We need to show that for almost all types, the principal and the agent get the same payoffs in both mechanisms. As before, there are three possible cases: $\sum_{i=1}^N h_i[x_i - w_i^\theta] \geq 0$ for all θ; $\sum_{i=1}^N h_i[x_i - w_i^\theta] \leq 0$ for all $\theta \in \Theta$;

or $\sum_{i=1}^N h_i(x_i - w_i^{\theta_+}) \geq 0 \geq \sum_{i=1}^N h_i(x_i - w_i^{\theta_-})$ for some θ_+ and θ_-. We will consider the first case (the other two are analogous).

By construction, $u^\theta_{e^+(\theta)}(w^+) \geq u^\theta_{e(\theta)}(w^\theta)$ for all θ. By the optimality of (w, e), this inequality cannot hold on a set of types with positive measure. Therefore, the principal must be indifferent between these mechanisms except for a set of types with measure zero. Now consider the agent's payoff. Since w^+ is the limit of sequence in $\mathcal{M}$, we can proceed as in (12) to obtain

$$v^\theta_{e(\theta)}\big(w^\theta\big) \geq v^\theta_{e^+(\theta)}\big(w^+\big) \tag{21}$$

for all θ. Therefore, we must have

$$v^\theta_{e(\theta)}\big(w^\theta\big) \geq v^\theta_{e^+(\theta)}\big(w^+\big) \geq v^\theta_{e(\theta)}\big(w^+\big)$$

for all θ (where the last inequality follows from incentive compatibility of $(\tilde{w}, \tilde{e})$). Let $\tilde{\Theta}$ be the set of types for which

$$v^\theta_{e(\theta)}\big(w^\theta\big) > v^\theta_{e(\theta)}\big(w^+\big).$$

Using the expression for the agent's utility and rearranging, we obtain

$$u^\theta_{e(\theta)}\big(w^+\big) = \sum_{i=1}^N p^\theta_{e(\theta),i}\big(x_i - w_i^+\big) \geq \sum_{i=1}^N p^\theta_{e(\theta),i}\big(x_i - w_i^\theta\big) = u^\theta_{e(\theta)}\big(w^\theta\big),$$

with strict inequality exactly on $\tilde{\Theta}$. Use MS to write

$$u^\theta_{e^+(\theta)}\big(w^+\big) - u^\theta_{e(\theta)}\big(w^+\big) = \big[I\big(e(\theta), \theta\big) - I\big(e^+(\theta), \theta\big)\big] \sum_{i=1}^N h_i\big(x_i - w_i^+\big) \geq 0,$$

where, as in the previous part of the proof, the inequality follows from $I(e(\theta), \theta) \geq I(e^+(\theta), \theta)$ and $\sum_{i=1}^N h_i(x_i - w_i^+) \geq 0$. Combine these two inequalities to obtain $u^\theta_{e^+(\theta)}(w^+) \geq u^\theta_{e(\theta)}(w^\theta)$ for all θ, with strict inequality exactly on $\tilde{\Theta}$. Again, by the optimality of mechanism (w, e), $\tilde{\Theta}$ must have zero measure, which concludes the proof. □

Proving Proposition 1

When the principal offers a single contract, we can use MS to rewrite IC as

$$\big[I(e, \theta) - I\big(e(\theta), \theta\big)\big] \sum_{i=1}^N h_i w_i \geq c_{e(\theta)} - c_e \quad \forall e.$$

Notice that if it is optimal for the agent to pick effort $e(\theta)$ when he is offered contract w, it is also optimal to do so it when offered any contract $\tilde{w}$ with $\sum_{i=1}^N h_i w_i = \sum_{i=1}^N h_i \tilde{w}_i$. We are now ready to present the proof of the proposition.

PROOF OF PROPOSITION 1. Let w^* be an optimal contract, let $e(\theta)$ be the effort chosen by type θ when offered this contract, and let $\bar{p}_i \equiv \int_\Theta p^\theta_{e(\theta),i}\, d\mu(\theta)$ denote the (marginal) probability of outcome i. Then, for $K := \sum_{i=1}^N h_i w_i^*$, w^* must also solve the program

$$\min_w \sum_{i=1}^N \bar{p}_i w_i$$

subject to

$$\sum_{i=1}^N h_i w_i = K \tag{22}$$

$$w_i \geq 0, \quad \text{for all } i = 1, \ldots, N. \tag{LL}$$

As argued above, (22) ensures that effort $e(\theta)$ is still optimal for the agent. This is a restricted program: any contract that satisfies these constraints is feasible, but not every feasible contract satisfies these constraints. Since w^* is optimal among all feasible contracts, it must also solve this more restricted program that only includes a subset of feasible contracts.

The first-order conditions of this program are

$$-\bar{p}_j + \lambda h_j \leq 0 \tag{23}$$

for all j with $=$ if $w_j^* > 0$. There are three possible cases: (i) $\lambda > 0$, (ii) $\lambda < 0$, and (iii) $\lambda = 0$. First, suppose that $\lambda > 0$ and let $w_i^* > 0$ in some outcome i. Then, by (23),

$$\frac{h_j}{\bar{p}_j} \leq \frac{h_i}{\bar{p}_i} = \frac{1}{\lambda}$$

for all j, so that $w_j^* = 0$ whenever $j \notin \arg\max_{\tilde{j}} \{\frac{h_{\tilde{j}}}{\bar{p}_{\tilde{j}}}\}$. Next, suppose that $\lambda < 0$ and let $w_i^* > 0$ in some outcome i. Then, by (23), $w_j^* = 0$ whenever $j \notin \arg\min_{\tilde{j}} \{\frac{h_{\tilde{j}}}{\bar{p}_{\tilde{j}}}\}$. Last, notice that when $\lambda = 0$, condition (23) implies that $w_j^* = 0$ for all j (because $\bar{p}_j > 0$ for all j). In all three cases, the claim in the proposition is true. □

Proving Proposition 2

Let $\mathcal{P}$ denote the space of distributions satisfying $p^\theta_e(x_i) > \underline{p}$ for all e, θ, i, where $\underline{p} > 0$ is given in Assumption 2. Let $\#\Theta$ denote the number of elements in Θ and let $\Delta c^\theta_{e,\hat{e}} := c^\theta_e - c^\theta_{\hat{e}}$. Throughout the proof, we take any of the equivalent Euclidean norms for the (finite-dimensional) distribution $(p^\theta_{e,i})$, cost function (c^θ_e), and contract $w \in \mathbb{R}^{\#\Theta \times N}$.

Let $\Psi : \mathcal{P} \times E^{\#\Theta} \mapsto \mathbb{R}^{\#\Theta \times N}$ denote the feasibility correspondence,

$$\Psi(p, e) := \left\{ \begin{array}{c} \tilde{w} \in \mathbb{R}_+^{\#\Theta \times N};\ \forall \hat{e} \in E, \forall \theta, \hat{\theta} \in \Theta \\ \sum_{i=1}^N \left[p^\theta_{e(\theta),i} \tilde{w}_i^\theta - p^\theta_{\hat{e},i} \tilde{w}_i^{\hat{\theta}} \right] \geq \Delta c^\theta_{e(\theta),\hat{e}} \end{array} \right\},$$

that is, the set of incentive-compatible mechanisms under p. Let $\Gamma : \mathcal{P} \times E^{\#\Theta} \mapsto \times \mathbb{R}^{\#\Theta \times N}$ denote the policy correspondence of the principal's program,

$$\Gamma(p, e) = \underset{\tilde{w} \in \Psi(p,e)}{\arg\max} \sum_{\theta} \mu^{\theta} \sum_{i=1}^{N} p^{\theta}_{e(\theta),i}\big(x_i - \tilde{w}^{\theta}_i\big),$$

and $V : \mathcal{P} \times E^{\#\Theta} \to \mathbb{R}$ denote its optimal value,

$$V(p, e) = \max_{\tilde{w} \in \Psi(p,e)} \sum_{\theta} \mu^{\theta} \sum_{i=1}^{N} p^{\theta}_{e(\theta),i}\big(x_i - \tilde{w}^{\theta}_i\big).$$

LEMMA 2. *For each $e \in E^{\#\Theta}$, $V(\cdot, e)$ is a continuous function and $\Gamma(\cdot, e)$ is a upper semi-continuous correspondence at any p for which the interior of $\Psi(p, e)$ is nonempty.*

PROOF. As shown in the existence part of the proof of Theorem 1, we can assume, without loss of generality, that all feasible contracts belong to $[0, L]^{\#\Theta \times N}$ for some $L > 0$. Notice that $\Psi(\cdot, e)$ is a compact-valued correspondence. The proof proceeds by verifying the conditions for the maximum theorem (Berge (1963)). For this, we show that $\Psi(\cdot, e)$ is a continuous correspondence.

(a) Ψ is upper semi-continuous (u.s.c.): Let (p^n) be a sequence of distributions in $\mathcal{P}$ converging to p. For any $\tilde{w}_n \in \Psi(p_n, e)$, the finiteness of E and Θ, the compactness of $[0, L]^N$ (and passing to a convergent subsequence if necessary), we can suppose that $\tilde{w}_n$ converges to $\tilde{w} \in [0, L]^{\#\Theta \times N}$. By the continuity of the objective function and the constraints of the maximization problem that defines Γ, we have that $\tilde{w} \in \Psi(p, e)$. Therefore, Ψ is u.s.c.

(b) Ψ is lower semi-continuous (l.s.c.): Let (p^n) be a sequence of distributions in $\mathcal{P}$ that converges to p. Let $\tilde{w}$ be an interior point of $\Psi(p, e)$, i.e.,

$$\sum_{i=1}^{N} \big[p^{\theta}_{e(\theta),i}\tilde{w}^{\theta}_i - p^{\theta}_{\hat{e},i}\tilde{w}^{\hat{\theta}}_i\big] > \Delta c^{\theta}_{\tilde{e}(\theta),\hat{e}} \quad \text{for all } (\hat{\theta}, \hat{e}) \notin \big\{\big(\theta, e(\theta)\big);\, \theta \in \Theta\big\}.$$

Then, for n sufficiently large, we have that the previous inequality is also true for p^n instead of p. This implies that the constant sequence $\tilde{w} \in \Psi(p_n, e)$ converges to $\tilde{w} \in \Psi(p, e)$, which shows that Ψ is l.s.c. Let w be a frontier point of $\Psi(p, e)$. Since $\Psi(p, e)$ is a convex set with a nonempty interior, we can find a sequence (w_k) in the interior of $\Psi(p, e)$ converging to w. Now, for every n, we can then find k_n such that w_{k_n} belongs to the interior of $\Psi(p_n, e)$. Since (w_{k_n}) is a subsequence of (w_k), it also converges to w, establishing that Ψ is l.s.c.

Because the objective function of the maximization program in $V(p, e)$ is continuous and $\Psi(\cdot, e)$ is a continuous correspondence, it follows from the maximum theorem that $V(\cdot, e)$ is a continuous function and $\Gamma(\cdot, e)$ is u.s.c. □

In what follows we use the convention that $V(p, e) = -\infty$ when $\Psi(p, e) = \emptyset$.

COROLLARY 2. *Let $e_i \in E^{\#\Theta}$, $i = 1, 2$. If $V(p, e_1) > V(p, e_2)$ and $\Psi(p, e_1)$ has nonempty interior for some distribution $p \in \mathcal{P}$, then there exists a neighborhood $\mathcal{N}$ of p such that $V(\tilde{p}, e_1) > V(\tilde{p}, e_2)$ for all $\tilde{p} \in \mathcal{N}$.*

PROOF. To obtain a contradiction, let (p_n) be a sequence converging to p such that

$$V(p_n, e_2) \geq V(p_n, e_1)$$

for all $n \in \mathbb{N}$. Let $w_n \in \Psi(p_n, e_2)$ be a sequence that attains value $V(p_n, e_2)$. Passing to a convergent subsequence if necessary, let $w = \lim_n w_n$. Since Ψ is a compact-valued correspondence, $w \in \Psi(p, e_2)$. Hence, $V(p, e_2)$ is at least as high as the value attained at w. By Lemma 2 (and passing convergent subsequence if necessary), $\lim_{n\to\infty} V(p_n, e_1) = V(p, e_1)$ and, therefore,

$$V(p, e_2) \geq V(p, e_1),$$

which contradicts the hypothesis that $V(p, e_1) > V(p, e_2)$. □

LEMMA 3. *Let $\mathcal{P}_{MS}$ be the set of distributions in $\mathcal{P}$ that satisfy MS. Then the subset of $\mathcal{P}_{MS}$ for which the optimal contract is nonnull and pays a positive amount in one outcome realization only is generic (i.e., open and dense in $\mathcal{P}_{MS}$).*

PROOF. Fix $p \in \mathcal{P}_{MS}$ for which the optimal contract is nonnull. By Proposition 1, each solution of the principal–agent problem can be represented by a triple (i, w, e) defined by the outcome i at which the contract pays a positive amount $w \geq 0$ and by the recommended effort profile e. Denote $\mathcal{S}$ the set of all these feasible triples. Fix $(i^*, w^*, e^*) \in \mathcal{S}$ such that

$$w^* \in \arg\max\big\{w;\ (i, w, e) \in \mathcal{S} \text{ is a solution of the principal–agent problem}\big\}.$$

By our assumption, $w^* > 0$. Suppose without loss that $h_{i^*} > 0$ (the proof is analogous if $h_{i^*} < 0$). The proof is divided into several steps.

(a) Constructing the distribution for which all optimal contracts pay a positive amount at a unique outcome. Let $\epsilon > 0$ be sufficiently small and define the distribution $\tilde{p} \in \mathcal{P}$ as

$$\tilde{p}^{\theta}_{e^*(\theta),i} = \begin{cases} p^{\theta}_{e^*(\theta),i} - \epsilon & \text{if } i = i^* \\ p^{\theta}_{e^*(\theta),i} + \dfrac{\epsilon}{k} & \text{if } h_i \leq 0 \end{cases}$$

for each pair $(\theta, e^*(\theta))$, where $\theta \in \Theta$ and $k \geq 1$ is the cardinality of $\{i;\ h_i \leq 0\}$, which is not empty because $\sum_{i=1}^{N} h_i = 0$. For all $(\tilde{\theta}, \tilde{e}) \notin \{(\theta, e^*(\theta));\ \theta \in \Theta\}$, extend the definition of $\tilde{p}$ satisfying MS with the same functions h and I that define p.

(b) (i^*, w^*, e^*) is an optimal solution with distribution $\tilde{p}$. For any contract $w = (w_1, \ldots, w_N)$ and any effort recommendation profile $e(\theta)$, the incentive-compatibility constraint reads

$$\big[I(\tilde{e}, \theta) - I\big(e(\theta), \theta\big)\big] \sum_{i=1}^{N} h_i w_i \geq \Delta c^{\theta}_{e(\theta),\tilde{e}}$$

for all θ and $\tilde{e}$. Since h and I are unchanged, restricted to mechanisms with just one contract, incentive-compatibility constraints are equivalent under distributions p and $\tilde{p}$. The principal's payoff at (i^*, w^*, e^*) satisfies

$$\sum_\theta \mu^\theta \left[\sum_{i=1}^N p^\theta_{e^*(\theta),i} x_i - p^\theta_{e^*(\theta),i^*} w^* \right] \geq \sum_\theta \mu^\theta \left[\sum_{i=1}^N p^\theta_{\hat{e}(\theta),i} x_i - p^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} \right]$$

for all $(\hat{i}, \hat{w}, \hat{e}) \in \mathcal{S}$. Using MS, this last condition is equivalent to

$$\sum_\theta \mu^\theta \big[I\big(\hat{e}(\theta), \theta\big) - I\big(e^*(\theta), \theta\big) \big] \sum_{i=1}^N h_i x_i + \sum_\theta \mu^\theta \big[p^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - p^\theta_{e^*(\theta),i^*} w^* \big] \geq 0. \quad (24)$$

If $\hat{i} \neq i^*$, substituting $\tilde{p}$ for p, the inequality in (24) becomes strict, i.e.,

$$\sum_\theta \mu^\theta \big[I\big(\hat{e}(\theta), \theta\big) - I\big(e^*(\theta), \theta\big) \big] \sum_{i=1}^N h_i x_i + \sum_\theta \mu^\theta \big[\tilde{p}^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - \tilde{p}^\theta_{e^*(\theta),i^*} w^* \big] > 0,$$

because if $h_{\hat{i}} \leq 0$,

$$\begin{aligned}
&\tilde{p}^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - \tilde{p}^\theta_{e^*(\theta),i^*} w^* \\
&\quad = \big(\tilde{p}^\theta_{e^*(\theta),\hat{i}} + \big(I\big(e^*(\theta), \theta\big) - I\big(\hat{e}(\theta), \theta\big) \big) h_{\hat{i}} \big) \hat{w} - \tilde{p}^\theta_{e^*(\theta),i^*} w^* \\
&\quad = \left(p^\theta_{e^*(\theta),\hat{i}} + \frac{\epsilon}{k} + \big(I\big(e^*(\theta), \theta\big) - I\big(\hat{e}(\theta), \theta\big) \big) h_{\hat{i}} \right) \hat{w} - \big(p^\theta_{e^*(\theta),i^*} - \epsilon \big) w^* \\
&\quad > p^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - p^\theta_{e^*(\theta),i^*} w^*,
\end{aligned}$$

and if $h_{\hat{i}} > 0$,

$$\begin{aligned}
\tilde{p}^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - \tilde{p}^\theta_{e^*(\theta),i^*} w^* &= p^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - \big(p^\theta_{e^*(\theta),i^*} - \epsilon \big) w^* \\
&= p^\theta_{\hat{e}(\theta),\hat{i}} \hat{w} - p^\theta_{e^*(\theta),i^*} w^* + \epsilon w^* \\
&> p^\theta_{e^*(\theta),\hat{i}} \hat{w} - p^\theta_{e^*(\theta),i^*} w^*,
\end{aligned}$$

where we have used MS and the definition of $\tilde{p}$. If $\hat{i} = i^*$, substituting $\tilde{p}$ for p, the inequality (24) is equivalent to

$$\begin{aligned}
&\sum_\theta \mu^\theta \big[I\big(\hat{e}(\theta), \theta\big) - I\big(e^*(\theta), \theta\big) \big] \sum_{i=1}^N h_i x_i \\
&\quad + \sum_\theta \mu^\theta \big[p^\theta_{\hat{e}(\theta),i^*} \hat{w} - p^\theta_{e^*(\theta),i^*} w^* \big] + \epsilon \big(w^* - \hat{w} \big) \geq 0,
\end{aligned}$$

where we have used MS and the definitions of $\tilde{p}$ and w^*. Therefore, under $\tilde{p}$, (i^*, w^*, e^*) provides a strict higher payoff for the principal than $(\hat{i}, \hat{w}, \hat{e})$ for all $\hat{i} \neq i^*$.

(c) i^* is the unique outcome at which the optimal contract pays a positive amount with distribution $\tilde{p}$. From the proof of Proposition 1, i^* is characterized by

$$\sum_\theta \mu^\theta p^\theta_{e^*(\theta),i^*}\frac{h_i}{h_{i^*}} \le \sum_\theta \mu^\theta p^\theta_{e^*(\theta),i} \quad \text{for all } i \neq i^*.$$

Replacing p by $\tilde{p}$, these inequalities become

$$\begin{aligned}\sum_\theta \mu^\theta \tilde{p}^\theta_{e^*(\theta),i^*}\frac{h_i}{h_{i^*}} - \sum_\theta \mu^\theta \tilde{p}^\theta_{e^*(\theta),i} &\le \sum_\theta \mu^\theta p^\theta_{e^*(\theta),i^*}\frac{h_i}{h_{i^*}} - \sum_\theta \mu^\theta p^\theta_{e^*(\theta),i} - \frac{h_i}{h_{i^*}}\epsilon \\ &< 0 \end{aligned} \tag{25}$$

for all $i \neq i^*$. Therefore, i^* is the unique outcome where the optimal contract pays a positive amount at (i^*, w^*, e^*) with distribution $\tilde{p}$. Let $(i^*, \tilde{w}, \tilde{e})$ be another solution of the principal-agent problem with distribution $\tilde{p}$. Again from Proposition 1, i^* is the unique outcome this optimal contract pays a positive amount if and only if

$$\frac{h_i}{h_{i^*}}\sum_\theta \mu^\theta \tilde{p}^\theta_{\tilde{e}(\theta),i^*} < \sum_\theta \mu^\theta \tilde{p}^\theta_{\tilde{e}(\theta),i} \quad \text{for all } i \neq i^*.$$

Adding $\sum_\theta \mu^\theta (I(e^*(\theta), \theta) - I(\tilde{e}(\theta), \theta))$ on both sides and using MS, this previous inequality is equivalent to (25). Therefore, at any optimal solution with distribution $\tilde{p}$, i^* is the unique outcome realization where the optimal contract pays a positive amount.

(d) The set of distributions in $\mathcal{P}_{MS}$ for which any optimal contract pays a positive amount at a single outcome is generic. Indeed, from steps (a)–(c), for every neighborhood of a distribution in $\mathcal{P}_{MS}$, there exists a distribution in $\mathcal{P}_{MS}$ for which the uniqueness property holds. Moreover, it is straightforward to see that the set of distributions satisfying the uniqueness property is open in $\mathcal{P}_{MS}$. This step concludes the proof. □

LEMMA 4. *For a generic set of distributions in $\mathcal{P}_{MS}$ and a generic set of cost functions, there is a neighborhood in $\mathcal{P}$ for which the optimal mechanism is implemented by only one contract that pays a positive amount in only one outcome realization.*

PROOF. Let $p \in \mathcal{P}_{MS}$ and (i^*, w^*, e^*) be any solution of the principal–agent problem (we are using the notation in the proof of Lemma 3). Notice that incentive compatibility at this optimal solution is equivalent to

$$\big[I(\tilde{e}, \theta) - I\big(e^*(\theta), \theta\big)\big]h_{i^*}w^* \ge \Delta c^\theta_{e^*(\theta),\tilde{e}} \tag{26}$$

for all θ and $\tilde{e} \neq e^*(\theta)$. We claim that for a generic set of cost functions, the inequality (26) is strict. Indeed, let $\epsilon > 0$ be sufficiently small. Let us define the cost function $\tilde{c}^\theta_{\tilde{e}}$ exactly the same as $c^\theta_{\tilde{e}}$ except at $\tilde{e}$ and θ that bind constraint (26), where we define

$$\tilde{c}^\theta_{\tilde{e}} = c^\theta_{\tilde{e}} \begin{cases} +\epsilon & \text{if } \Delta c^\theta_{e^*(\theta),\tilde{e}} \ge 0 \\ -\epsilon & \text{if } \Delta c^\theta_{e^*(\theta),\tilde{e}} < 0. \end{cases}$$

For the model with cost function $\tilde{c}$ and distribution p, the constraint (26) is slack at contract w^* for all θ and $\tilde{e} \neq e^*(\theta)$.

There are two cases to be considered.

(i) The optimal contract is a null contract (i.e., $w^* = 0$) and there is no other nonnull contract that gives the same payoff to the principal. Since the interior of $\Psi(p, e^*)$ is nonempty, by the corollary of Lemma 2, there exists a neighborhood $\mathcal{N}$ of p such that $0 \in \Gamma(\tilde{p}, e^*)$ for all $\tilde{p} \in \mathcal{N}$. Moreover, we can take $\mathcal{N}$ such that no other nonnull contract can give higher payoff to the principal than the null contract in $\mathcal{N}$.

(ii) The optimal contract is nonnull, i.e., $w^* > 0$. By Lemma 3, we can generically assume that outcome i^* is the unique outcome at which the optimal contract pays a positive amount among all possible solutions. Let us assume without loss that $h_{i^*} > 0$ (as in the proof of that lemma). By the corollary of Lemma 2, there exists a neighborhood $\mathcal{N}$ of p such that e^* is the optimal effort profile and the optimal contract under $\tilde{p}$ is also nonnull for all $\tilde{p} \in \mathcal{N}$. Fixing $\tilde{p} \in \mathcal{N}$, define the relaxed version of the cost minimization program of implementing the effort recommendation profile $e^*(\theta)$:

$$\begin{aligned}
\min_{w^\theta} \quad & \sum_\theta \mu^\theta \sum_{i=1}^N \tilde{p}^\theta_{e^*(\theta),i} w_i^\theta \\
\text{s.t.} \quad & \sum_{i=1}^N [\tilde{p}^\theta_{e^*(\theta),i} - \tilde{p}^\theta_{\hat{e},i}] w_i^\theta - \Delta c^\theta_{e^*(\theta),\hat{e}} \geq 0 \quad \forall \theta, \hat{e} \\
& \sum_{i=1}^N \tilde{p}^{\hat{\theta}}_{e^*(\hat{\theta}),i} \big(w_i^{\hat{\theta}} - w_i^\theta\big) \geq 0 \quad \forall \hat{\theta} \neq \theta \\
& w^\theta \geq 0 \quad \forall \theta.
\end{aligned}$$

By the definition of i^*, there exists a mechanism formed by only one contract paying a positive amount at outcome i^*, which is feasible for the above program. We now show that the optimal mechanism that implements e^* is also in this class. Restricting to mechanisms with only one contract paying a positive amount at outcome i^*, let θ^* and $\tilde{e}^*$ be a type and an effort recommendation at which the first constraint of the above program binds at the optimal contract. Hence, we must have $\Delta c^\theta_{e^*(\theta),\tilde{e}^*} > 0$ and, consequently, $\tilde{p}^{\theta^*}_{e^*(\theta^*),i^*} - \tilde{p}^{\theta^*}_{\tilde{e}^*,i^*} > 0$, since otherwise the optimal mechanism would be the null one. Let $\tilde{\lambda}$, $\tilde{\sigma}$, and $\tilde{\gamma}$ be the Lagrangian multipliers of the first, second, and third constraints, respectively. The necessary and sufficient first-order conditions of the linear Lagrangian of the above program are

$$\mu^\theta \tilde{p}^\theta_{e^*(\theta)} - \sum_{\hat{e}} \tilde{\lambda}^{\theta,\hat{e}} [\tilde{p}^\theta_{e^*(\theta)} - \tilde{p}^\theta_{\hat{e}}] + \sum_{\hat{\theta}\neq\theta} \tilde{\sigma}^{\theta,\hat{\theta}} \tilde{p}^{\hat{\theta}}_{e^*(\hat{\theta})} - \Bigg(\sum_{\hat{\theta}\neq\theta} \tilde{\sigma}^{\hat{\theta},\theta}\Bigg) \tilde{p}^\theta_{e^*(\theta)} - \tilde{\gamma}^\theta = 0.$$

To conclude the proof, it is enough to construct multipliers for which the mechanism formed by one contract paying a positive amount only at outcome i^* is a

critical point of the Lagrangian and satisfy the complementary slackness conditions. For this, let us define the following continuous mapping on $\mathcal{N}$:

$$\tilde{\lambda}^{\theta,\hat{e}}(\tilde{p}) = \begin{cases} \dfrac{\sum_{\hat{\theta}} \mu^{\hat{\theta}} \tilde{p}^{\hat{\theta}}_{e^*(\hat{\theta}),i^*}}{\tilde{p}^{\theta}_{e^*(\theta),i^*} - \tilde{p}^{\theta}_{\hat{e},i^*}} & \text{if } \theta = \theta^* \text{ and } \hat{e} = \tilde{e}^* \\ 0 & \text{if otherwise} \end{cases}$$

and

$$\tilde{\sigma}^{\hat{\theta},\theta} = \begin{cases} \mu^{\theta} & \text{if } \hat{\theta} = \theta^* \text{ and } \theta \neq \theta^* \\ 0 & \text{if otherwise.} \end{cases}$$

For $\theta = \theta^*$, the first-order condition reads

$$\tilde{\gamma}^{\theta^*} = \sum_{\hat{\theta}} \mu^{\hat{\theta}} \tilde{p}^{\hat{\theta}}_{e^*(\hat{\theta})} - \sum_{\hat{e}} \tilde{\lambda}^{\theta^*,\hat{e}}(\tilde{p}) \big[\tilde{p}^{\theta^*}_{e^*(\theta^*)} - \tilde{p}^{\theta^*}_{\hat{e}}\big],$$

which defines $\tilde{\gamma}^{\theta^*}$. In particular, by the definition of $\tilde{\lambda}$, we have $\tilde{\gamma}^{\theta^*}_{i^*} = 0$. For $\theta \neq \theta^*$, the first-order condition reads

$$\tilde{\gamma}^{\theta} = \mu^{\theta} \tilde{p}^{\theta}_{e^*(\theta)} - \left(\sum_{\hat{\theta} \neq \theta} \tilde{\sigma}^{\hat{\theta},\theta}\right) \tilde{p}^{\theta}_{e^*(\theta)} = 0,$$

which completes the definition of $\tilde{\gamma}$, where the last equality is a consequence of the definition of $\tilde{\sigma}$. The only remaining slackness condition that must be checked is $\tilde{\gamma}^{\theta^*} \geq 0$. For this, define the mapping $T : \Delta^{\#\Theta \times \#E} \to \mathbb{R}^N$ by

$$T(\tilde{p}) = \sum_{\hat{\theta}} \mu^{\hat{\theta}} \tilde{p}^{\hat{\theta}}_{e^*(\hat{\theta})} - \sum_{\hat{e}} \tilde{\lambda}^{\theta^*,\hat{e}}(\tilde{p}) \big[\tilde{p}^{\theta^*}_{e^*(\theta^*)} - \tilde{p}^{\theta^*}_{\hat{e}}\big],$$

where Δ is the $(N-1)$-dimensional simplex. We have that T is a continuous mapping, $T(p)_{i^*} = 0$, and

$$T(p)_i = \sum_{\hat{\theta}} \mu^{\hat{\theta}} p^{\hat{\theta}}_{e^*(\hat{\theta}),i} - \sum_{\hat{\theta}} \mu^{\hat{\theta}} p^{\hat{\theta}}_{e^*(\hat{\theta}),i^*} \frac{h_i}{h_{i^*}} > 0$$

for all $i \neq i^*$. By the continuity of T, we can find an even smaller $\mathcal{N}$ such that $T(\tilde{p})_{i^*} = 0$ and $T(\tilde{p})_i > 0$ for all $i \neq i^*$ and $\tilde{p} \in \mathcal{N}$. Therefore, the effort profile $e^*(\theta)$ is implemented by a single contract paying a positive amount only at outcome i^*. □

Proving Theorem 2

It is straightforward to adapt the proof of Theorem 1 to show that it is still optimal to offer a single contract when free disposal is imposed. Therefore, there is no loss of generality in assuming that the optimal mechanism offers a single contract.

Notice that whenever MS holds, we can write

$$p^{\theta}_{e,i} + I(e,\theta)h_i = p^{\theta}_{\tilde{e},i} + I(\tilde{e},\theta)h_i \quad \forall e,\tilde{e},\theta,i.$$

Rearranging this expression, gives

$$\frac{p^{\theta}_{\tilde{e},i}}{p^{\theta}_{e,i}} - 1 = -\big[I(\tilde{e},\theta) - I(e,\theta)\big]\frac{h_i}{p^{\theta}_{e,i}}.$$

Thus, MLRP holds if and only if $\frac{h_i}{p^{\theta}_{e,i}}$ is nondecreasing in i for any e, θ.

Recall that, as shown in the proof of Proposition 1, if it is optimal for the agent to pick effort $e(\theta)$ when offered contract w, it is also optimal to do so when offered any contract $\tilde{w}$ with $\sum_{i=1}^{N} h_i w_i = \sum_{i=1}^{N} h_i \tilde{w}_i$. We are now ready to present the proof:

Proof of Theorem 2. Let w^* be an optimal contract and let $e(\theta)$ denote the effort chosen by type θ when offered this contract. Then, for $K = \sum_{i=1}^{N} h_i w_i^*$, this contract must solve the program

$$\min_{w} \sum_{i=1}^{N} w_i \int_{\Theta} p^{\theta}_{e(\theta),i}\, d\mu(\theta)$$

subject to

$$\sum_{i=1}^{N} h_i w_i = K \tag{IC′}$$

$$w_i \geq 0 \tag{LL}$$

$$x_i - x_{i-1} \geq w_i - w_{i-1}, \tag{M}$$

where we are using the convention $x_0 = w_0 = 0$. As argued above, the first constraint ensures that effort $e(\theta)$ is still optimal for the agent. The second and third constraints are LL and FD. This is a restricted program: any contract that satisfies these constraints is feasible, but not every feasible contract satisfies these constraints. Therefore, since w^* is optimal among all feasible contracts, it must also solve this more restricted program that only includes a subset of feasible contracts.

Let $\bar{p}_i \equiv \int_{\Theta} p^{\theta}_{e(\theta),i}\, d\mu(\theta)$ denote the marginal distribution of outputs induced by effort $e(\cdot)$. It is convenient to rewrite the program above in terms of increments:

$$\min_{\{\Delta w_i\}} \sum_{j=1}^{N} \overline{p}_j \sum_{i=1}^{j} \Delta w_i \tag{27}$$

subject to

$$\sum_{j=1}^{N} h_j \sum_{i=1}^{j} \Delta w_i = K \tag{IC′}$$

$$\sum_{i=1}^{j} \Delta w_i \geq 0 \quad \forall j \tag{LL$_j$}$$

$$\Delta x_j - \Delta w_j \geq 0 \quad \forall j. \tag{M$_j$}$$

We claim that if (LL$_j$) holds with equality, then (M$_j$) holds with strict inequality and vice versa. To see this, notice that if (LL$_j$) holds with equality, then

$$\sum_{i=1}^{j} \Delta w_i = \underbrace{\sum_{i=1}^{j-1} \Delta w_i}_{\geq 0 \text{ by } M_{j-1}} + \Delta w_j = 0 \therefore \Delta w_j \leq 0 < \Delta x_j,$$

showing that (M$_j$) holds with inequality. Conversely, if (M$_j$) holds with equality, then

$$\Delta x_j = \Delta w_j \therefore \Delta w_j + \underbrace{\sum_{i=1}^{j-1} \Delta w_i}_{\geq 0 \text{ by } LL_{j-1}} \geq \Delta x_j > 0,$$

limit so (LL$_j$) holds with inequality.

The necessary first-order conditions associated with program (27) are

$$-\sum_{j=i}^{N} \bar{p}_j + \lambda^{IC} \sum_{j=i}^{N} h_j + \sum_{j=i}^{N} \mu_j^{LL} - \mu_i^{M} = 0 \quad \forall i \tag{28}$$

along with the usual complementary slackness conditions.

Let $\xi_j \equiv \bar{p}_j - \lambda^{IC} h_j$ and notice that

$$\xi_j > 0 \quad \Longleftrightarrow \quad \frac{1}{\lambda^{IC}} > \frac{h_j}{\bar{p}_j}.$$

Notice that MLRP implies that $\frac{h_j}{p^{\theta}_{e(\theta),j}}$ is nondecreasing in j for all θ, so that $\frac{h_j}{\int_\Theta p^{\theta}_{e(\theta),j}\, d\mu(\theta)} = \frac{h_j}{\bar{p}_j}$ is also nondecreasing in j. Hence, there exists $k \in \{1, \ldots, N\}$ such that $\xi_j \geq (\leq) 0$ for all $j \leq (\geq) k$. Define $\underline{k}$ and $\bar{k}$ as the lowest and highest values that satisfy this property.

Substituting ξ_i in (28), we obtain

$$\begin{aligned} \xi_i &= \mu_i^{LL} - \mu_i^{M} + \mu_{i+1}^{M}, \quad i < N \\ \xi_N &= \mu_N^{LL} - \mu_N^{M}. \end{aligned}$$

There are three cases: $\xi_N > 0$, $\xi_N < 0$, and $\xi_N = 0$.

Case 1: $\xi_N > 0$. In this case, (LL$_N$) binds and (M$_N$) does not. Because ξ_i crosses 0 from above, $\xi_i > 0$ for all i and, by induction, none of the free disposal constraints binds. As a result, the solution is $w_i = 0$ for all i.

Case 2: $\xi_N < 0$. In this case, (M_N) binds. For $N-1$, we have

$$\xi_{N-1} = \mu^{LL}_{N-1} - \mu^{M}_{N-1} - \xi_N \geq \mu^{LL}_{N-1} - \mu^{M}_{N-1}.$$

If $N-1 \geq \bar{k}$ so that $\xi_{N-1} \leq 0$, it then follows that $\mu^{LL}_{N-1} - \mu^{M}_{N-1} \leq 0$ so that (M_{N-1}) binds (and (LL_{N-1}) does not). Inductively, it follows that (M_i) binds for all $i \geq k$.

Let j be such that (LL_j) binds (and, therefore $\mu^{LL}_j = 0$). By the previous argument, it must be the case that $j < \underline{k}$ so that $\xi_j > 0$. We have that

$$0 < \xi_{j-1} = \mu^{LL}_{j-1} - \mu^{M}_{j-1} + \mu^{M}_{j} = \mu^{LL}_{j-1} - \mu^{M}_{j-1}.$$

Then we must have that (LL_{j-1}) binds and (M_{j-1}) does not. Hence, there exists i^* between $\underline{k}$ and $\bar{k}$ such that (LL_i) binds on all $i \leq i^*$ and (M_i) binds on all $i > i^*$. The contract must, therefore, be an option with a strike price $x^* \in \{x_{i^*}, x_{i^*+1}\}$.

Case 3: $\xi_N = 0$. In this case, $\mu^{LL}_N = \mu^{M}_N$. Since we have previously shown that we cannot have both constraints holding with equality, we must have $\mu^{LL}_N = \mu^{M}_N = 0$. Substituting at the condition for the previous output gives

$$\xi_{N-1} = \mu^{LL}_{N-1} - \mu^{M}_{N-1} > 0,$$

where we used the fact that ξ_i crosses 0 from above and $\mu^{M}_N = 0$. Thus, $\mu^{LL}_{N-1} > 0 = \mu^{M}_{N-1}$ so that (LL_{N-1}) binds, i.e., $N-1 \leq \underline{k}$. Substituting inductively shows that all other LL constraints also bind. Hence, the solution in this case is an option contract with a strike price $x^* \in \{x_{N-1}, x_N\}$.

The last part of the proposition follows an argument similar to the proof of Theorem 1. □

Remark 1. Notice that Theorems 1 and 2 still hold with a continuum of outputs if we impose that the space of feasible contracts is uniformly bounded. The maximization and minimization problems defined in the proof of Theorem 1 have solutions if we consider the $(L^\infty(X), L^1(X))$ weak* topology on the space of feasible contracts. Indeed, by the Banach–Alaoglu theorem (see Rudin (1991)), the set $\mathcal{M}$ defined for those problems has a weak*-compact closure, and their objective functions are continuous with respect to weak* topology. The rest of the proof is a simple adaptation of the arguments from Theorems 1 and 2.

Proof of Proposition 3

Let $(w^\theta)_{\theta\in\Theta}$ be an incentive-compatible mechanism satisfying LL and BFD. Let $\Theta_e \subset \Theta$ denote the set of types who are recommended effort $e \in E$ in this mechanism. By incentive compatibility,

$$\sum_{i=1}^{N} \Delta p^\theta_i w^\theta_i \geq \Delta c^\theta$$

for all $\theta \in \Theta_1$. There are two cases: (i) $\Theta_1 \neq \emptyset$ and (ii) $\Theta_1 = \emptyset$.

Case (i). By the compactness of Θ and the ordering condition, Θ_1 has an infimum θ^*. In other words, there exist θ^* such that $\theta \succcurlyeq \theta^*$ for all $\theta \in \Theta_1$, and for every neighborhood N of θ^* in Θ, there exists $\tilde{\theta} \in N$ such that $\theta^* \succcurlyeq \tilde{\theta}$ and it is not the case that $\tilde{\theta} \succcurlyeq \theta^*$.

Suppose first that $\theta^* \in \Theta_1$ and consider the mechanism consisting of the single contract $w^* = w^{\theta^*}$. We claim that the original mechanism cannot be optimal if IC does not bind. Otherwise, there exists $\alpha \in (0, 1)$ such that αw^* still satisfies the IC for type θ^*. We claim that the mechanism that substitutes all contracts for the single contract αw^* also satisfies IC (with the same effort recommendation as in the original mechanism). Indeed, types in Θ_0 do not deviate from this recommendation because they now have even lower incentive to exert effort. For $\theta \in \Theta_1$, our ordering condition yields $\theta \succcurlyeq \theta^*$, so that

$$\sum_{i=1}^{N} \Delta p_i^{\theta} \alpha w_i^* \geq \sum_{i=1}^{N} \Delta p_i^{\theta^*} \alpha w_i^* \geq \Delta c^{\theta^*} \geq \Delta c^{\theta},$$

where the first and the last inequalities result from our assumption on the incremental distributions and costs. This new mechanism implements the same effort as in the original mechanism at a strictly lower cost and, therefore, the original mechanism cannot be optimal. Hence, without loss of generality we can assume that

$$\sum_{i=1}^{N} \Delta p_i^{\theta^*} w_i^* = \Delta c^{\theta^*}.$$

By our assumptions on the incremental distributions and costs, and the fact that w^* is nondecreasing, we have

$$\sum_{i=1}^{N} \Delta p_i^{\theta} w_i^* \geq \Delta c^{\theta}$$

if and only if $\theta \succcurlyeq \theta^*$. If we denote such set by Θ_1^*, then $\Theta_1 \subset \Theta_1^*$, i.e., the new mechanism still recommends effort to all types in Θ_1 and eventually to some new ones. Moreover, if $\theta^* \succ \theta$, then

$$\sum_{i=1}^{N} \Delta p_i^{\theta} w_i^* < \sum_{i=1}^{N} \Delta p_i^{\theta^*} w_i^* = \Delta c^{\theta^*} < \Delta c^{\theta},$$

i.e., $\Theta_0^* = \Theta \backslash \Theta_1^* = \{\theta;\, \theta^* \succ \theta\} \subset \Theta_0$. Finally, the incentive-compatibility constraint gives

$$\sum_{i=1}^{N} p_{e,i}^{\theta} w_i^* \leq \sum_{i=1}^{N} p_{e,i}^{\theta} w_i^{\theta}$$

for all $\theta \in \Theta_e$ and for each $e \in \{0, 1\}$, i.e., the cost of the new mechanism reduces type-by-type.

Now we will compare the expected profit between the original and the new mechanisms in three possible regions of the type space:

(a) $\theta \in \Theta_0^*$: Since $\sum_{i=1}^N p_{0,i}^\theta w_i^* \leq \sum_{i=1}^N p_{0,i}^\theta w_i^\theta$, we have

$$\int_{\Theta_0^*} \sum_{i=1}^N \big(x_i - w_i^\theta\big) p_{0,i}^\theta \, d\mu(\theta) \leq \int_{\Theta_0^*} \sum_{i=1}^N \big(x_i - w_i^*\big) p_{0,i}^\theta \, d\mu(\theta).$$

(b) $\theta \in \Theta_1^* \backslash \Theta_1 \subset \Theta_0$: Since x and $x - w^*$ are nondecreasing, $\sum_{i=1}^N \Delta p_i^\theta x_i \geq 0$ and $\sum_{i=1}^N \Delta p_i^\theta (x_i - w_i^*) \geq 0$, and since $\sum_{i=1}^N p_{0,i}^\theta w_i^* \leq \sum_{i=1}^N p_{0,i}^\theta w_i^\theta$, we have

$$\int_{\Theta_1^* \backslash \Theta_1} \sum_{i=1}^N \big(x_i - w_i^\theta\big) p_{0,i}^\theta \, d\mu(\theta) \leq \int_{\Theta_1^* \backslash \Theta_1} \sum_{i=1}^N \big(x_i - w_i^*\big) p_{1,i}^\theta \, d\mu(\theta).$$

(c) $\theta \in \Theta_1$: Since $\sum_{i=1}^N p_{1,i}^\theta w_i^* \leq \sum_{i=1}^N p_{1,i}^\theta w_i^\theta$, we have

$$\int_{\Theta_1} \sum_{i=1}^N \big(x_i - w_i^\theta\big) p_{1,i}^\theta \, d\mu(\theta) \leq \int_{\Theta_1} \sum_{i=1}^N \big(x_i - w_i^*\big) p_{1,i}^\theta \, d\mu(\theta).$$

Therefore, the expected profit of the new mechanism weakly increases.

If $\theta^* \notin \Theta_1$, then we can find a sequence (θ_n) in Θ_1 that converges to θ^*. Since (w^{θ_n}) is a bounded sequence, there exists a subsequence that converges to some w^* satisfying LL and BFD. It is easy to adapt the proof above to this case.

In case (ii), the compactness of Θ and the properties of the order imply that there exists

$$\theta^* \in \min\{\Theta\},$$

where the minimum is with respect to $\succcurlyeq$, i.e.,

$$\eta \geq \theta^*$$

for all $\eta \in \Theta$. Consider the mechanism formed by the single contract $w^* = w^{\theta^*}$. By incentive compatibility, $p_0^\theta . w^* \leq p_0^\theta . w^\theta$ for all $\theta \in \Theta$ and, hence,

$$\int_{\Theta} \sum_{i=1}^N \big(x_i - w_i^\theta\big) p_{0,i}^\theta \, d\mu(\theta) \leq \int_{\Theta} \sum_{i=1}^N \big(x_i - w_i^*\big) p_{0,i}^\theta \, d\mu(\theta).$$

Therefore, again the expected profit of the new mechanism weakly increases.

Appendix B: Examples

Screening with pure adverse selection (Example 1 continuation)

Let w_j^i denote type i's payment in outcome j. The optimal mechanism must solve the program

$$\min_{w_H^i, w_L^i \geq 0} \quad 2w_H^A + w_L^A + w_H^B + 2w_L^B$$

$$\begin{aligned}\text{subject to}\quad & 2w_H^A + w_L^A \geq 2w_H^B + w_L^B\\ & w_H^B + 2w_L^B \geq w_H^A + 2w_L^A\\ & 2w_H^A + w_L^A \geq 3\\ & w_H^B + 2w_L^B \geq 2.\end{aligned}$$

The two first constraints require A and B to prefer to report their types truthfully (IC constraints). Since effort is observable, the principal does not need to worry about deviations on effort. Then it is no longer the case that LL implies IR. The last two constraints are precisely the IR constraints. Disregarding the IC constraints, the IR constraints must bind for both types at optimal mechanism, which gives the one described in the text. It is straightforward to check that the IC constraints are satisfied for this mechanism.

Example 3 (continuation)

We will calculate the cheapest way of implementing high effort from all types. Consider the relaxed program

$$\min_{w_i^j \geq 0} \sum_{i=1}^{N-1} \frac{\mu_i}{2}\left(w_{i+1}^i + \frac{\sum_{j\neq i+1} w_j^i}{N-1}\right)$$

subject to

$$\frac{1}{2}\left(w_{i+1}^i + \frac{\sum_{j\neq i+1} w_j^i}{N-1}\right) - \frac{N-2}{N-1} \geq \frac{1}{2}\left(w_1^i + \frac{\sum_{j\neq 1} w_j^i}{N-1}\right)$$

for $i = 1, \ldots, N-1$. This is a relaxed program in that it only considers the "pure moral hazard" ICs. The unique solution is

$$w_j^i = \begin{cases} 2 & \text{if } j = i+1\\ 0 & \text{if } j \neq i+1.\end{cases}$$

The relaxed program omitted the ICs that require that each type does not benefit by picking a different contract while exerting high effort. However, those constraints are satisfied at the solution obtained above, since

$$0 = \frac{1}{2}\cdot 2 - \frac{N-2}{N-1} > \frac{1}{2(N-1)}\cdot 2 - \frac{N-2}{N-1}.$$

Thus, all omitted ICs are satisfied.

Sufficient conditions for it to be optimal to recommend high effort to all types are

$$x_{i+1} - 2 + \frac{1}{N-1}\sum_{j\neq i+1} x_j > x_1 + \frac{1}{N-1}\sum_{j\neq 1} x_j$$

or

$$x_{i+1} - 2\frac{N-1}{N-2} > x_1$$

for all $i = 1, \ldots, N-1$.

Appendix C: Multiplicative separability

This appendix examines the multiplicative separability condition (MS). We provide two results. The first is a characterization of MS in terms of the ordering of incentives. The second one is the graphical interpretation given in the text regarding how distributions satisfying MS change with effort.

Recall the condition presented in the text, which allows contracts to be ranked by their incentives:

Definition 3. A distribution of outputs is *ordered in terms of incentives* if, for given w and $\tilde{w}$ satisfying LL, if there exist $\theta_0 \in \Theta$, $e_0, \tilde{e}_0 \in E$, with $p^{\theta_0}_{e_0} \neq p^{\theta_0}_{\tilde{e}_0}$,

$$\sum_{i=1}^{N} (p^{\theta_0}_{e_0,i} - p^{\theta_0}_{\tilde{e}_0,i}) w_i = \sum_{i=1}^{N} (p^{\theta_0}_{e_0,i} - p^{\theta_0}_{\tilde{e}_0,i}) \tilde{w}_i,$$

then

$$\sum_{i=1}^{N} (p^{\theta}_{e,i} - p^{\theta}_{\tilde{e},i}) w_i = \sum_{i=1}^{N} (p^{\theta}_{e,i} - p^{\theta}_{\tilde{e},i}) \tilde{w}_i$$

for all $\theta \in \Theta$ and $e, \tilde{e} \in E$.

Substitution shows that any distribution that satisfies MS must be ordered in terms of incentives. The next proposition shows that the reverse is also true, so these are equivalent conditions.

Proposition 4. *The following statements are equivalent:*

(i) The distribution of outputs $p^{\theta}_{e}(x)$ is ordered in terms of incentives.

(ii) The distribution of outputs $p^{\theta}_{e}(x)$ satisfies MS.

(iii) For all i, j, there exist constants $\phi_{i,j}$ (that depend on the outputs x_i and x_j, but not on type or effort) such that

$$\frac{p^{\theta}_{e,i} - p^{\theta}_{\tilde{e},i}}{p^{\theta}_{e,j} - p^{\theta}_{\tilde{e},j}} = \phi_{i,j} \tag{29}$$

whenever $p^{\theta}_{e,j} - p^{\theta}_{\tilde{e},j} \neq 0$.

PROOF. We first show that (i) implies (ii). Define the linear functional $\varphi^{\theta,e,\tilde{e}}:\mathbb{R}^N\to\mathbb{R}$ as

$$\varphi^{\theta,e,\tilde{e}}(w)\equiv\sum_{i=1}^{N}\big(p^{\theta}_{e,i}-p^{\theta}_{\tilde{e},i}\big)w_i.$$

Since $w(x)=w_+(x)-w_-(x)$, where $w_+(x)=\max\{w(x),0\}$ and $w_-(x)=\max\{-w(x),0\}$, Definition 3 implies that for any $\theta_0\in\Theta$ and $e_0,\tilde{e}_0\in E$, with $\varphi^{\theta_0,e_0,\tilde{e}_0}\neq 0$,

$$\varphi^{\theta_0,e_0,\tilde{e}_0}(w)=0\quad\text{implies that}\quad\varphi^{\theta,e,\tilde{e}}(w)=0\quad\forall\theta\in\Theta,\forall e,\tilde{e}.$$

Hence, functionals $\varphi^{\theta,e,\tilde{e}}$ are equivalent for all $\theta\in\Theta$ and $e,\tilde{e}\in E$, i.e., there exist constants $\lambda^{\theta,e,\tilde{e}}\in\mathbb{R}$ and a linear functional $\bar{\varphi}$ in $\mathbb{R}^N$ such that $\varphi^{\theta,e,\tilde{e}}=\lambda^{\theta,e,\tilde{e}}\bar{\varphi}$. Indeed, we have that the null spaces of $\varphi^{\theta,e,\tilde{e}}$ are all the same, which we denote by $\mathcal{N}$. By the rank-nullity theorem, there exists $v\in\mathbb{R}^N\backslash\mathcal{N}$ such that $\mathbb{R}^N=[v]\oplus\mathcal{N}$, where $[v]$ is the subspace generated by vector v and $\oplus$ represents the direct sum between vector spaces. Let $\bar{\varphi}$ be the unique linear functional such that $\bar{\varphi}(v)=1$ and $\bar{\varphi}(n)=0$ for all $n\in\mathcal{N}$. Hence,

$$\lambda^{\theta,e,\tilde{e}}=\varphi^{\theta,e,\tilde{e}}(v)=\sum_{i=1}^{N}v_i\big(p^{\theta}_{e,i}-p^{\theta}_{\tilde{e},i}\big).$$

Defining $I(e,\theta)=-\sum_{i=1}^{N}v_ip^{\theta}_{e,i}$ and $h=\bar{\varphi}$, we have that

$$p^{\theta}_{e}-p^{\theta}_{\tilde{e}}=\varphi^{\theta,e,\tilde{e}}=\lambda^{\theta,e,\tilde{e}}\bar{\varphi}=\big[I(\tilde{e},\theta)-I(e,\theta)\big]h,$$

which implies MS.

Next, we show that (ii) implies (iii). From Definition 1, a distribution satisfies MS if and only if, for all $(x,\theta,e,\tilde{e})$,

$$p^{\theta}_{e}(x)-p^{\theta}_{\tilde{e}}(x)=\big[I(\tilde{e},\theta)-I(e,\theta)\big]h(x).\tag{30}$$

Since (ii) and (iii) trivially hold if $p^{\theta}_{e,i}=p^{\theta}_{\tilde{e},i}$ for all i,θ,e, and $\tilde{e}$, it suffices to consider the case where $p^{\theta}_{e,j}\neq p^{\theta}_{\tilde{e},j}$ for some $(j,\theta,e,\tilde{e})$.

Suppose (ii) holds. Then we must have $h_j\neq 0$. Multiplying both sides of (30) by $\frac{h_i}{h_j}$, we obtain

$$\big(p^{\theta}_{e,j}-p^{\theta}_{\tilde{e},j}\big)\frac{h_i}{h_j}=\big[I(\tilde{e},\theta)-I(e,\theta)\big]h_i=p^{\theta}_{e,i}-p^{\theta}_{\tilde{e},i}.$$

Setting $\phi_{i,j}\equiv\frac{h_i}{h_j}$ establishes the equation in the statement.

Finally, we establish that (iii) implies (i). To do so, suppose that (iii) holds, and let w and $\tilde{w}$ be two payment vectors satisfying LL such that for some $\theta_0\in\Theta$,

$$\sum_{i=1}^{N}\big(p^{\theta_0}_{e,i}-p^{\theta_0}_{\tilde{e},i}\big)w_i=\sum_{i=1}^{N}\big(p^{\theta_0}_{e,i}-p^{\theta_0}_{\tilde{e},i}\big)\tilde{w}_i\tag{31}$$

for all $e,\tilde{e}\in E$. Without loss of generality, suppose that $p^{\theta_0}_{e_0,j}\neq p^{\theta_0}_{\tilde{e}_0,j}$ for some $(j,\theta_0,e_0,\tilde{e}_0)$ (the distribution would immediately be ordered in terms of incentives if $p^{\theta}_{e,i}=p^{\theta}_{\tilde{e},i}$ for all $(i,\theta,e,\tilde{e})$). For notational simplicity (relabeling if needed), let $j=1$.

Use (29) to write

$$\sum_{i=1}^{N}(p_{e_0,i}^{\theta_0}-p_{\tilde{e}_0,i}^{\theta_0})w_i=(p_{e_0,1}^{\theta_0}-p_{\tilde{e}_0,1}^{\theta_0})\sum_{i=1}^{N}\phi_{i,1}w_i$$

and

$$\sum_{i=1}^{N}(p_{e_0,i}^{\theta_0}-p_{\tilde{e}_0,i}^{\theta_0})\tilde{w}_i=(p_{e_0,1}^{\theta_0}-p_{\tilde{e}_0,1}^{\theta_0})\sum_{i=1}^{N}\phi_{i,1}\tilde{w}_i.$$

Substitute these equations into (31),

$$(p_{e_0,1}^{\theta_0}-p_{\tilde{e}_0,1}^{\theta_0})\sum_{i=1}^{N}\phi_{i,1}w_i=(p_{e_0,1}^{\theta_0}-p_{\tilde{e}_0,1}^{\theta_0})\sum_{i=1}^{N}\phi_{i,1}\tilde{w}_i,$$

which, since $p_{e_0,1}^{\theta_0}\neq p_{\tilde{e}_0,1}^{\theta_0}$, implies

$$\sum_{i=1}^{N}\phi_{i,1}w_i=\sum_{i=1}^{N}\phi_{i,1}\tilde{w}_i.$$

For each θ, e, and $\tilde{e}$, multiply both sides by $p_{e,1}^{\theta}-p_{\tilde{e},1}^{\theta}$ to obtain

$$(p_{e,1}^{\theta}-p_{\tilde{e},1}^{\theta})\sum_{i=1}^{N}\phi_{i,1}w_i=(p_{e,1}^{\theta}-p_{\tilde{e},1}^{\theta})\sum_{i=1}^{N}\phi_{i,1}\tilde{w}_i.$$

Then using condition (29) again, gives

$$\sum_{i=1}^{N}(p_{e,i}^{\theta}-p_{\tilde{e},i}^{\theta})w_i=\sum_{i=1}^{N}(p_{e,i}^{\theta}-p_{\tilde{e},i}^{\theta})\tilde{w}_i,$$

which establishes (i). □

Proposition 4 shows the equivalence between three properties of a probability distribution. As discussed in the main text, property (i) states that if one type has the same incentives to exert two efforts under contracts w and $\tilde{w}$, so do all other types. In other words, all contracts can be ranked in terms of the incentives that they provide. Property (ii) is the MS condition used in the text.

Property (iii) provides a straightforward graphic interpretation of what it means for a distribution to be multiplicatively separable as explained in the text.

Appendix D: Binding individual rationality constraint

We showed that if the participation constraint does not bind, which in the model in the text is implied by limited liability, the optimal mechanism offers a single contract to all types. This appendix addresses situations in which the participation constraint may bind.

To simplify the analysis, suppose the agent exerts two levels of efforts $e \in \{0, 1\}$, which costs $0 = c_0^\theta < c_1^\theta$. Notice that the assumption that the least-costly effort has a nonpositive cost (i.e., $\min_{e\in E} c_e^\theta \leq 0$ for all θ) does not hold anymore. In this case, limited liability does no longer imply that the participation constraint does not bind.

The principal does not observe the effort chosen by the agent. She does, however, observe the output from the partnership $x \in \{x_L, x_H\}$, which is stochastically affected by the agent's effort. We refer to x_H as a high output or as success, to x_L as a low output or failure, and to $\Delta x := x_H - x_L > 0$ as the incremental output. Given effort e, a high output happens with probability p_e^θ. Let us define the incremental probability and incremental cost as

$$\Delta p^\theta = p_1^\theta - p_0^\theta \quad \text{and} \quad \Delta c^\theta = c_1^\theta - c_0^\theta.$$

By the revelation principle, we can focus on direct mechanisms. A direct mechanism is a triple of $\mathcal{B}(\Theta)$-measurable functions $(s, b, e) : \Theta \to \mathbb{R}^2 \times \{0, 1\}$, consisting of fixed payments s (or salaries), bonuses b, and effort recommendations e. An agent who reports type θ agrees to exert effort e^θ, and receives s^θ in case of failure and $s^\theta + b^\theta$ in case of success. A pair of payments s^θ and b^θ is called a *contract*.

Moreover, let us suppose that the gain from exerting high level of effort is sufficiently large (i.e., Δx is high enough) that makes the principal optimally elicit high effort for all types (i.e., $e^\theta = 1$ for all θ). Hence, as in Grossman and Hart (1983), the relevant incentive problem is the cost minimization problem of implementing high effort for all types. Therefore, define the cost minimization problem as

$$\min_{(s,b)} \int_\Theta \left(s^\theta + p_1^\theta b^\theta\right) d\mu(\theta) \tag{32}$$

subject to

$$\Delta p^\theta b^\theta \geq \Delta c^\theta \tag{IC$_1^\theta$}$$

$$s^\theta + p_1^\theta b^\theta \geq s^{\hat\theta} + p_1^\theta b^{\hat\theta} \tag{IC$_2^\theta$}$$

$$s^\theta + p_1^\theta b^\theta - \left[s^{\hat\theta} + p_0^\theta b^{\hat\theta}\right] \geq \Delta c^\theta \tag{IC$_3^\theta$}$$

$$s^\theta \geq 0, s^\theta + b^\theta \geq 0 \tag{LL$^\theta$}$$

$$s^\theta + p_1^\theta b^\theta \geq c_1^\theta \tag{IR$^\theta$}$$

for all θ, $\hat\theta$. The first three constraints refer to the incentive-compatibility constraint: (IC$_1^\theta$) is the incentive of not reducing effort conditional on truth-telling; (IC$_2^\theta$) is truth-telling conditional on obedience of the high effort recommendation; (IC$_3^\theta$) prevents the double deviation on truthful announcement and effort recommendation. Constraints (LL$^\theta$) and (IR$^\theta$) are the limited liability and participation constraints.

Our first result shows that if we can identify a least efficient type[21] such that all other types that have lower probability of success conditional on high effort are the ones that

[21] That is, a type with the highest cost of effort provision and with highest incremental cost per incremental probability.

also have a lower cost of effort provision per unit of probability, then every feasible mechanism is weakly dominated by the mechanism with the single contract given by this least efficient type. Let us state first the existence of this least efficient type.

ASSUMPTION 2. *There exists $\theta^* \in \Theta$ such that for all $\theta \in \Theta$: (i) $c_1^{\theta^*} \geq c_1^{\theta}$; (ii) $\frac{\Delta c^{\theta^*}}{\Delta p^{\theta^*}} \geq \frac{\Delta c^{\theta}}{\Delta p^{\theta}}$; (iii) $\frac{c_1^{\theta^*}}{p_1^{\theta^*}} \geq \frac{c_1^{\theta}}{p_1^{\theta}}$ if and only if $p_1^{\theta^*} \geq p_1^{\theta}$.*

Under this assumption, we can show the following proposition.

PROPOSITION 5. *Under Assumptions 1 and 2, a single contract is the optimal mechanism that solves problem (32).*

PROOF. Take any feasible mechanism $(s^\theta, b^\theta)_{\theta\in\Theta}$ for problem (32). Let (s^*, b^*) be the contract associated to the type θ^*. We claim that the mechanism formed by the single contract (s^*, b^*) is feasible for the program. In this case, it obviously generates a lower cost for the principal since we are only withdrawing contracts from the original menu. This new mechanism obviously satisfies (IC_1^θ)–(IC_3^θ) and (LL^θ). The only constraint that must be checked is (IR^θ), which can be equivalently rewritten as

$$s^\theta \geq p_1^\theta\left(\frac{c_1^\theta}{p_1^\theta} - b^\theta\right).$$

By construction, this constraint is satisfied for type θ^*. Now, from (IC_2^θ), we have that

$$p_1^\theta\left(b^\theta - b^{\theta^*}\right) \geq p_1^{\theta^*}\left(b^\theta - b^{\theta^*}\right).$$

Therefore, for θ such that $b^\theta < b^{\theta^*} = b^*$, we have that $p_1^\theta \leq p_1^{\theta^*}$ and, by Assumption 2, we get

$$s^* \geq p_1^\theta\left(\frac{c_1^\theta}{p_1^\theta} - b^*\right).$$

Using the same argument, for θ such that $b^\theta > b^{\theta^*} = b^*$, we have that $p_1^\theta \geq p_1^{\theta^*}$ and, by Assumption 2, we get

$$s^* \geq c_1^{\theta^*}\left(1 - \frac{p_1^{\theta^*}}{c_1^{\theta^*}}b^*\right) \geq c_1^\theta\left(1 - \frac{p_1^\theta}{c_1^\theta}b^*\right).$$

In both cases we have that (IR^θ) holds at contract (s^*, b^*). □

This proposition shows that in a two-by-two model, under the strong Assumption 2, all feasible mechanisms are dominated by one entailing full pooling. In the Suplemental Material we provide a complete characterization of the special case of two types (including when Assumption 2 fails). As we show there, the optimal mechanism may feature screening when this assumption fails.

References

Attar, Andrea, Thomas Marriotti, and François Salanié (2011), "Nonexclusive competition in the market for lemons." *Econometrica*, 79, 1869–1918. [1360]

Bajari, Patrick and Steven Tadelis (2001), "Incentives versus transaction costs: A theory of procurement contracts." *RAND Journal of Economics*, 32, 287–307. [1359]

Berge, Claude (1963), *Topological Spaces*. MacMillan. [1382]

Bergemann, Dirk and Stephen Morris (2005), "Robust mechanism design." *Econometrica*, 73, 1521–1534. [1358]

Biais, Bruno and Thomas Mariotti (2005), "Strategic liquidity supply and security design." *Review of Economic Studies*, 72, 615–649. [1360]

Bolton, Patrick and David S. Scharfstein (1990), "A theory of predation based on agency problems in financial contracting." *American Economic Review*, 80, 93–106. [1360]

Burguet, Roberto, Juan-José Ganuza, and Esther Hauk (2012), "Limited liability and mechanism design in procurement." *Games and Economic Behavior*, 76, 15–25. [1359]

Caillaud, Bernard, Roger Guesnerie, and Patrick Rey (1992), "Noisy observation in adverse selection models." *Review of Economic Studies*, 59, 595–615. [1361]

Carroll, Gabriel (2019), "Robustness in mechanism design and contracting." *Annual Review of Economics*, 11, 139–166. [1358]

Castro-Pires, Henrique and Humberto Moreira (2021), "Limited liability and non-responsiveness in agency models." *Games and Economic Behaviour*, 128, 73–103. [1369]

Chade, Hector and Edward Schlee (2020), "Insurance as a lemons market: Coverage denials and pooling." Journal of Economic Theory, 189. [1361]

Chade, Hector and Jeroen Swinkels (2019), "Disentangling moral hazard and adverse selection." Report. [1361]

Chaigneau, Pierre and Alex Edmans and Daniel Gottlieb (2018), "Does improved information improve incentives?" *Journal of Financial Economics*, 130 (2), 291–307. [1360]

Chiappori, Piere-André and Bernard Salanié (2003), "Testing contract theory: A survey of some recent work." In *Advances in Economics and Econometrics*, volume 1 (Mathias Dewatripont, Lars P. Hansen, and Stephen T. Turnovsky, eds.). Cambridge University Press, Cambridge. [1357, 1358]

Chu, Leon Yang and David Sappington (2007), "Simple cost-sharing contracts." *American Economic Review*, 97, 419–428. [1359]

Demarzo, Peter and Darrell Duffie (1999), "A liquidity-based model of security design." *Econometrica*, 67, 65–99. [1360]

DeMarzo, Peter M. (2005), "The pooling and tranching of securities: A model of informed intermediation." *Review of Financial Studies*, 18, 1–35. [1360]

DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz (2005), "Bidding with securities: Auctions and security design." *American Economic Review*, 95, 936–959. [1360]

Dewatripont, Mathias, Patrick Legros, and Steven A. Matthews (2003), "Moral hazard and capital structure dynamics." *Journal of the European Economic Association*, 1, 890–930. [1360]

Diamond, Douglas W. (1984), "Financial intermediation and delegated monitoring." *Review of Economic Studies*, 51, 393–414. [1360]

Faure-Grimaud, Antoine and Thomas Mariotti (1999), "Optimal debt contracts and the single-crossing condition." *Economic Letters*, 65, 85–89. [1360]

Gottlieb, Daniel and Humberto Moreira (2014), "Simultaneous adverse selection and moral hazard." Tech. rep, Wharton School and EPGE/FGV. [1373]

Grossman, Sanford J. and Oliver D. Hart (1983), "An analysis of the principal-agent problem." *Econometrica*, 51, 7–45. [1358, 1363, 1370, 1397]

Guesnerie, Roger and Jean-Jacques Laffont (1984), "A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm." *Journal of Public Economics*, 25, 329–369. [1360]

Hart, Oliver D. and Bengt Holmstrom (1987), "The theory of contracts." In *Advances in Economic Theory, Fifth World Congress* (Thomas Bewley, ed.). Cambridge University Press, Cambridge. [1357, 1363]

Holmstrom, Bengt (1979), "Moral hazard and observability." *The Bell Journal of Economics*, 10, 74–91. [1363, 1367, 1370]

Innes, Robert D. (1990), "Limited liability and incentive contracting with ex-ante action choices." *Journal of Economic Theory*, 52, 45–67. [1359, 1360, 1369, 1370, 1371]

Jewitt, Ian, Ohad Kadan, and Jeroen M. Swinkels (2008), "Moral hazard with bounded payments." *Journal of Economic Theory*, 143, 59–82. [1360]

Kadan, Ohad, Philip Reny, and Jeroen M. Swinkels (2017), "Existence of optimal mechanisms in principal-agent problems." *Econometrica*, 85, 769–823. [1375]

Laffont, Jean-Jacques and David Martimort (2002), *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press. [1361]

Laffont, Jean-Jacques and Jean Tirole (1986), "Using cost observation to regulate firms." *Journal of Political Economy*, 94, 614–641. [1359, 1361, 1371]

Laffont, Jean-Jacques and Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*. MIT press. [1359, 1361, 1371]

Matthews, Steven A. (2001), "Renegotiating moral hazard contracts under limited liability and monotonicity." *Journal of Economic Theory*, 97, 1–29. [1360, 1370]

Melumad, Nahum D. and Stefan Reichelstein (1989), "Value of communication in agencies." *Journal of Economic Theory*, 47, 334–368. [1361]

Nachman, David C. and Thomas H. Noe (1994), "Optimal design of securities under asymmetric information." *Review of Financial Studies*, 7, 1–44. [1360]

Netz, Janet S. (2000), "Price regulation: A (non-technical) overview." *Encyclopedia of Law and Economics*, 3, 396–466. [1359]

Ollier, Sandrine and Lionel Thomas (2013), "Ex post participation constraint in a principal–agent model with adverse selection and moral hazard." *Journal of Economic Theory*, 148, 2383–2403. [1361]

Picard, Pierre (1987), "On the design of incentive schemes under moral hazard and adverse selection." *Journal of Public Economics*, 33, 305–331. [1361]

Poblete, Joaquín and Daniel Spulber (2012), "The form of incentive contracts: Agency with moral hazard, risk neutrality, and limited liability." *RAND Journal of Economics*, 43, 215–234. [1360]

Riley, John and Richard Zeckhauser (1983), "Optimal selling strategies: When to haggle, when to hold firm." *The Quarterly Journal of Economics*, 15, 267–289. [1360]

Rochet, Jean-Charles and Lars A. Stole (2002), "Nonlinear pricing with random participation." *Review of Economic Studies*, 69, 277–311. [1362]

Rogerson, William P. (2003), "Simple menus of contracts in cost-based procurement and regulation." *American Economic Review*, 93, 919–926. [1359]

Rudin, Walter (1991), *Functional Analysis*, second edition. McGraw-Hill, Inc. [1390]

Tirole, Jean (2005), *The Theory of Corporate Finance*. Princeton University Press. [1359]

Townsend, Robert M. (1979), "Optimal contracts and competitive markets with costly state verification." *Journal of Economic Theory*, 21, 265–293. [1360]

Co-editor Thomas Mariotti handled this manuscript.

Manuscript received 7 September, 2017; final version accepted 23 July, 2021; available online 2 September, 2021.