Security Design with Flexible Moral Hazard and Limited Liability

Daniel Krähmer*

March 8, 2026

Abstract

I study security design with a risk-neutral entrepreneur and a risk-neutral investor who are both protected by limited liability. The project return is determined by an unobservable action by the entrepreneur (moral hazard). Moral hazard is flexible: the entrepreneur can choose any distribution of returns subject to a cost. I characterize the set of implementable distributions and when the first-best is implementable (and optimal). I derive optimal distributions for cost functions that are increasing or decreasing in risk or depend only on moments of the distribution, and show that they are first order stochastically dominated by the first-best. Securities that implement optimal distributions are not unique.

Keywords: Security Design, Flexible Moral Hazard, Limited Liability

JEL: D21, D82, D86, G32

*Universität Bonn (kraehmer@hcm.uni-bonn.de); I thank Martin Hellwig and Roland Strausz for very valuable discussions. Special thanks go to Mathijs Janssen. I gratefully acknowledge financial support from the German Research Foundation (DFG) through Germany's Excellence Strategy EXC 2126/1–390838866 and CRC TR 224 (project B03).

1 Introduction

I study a canonical security design problem where an entrepreneur seeks outside funding for a project by issuing a security that promises an investor a portion of the realized project return. The probability distribution of project returns depends on a hidden action by the entrepreneur, that is, there is moral hazard. As emphasized in the seminal work of Jensen and Meckling (1976), and as formalized by more recent contributions (e.g., Biais and Casamatta, 1999; Hellwig, 2009; Hébert, 2018), an important issue of the moral hazard problem in security design is that the entrepreneur's available actions may not only include costly effort that increases project returns but also unproductive actions that only affect the risk of returns. This can give rise to the "risk-shifting problem" that the entrepreneur may take on excessive risk in an attempt to conceal lax effort behind a high return realization.

In the spirit of Hébert (2018), the present paper studies this issue by adopting a flexible moral hazard approach where the entrepreneur can choose any return distribution subject to a cost. My paper differs from Hébert (2018) in two respects: First, Hébert (2018) is interested in the optimality of certain financial instruments, notably debt, whereas the focus of my paper is to study the economic distortions that arise from flexible moral hazard. Second, Hébert (2018) restricts attention to cost functions that are in the class of divergences, whereas I adopt the "smooth" flexible moral hazard approach by Georgiadis et al. (2024) which allows for all cost functions that are convex, increasing (in first-order stochastic dominance), and smooth in the sense that they admit a Gateaux derivative.

Specifically, I apply the smooth flexible approach to the security design model by Innes (1990), in which, instead, the entrepreneur chooses a one-dimensional action that captures productive effort. Next to moral hazard, there is two-sided limited liability and a funding constraint: The entrepreneur's liability is limited to the project return, and the investor is not liable for losses incurred by the entrepreneur. Moreover, to participate in the venture, the investor requires the expected security payout cover the cost of capital he provides.

The main results of my paper are as follows. I provide a general condition that characterizes when the first-best outcome is a solution to the security design problem. I argue that these conditions are harder to satisfy the higher the degree of convexity of the cost function. In fact, when the cost function is linear, the condition is always satisfied, and hence there are no distortions.

Economically, linearity means that the cheapest way to generate a stochastic return distribution is to randomize over the deterministic outcomes in its support. An alternative interpretation is that the marginal costs to alter a return distribution are constant (that is, independent of this distribution).

If the first-best outcome is not optimal, I derive second-best solutions for cost functions which, for a given mean of the distribution, are monotone in its risk (in the sense of a mean preserving spread). I also consider "moment-based" cost functions that depend on finitely many generalized moments of the distribution but might neither be increasing nor decreasing in risk. Under these specifications, the problem is remarkably tractable, and I show that when costs are monotone in risk and moment-based, or when costs depend on a single moment only, second-best solutions are first-order stochastically dominated by the first-best. In this sense, the second-best does not display socially excessive, but too little risk. This, in particular, is true when costs are decreasing in risk where the risk-shifting problem seems perhaps most severe. Note, however, that when costs are decreasing in risk, risk is also socially desirable, as it reduces costs.

To derive these results, I first ask which return distributions can be implemented by some security and then search for the optimal return distributions among these (in the spirit of Grossman and Hart, 1983).[1] This approach highlights the basic trade-off that underlies the security design problem: in a first-best world, the entrepreneur would like to implement the return distribution that maximizes the total expected return and just compensate the investor for his capital costs. Due to moral hazard and limited liability, however, there are agency costs: The entrepreneur must be afforded a moral hazard rent, and the remaining portion of the payout might not be sufficient to cover the investor's capital costs. An optimal distribution thus maximizes the expected return subject to agency costs being sufficiently low to allow the investor to recoup his investment.

When the cost function is linear, agency costs are shown to be zero, and thus the optimal design is efficient. The intuitive reason why agency costs are zero in this case, is familiar from standard moral hazard problems: to incentivize a target action, the entrepreneur is compensated for her *marginal* cost while her overall utility is (proportional to) her compensation minus her *average* cost. When costs are linear, marginal and average effort costs are the same, resulting in zero overall utility for the entrepreneur, hence zero agency costs.[2]

[1]The literature often imposes the constraint that the security be monotone. In this paper, I do not impose this constraint at any point.

[2]The same force drives related results in Krähmer (2025, 2026) and in Georgiadis et al. (2024).

To derive second-best solutions, I use fairly simple arguments that exploit the structures of the cost functions. Monotonicity in risk implies constant sign of the curvature of the Gateaux derivative, which leads to objective functions that are either convex or concave. This observation can be used to show that first- and second-best solutions are either degenerate or place all probability on the most extreme returns only. When costs are moment-based, I use extreme point arguments based on Winkler (1988) to show that the resulting first- and second-best distributions are discrete with a finite support. When costs depend on a single moment only, a by now familiar concavification argument can be used to rank the first- and the second-best.

A prominent question in the literature is whether the theory can explain real world securities such as debt or equity as the outcome of an optimal design.[3] In particular, in Innes' (1990) parametric and in Hébert's (2018) divergence cost based approach, debt arises as an optimal security when securities are required to be monotone.[4] By contrast, in my setting with little specific assumptions on the cost function, there is no unique security that implements the optimal return distribution, even if this distribution and the resulting security payout is unique (which is often the case). The reason is that in my setting, optimal securities are only pinned down on the support of optimal distributions but can be specified fairly arbitrarily off the support.[5] I show that under the cost functions I consider, all optimal distributions have finite support within the continuous interval of possible returns, thus leaving an abundance of degrees of freedom to specify optimal securities. By contrast, parametric approaches typically only consider full support distributions. Similarly, in Hébert (2018), the cost of a distribution without full support is infinite. This typically pins down the optimal security, and when combined with the additional requirement that the security be monotone, a debt security arises as optimal.[6]

Related literature

I contribute to the literatures on flexible moral hazard and security design by applying the smooth flexible moral hazard approach of Georgiadis et al. (2024) to the security design problem of Innes (1990). In Georgiadis et al. (2024), the principal offers a wage contract to an agent who covertly chooses a return distribution and is protected by limited liability. The security design problem differs in two aspects: First, it is the party who offers the contract (the entrepreneur) who also

[3] See Allen and Barbalau (2024) for a review.

[4] Hébert (2018) shows that monotonicity is not required in the special case of KL-divergence.

[5] On this point, see also Hellwig (2009).

[6] See also Yang and Zefentis (2024) who identify optimal monotone securities as extreme points of monotone functions.

chooses the hidden action. Second, there are additional constraints: both parties are protected by limited liability, and there is a funding constraint that requires the investor to break even. The first point implies that the objective function is different. As I will show, the second point implies that not all distributions are implementable by some contract, unlike in Georgiadis et al. (2024).[7]

Hébert (2018) considers security design within a flexible moral hazard framework similar to mine but considers a finite return space and focusses on the class of divergence cost functions. These are not included in the set of cost functions that I allow, as divergences are not monotone and are infinite for distributions that do not have full support. Moreover, Hébert (2018) focusses on the optimality of certain financial instruments whereas I am interested in the distortions caused by moral hazard.[8]

In Hellwig (2009), the entrepreneur can choose among two parameters, the failure risk of the venture as well as its mean return if it succeeds. While the return in case of failure is zero, the return in case of success depends positively on the entrepreneurs costly effort as well as on the chosen failure risk. Hellwig (2009) shows that the second-best exhibits inefficiently large failure risk. Failure risk is not well-defined in my model with full flexibility, and I compare first- and second-best solutions in terms of first-order stochastic dominance.

Biais and Casamatta (1999) consider a security design problem where the entrepreneur has two effort and two risk choices with a risk-return trade-off. Only the high-effort, low-risk project is viable, and either the agency costs necessary to prevent excessive risk taking are sufficiently low, or a complete funding break-down occurs. In my flexible approach, the risk-return trade-off is smoothly captured by the cost function, not by a restriction to the action set. This allows me to study how return distortions depend on the cost function.

Hellwig (2025) considers a flexible moral hazard problem where the agent's costs depend on the mean and (minus) the variance of the distribution and is thus decreasing in risk. The key difference to my paper is that the agent is risk averse and effort costs are monetary. Similar to my setting when costs depend on two moments of the distribution, second-best solutions have at most three points in the support. Hellwig (2025) finds that the agent's remuneration might be non-monotone in returns. In contrast, as already noted in Georgiadis et al. (2024), in

[7]For other applications of the smooth flexible moral hazard approach to contracting problems with moral hazard and adverse selection, see Krähmer (2025), Liu (2025), Castro-Pires et al. (2025).

[8]Flexible moral hazard problems with finite outcome spaces are also considered in Mattson and Weibull (2023) in the context of wage design, and in Kocherlakota (2025) in the context of taxation.

the smooth flexible framework with monotone costs, the agent's/entrepreneur's remuneration is always monotone on the support of returns.

The paper is organized as follows. Secton 2 presents the model. Section 3 provides a characterization of implementable distributions. Section 4 derives the security design problem and characterizes when the first-best is implementable and optimal. Section 5 solves the security design problem for various classes of cost functions. Section 6 concludes. All proofs are in the appendix.

2 Model

There are a risk neutral entrepreneur (she) and a risk neutral investor (he). The entrepreneur needs funds $K > 0$ to conduct a project that pays a contractible return $x \in X = [\underline{x}, \overline{x}]$, $0 \leq \underline{x} < \overline{x}$. The return x is distributed with a cdf F that is the result of the entrepreneur's effort choice. I assume that the entrepreneur's effort is flexible: she can choose any return distribution F subject to the effort cost $C(F)$.

The entrepreneur owns no capital (for simplicity), and to obtain financing for the project by the investor, she issues a security $S : X \to \mathbb{R}$ that promises to pay the investor the amount $S(x)$ if the project return x realizes.

Both parties are protected by limited liability: The entrepreneur cannot pay out more than the project return, that is, $S(x) \leq x$, and the investor is not liable for any losses incurred by the entrepreneur, that is, $0 \leq S(x)$.

The timing is as follows. The entrepreneur commits to a security. The investor decides whether to invest. If he does not invest, both parties get their outside option of zero. If the investor invests, the entrepreneur covertly chooses a return distribution F, that is, there is moral hazard. Finally, returns realize and are divided as specified by the security.

The entrepreneur's problem is

$$
\begin{aligned}
P_0: \qquad & \max_{S,F} \int x - S(x)\, dF - C(F) \quad s.t. \\
& F \in \arg\max_G \int x - S(x)\, dG - C(G), && (MH) \\
& 0 \le S(x) \le x \quad \forall x \in X, && (LL) \\
& \int S(x)\, dF \ge K. && (IR)
\end{aligned}
$$

Note that limited liability is required for all possible returns x, not only for x in the support of the distribution that the entrepreneur actually chooses. The reason is that if the entrepreneur were to deviate to a distribution with different support, the outcome would still be subject to limited liability.

Next, I state the assumptions on the cost function which parallel those in Georgiadis et al. (2024).

1. C is continuous, monotone, and convex.[9]

2. C is smooth in the sense that C is Gateaux-differentiable with continuous and differentiable Gateaux derivative $c_F : X \to \mathbb{R}$, that is, for $F, \tilde{F}$, we have[10]

$$
\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \left[C(F + \epsilon(\tilde{F} - F)) - C(F) \right] = \int c_F(x)\, d(\tilde{F} - F).
$$

3. The cost of the "smallest" distribution, which places mass 1 on $\underline{x}$, is normalized to 0: $C(\delta_{\underline{x}}) = 0$.[11]

Continuity is a technical condition that ensures the existence of various maximizers below. Monotonicity means that larger return distributions are more costly, that is, $C(F) \ge C(G)$ if F first order stochastically dominates G. This is a natural assumption in the effort provision context considered here. As pointed out by Georgiadis et al. (2024), convexity is without loss because the entrepreneur can randomize and thus $C(F)$ should be the expected cost of the cheapest randomization that generates it, implying convexity.

[9] Continuity refers to the weak topology.

[10] It is well-known that the Gateaux derivative is unique only up to a constant. The reason is that F and $\tilde{F}$ in (1) are both cdfs. Thus adding a constant to c_F does not affect the right-hand side.

[11] δ_x denotes the cdf that places mass 1 on x.

Smoothness captures a notion of differentiability which will make the analysis tractable. As is well-known, the Gateaux derivative is a functional derivative that generalizes the notion of a partial derivative from functions of vectors to functions of functions. Economically, the Gateaux derivative $c_F(x)$ evaluated at x measures the marginal cost of increasing the probability mass assigned to x given F. The final assumption is a normalization that ensures that "zero effort" has no cost.

I will make use of the well-known fact that for smooth costs, monotonicity is characterized by monotonicity of the Gateaux-derivative, that is, monotonicity is equivalent to $c_F(x)$ being increasing in x for all F (see, e.g., Cerreia-Vioglio et al., 2017).

A first-best distribution F^{fb} maximizes the total expected project value $\int x \, dF - C(F) - K$. Let

$$V^{fb} = \max_F \int x \, dF - C(F) - K$$

be the first-best project value. To make the problem non-trivial, I assume that V^{fb} is positive.

3 Implementability

I start the analysis by studying which distributions are implementable by some security. F is implementable if there is a security so that all the constraints in the entrepreneur's problem are satisfied.

As the first step, it is instructive to isolate the restrictions that arise from moral hazard and limited liability alone which are at the core of the incentive problem.[12] I say that F is incentive-feasible if there is a security S so that the constraints (MH) and (LL) in the entrepreneur's problem are satisfied. The next proposition characterizes incentive-feasibility. To state it, let

$$\underline{\lambda}_F = -c_F(\underline{x}), \quad \overline{\lambda}_F = \min_{x \in supp(F)} (x - c_F(x)). \tag{1}$$

Proposition 1 *F is incentive-feasible if and only if there is $\lambda \in \mathbb{R}$ so that*

$$\underline{\lambda}_F \le \lambda \le \overline{\lambda}_F. \tag{LL'}$$

[12] This incentive problem also arises if, for example, the investor has all the bargaining power and there is no participation constraint.

The key to understanding the proposition is a characterization of the moral hazard constraint (MH) in terms of a first-order condition that equates the entrepreneur's marginal benefits and marginal costs from effort. Specifically, following Georgiadis et al. (2024), the constraint (MH) is satisfied if and only if there is λ so that[13]

$$S(x) = x - c_F(x) - \lambda \quad \forall x \in supp(F), \tag{MH_{supp}}$$

$$S(x) \geq x - c_F(x) - \lambda \quad \forall x \in X. \tag{MH_{all}}$$

Intuitively, $c_F(x)$ measures the entrepreneur's cost of marginally increasing the probability mass assigned to the return x, given F, and $x - S(x)$ is the entrepreneur's benefit of marginally increasing the probability mass assigned to return x (which is independent of F). The constant λ corresponds to the marginal shadow costs that comes from the constraint that the total probability mass the entrepreneur can allocate is equal to one. A distribution F is therefore optimal if and only if for no possible return level x, the entrepreneur can strictly improve by marginally increasing the probability of x (condition (MH_{all})), and for all return levels in the support of F, the benefits and costs of marginally increasing the probability of x are the same (condition (MH_{supp})).

A useful interpretation is that equation (MH_{supp}) defines a security which specifies a "lump sum" λ (possibly negative) that the entrepreneur retains irrespective of the realized return, and when x realizes, the entrepreneur retains the additional amount $c_F(x)$ which is pinned down by the distribution F.

The moral hazard constraints *alone* do not restrict the magnitude of the lump sum, because, as is usual, the magnitude of the lump sum does not affect effort incentives.[14] However, the lump sum does affect whether the security satisfies the additional constraints (LL). The lump sum must not be too large (so that the investor's limited liability is not violated) and not be too small (so that the entrepreneur's limited liability is not violated). Condition (LL') specifies the range of λ's which are consistent with (LL).

It is instructive to illustrate the incentive-feasibility constraints (MH_{supp}), (MH_{all}), and (LL')

[13]Strictly speaking, (MH) has to hold for only F-almost all $x \in supp(F)$. Throughout the paper, I abstract from measure-theoretic subtleties. Moreover, the distributions that arise endogenously in my setting are all discrete and finite.

[14]In particular, the moral hazard constraints alone do not restrict the set of implementable distribution. This is a well-known feature of flexible moral hazard with a risk neutral agent (see Georgiadis et al., 2024).

graphically. The left panel of Figure 1 depicts a distribution F which satisfies these constraints. The support of F is $\{x_1, x_2, x_3\}$. The figure plots the curve $x - c_F(x) - \lambda$ for various values of λ. For a security to induce F with the specific lump sum λ_0, it has to pass through the curve $x - c_F(x) - \lambda_0$ on the support by (MH_{supp}) and lie above the curve elsewhere by (MH_{all}). Due to (LL), the security, moreover, has to be positive and smaller than the 45 degree line. Hence, any security which passes through the dotted points and is otherwise located in the shaded area satisfies the moral hazard and limited liability constraints for the lump sum λ_0. By decreasing (increasing) the lump sum, the curve shifts up (down). The lump sum $\underline{\lambda}_F$ is the smallest lump sum so that the security can be chosen to meet the entrepreneur's limited liability constraint $S(x) \leq x$, and the lump sum $\overline{\lambda}_F$ is the largest lump sum so that the security can be chosen to meet the investor's limited liability constraint $S(x) \geq 0$. Clearly, the depicted constellation has $\underline{\lambda}_F < \overline{\lambda}_F$, and ($LL'$) holds.

In the right panel of Figure 1, the support of the distribution F is $\{x_1, x_2\}$. This distribution is not incentive-feasible. Indeed, the blue curve now plots $x - c_F(x) - \lambda$ for the smallest possible lump sum $\lambda = \underline{\lambda}_F$ that is still consistent with the LL constraint $S(x) \leq x$. Any smaller lump sum would shift the curve upwards and imply a violation of the constraint $S(\underline{x}) \leq \underline{x}$. For the security to implement F, it would need to pass through the dotted points in order to satisfy the moral hazard constraint (MH_{supp}). But this is impossible without violating the constraint $S(x_2) \geq 0$. In other words, in the right panel, (LL') cannot be satisfied because $\underline{\lambda}_F > \overline{\lambda}_F$.

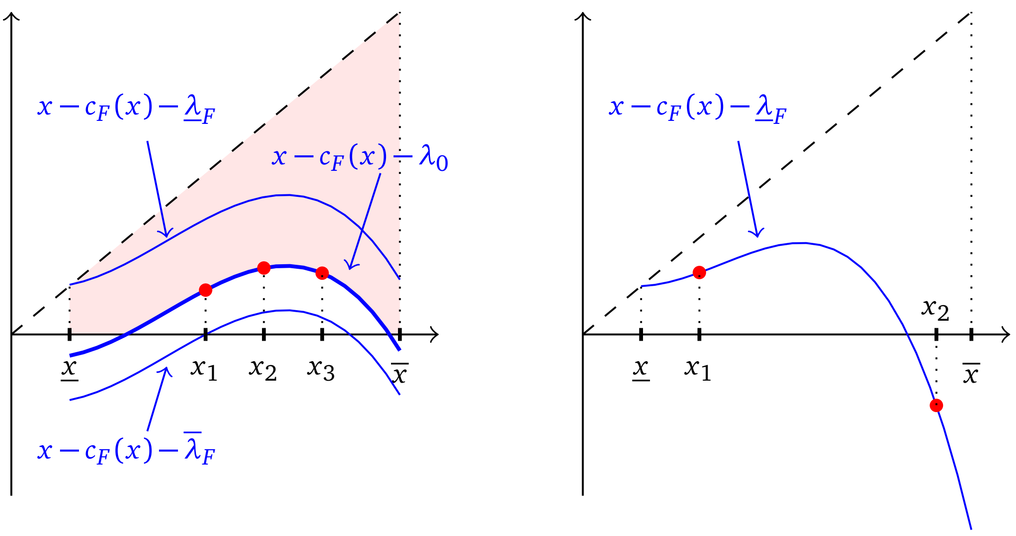


Figure 1: The figure illustrates when the incentive-feasibility conditions (MH_{supp}), (MH_{all}), and (LL') can (left) and cannot (right) be satisfied

It is now straightforward to characterize when a distribution is implementable, that is, when in addition to the moral hazard and limited liability constraints, individual rationality is satisfied. Indeed, observe that by (MH_{supp}), a security which induces the entrepreneur to choose a given F is pinned down on the support of F up to the lump sum λ. Therefore, the investor's expected payoff is $\int S(x)\, dF = \int x - c_F(x)\, dF - \lambda$, and the individual rationality constraint is therefore equivalent to

$$\lambda \leq \int x - c_F(x)\, dF - K \equiv \lambda_F^{IR}. \tag{IR'}$$

4 Optimal security design

I now turn to the problem of finding an optimal security and an optimal distribution. Because a security that implements a given F is pinned down on the support of F up to the lump sum λ by (MH_{supp}), the entrepreneur's expected payoff from implementing F, gross of effort costs $C(F)$, is $\int x - S(x)\, dF = \int c_F(x)\, dF + \lambda$.

In light of (LL') and (IR'), the entrepreneur's problem P_0 can therefore be re-written as

$$P: \qquad \max_{F,\lambda} \int c_F(x)\, dF + \lambda - C(F) \quad s.t. \quad \underline{\lambda}_F \leq \lambda \leq \overline{\lambda}_F, \quad \lambda \leq \lambda_F^{IR}.$$

Before I solve P formally, a reminder of the basic trade-off that underlies the security design problem is useful. Recall that $\lambda = \underline{\lambda}_F = -c_F(\underline{x})$ is the smallest lump sum that allows the entrepreneur to implement F and respect the limited liability constraint $S(x) \leq x$. The combination of moral hazard and this constraint implies that when F is implemented, the entrepreneur obtains *at least* the expected payout

$$\Pi_E(F) = \int c_F(x)\, dF - c_F(\underline{x}) \tag{2}$$

as a gross moral hazard rent (gross of effort cost). Accordingly, because the remaining portion of the return is paid out to the investor, the investor obtains *at most* the expected payout

$$\Pi_I(F) = \int x\, dF - \Pi_E(F).$$

Absent any frictions, the entrepreneur would optimally maximize the overall expected project return by committing to choose a first-best distribution F^{fb}. She could then just compensate the investor for her capital costs K and in this way obtain the first-best project value V^{fb}. The key issue, however, is that the entrepreneur's minimal payout $\Pi_E(F^{fb})$ that is required to induce the entrepreneur to choose the first-best under moral hazard and limited liability might be too large for the remaining payout $\Pi_I(F^{fb})$ to allow the investor to recoup his capital costs K.

Conversely, this suggests that if there is a first-best distribution F^{fb} with $\Pi_I(F^{fb}) \geq K$ which, in addition, is implementable, then F^{fb} is a solution to the security design problem. I now make these considerations more precise.

4.1 Implementability of the first-best

To examine when the first-best is implementable, I begin by characterizing first-best distributions.

Lemma 1 *A distribution F^{fb} is first-best if and only if there is μ so that*

$$x - c_{F^{fb}}(x) - \mu = 0 \quad \forall x \in supp(F^{fb}), \tag{FB_{supp}}$$

$$x - c_{F^{fb}}(x) - \mu \leq 0 \quad \forall x \in X. \tag{FB_{all}}$$

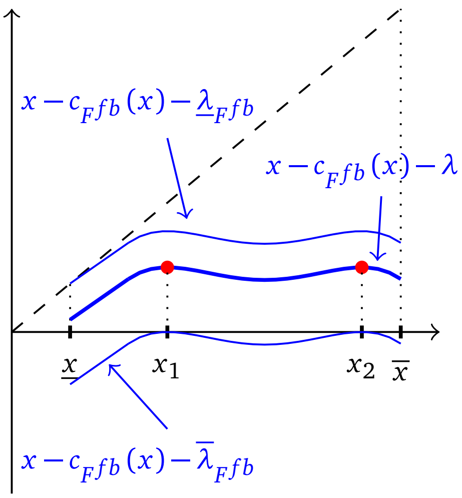
The conditions in the lemma are the first-order conditions that equate the marginal (first-best) project return and the marginal costs of effort. They are analogous to the moral hazard conditions, but in the first-best the benefit from marginally increasing the probability mass assigned to the return x is equal to x (instead of the entrepreneur's payout $x - S(x)$).

Next, I argue that a first-best distribution is incentive-feasible.

Lemma 2 *A first-best distribution F^{fb} is incentive-feasible, that is,*

$$\underline{\lambda}_{F^{fb}} \leq \overline{\lambda}_{F^{fb}}. \tag{3}$$

The reason is that the conditions (FB_{supp}) and (FB_{all}) imply that the function $x - c_{F^{fb}}(x)$ is maximized on the support of F^{fb}. This is illustrated in Figure 2 which displays a first-best distribution with support $\{x_1, x_2\}$. I now use Figure 2 to show that F^{fb} is incentive-feasible. As explained above, a security that satisfies (MH_{supp}), (MH_{all}), and (LL') for F^{fb} has to pass through the dotted points, lie above the curve $x - c_{F^{fb}}(x) - \lambda$, be positive and below the 45 degree line. It is



A first-best distribution

Figure 2: Incentive-feasibility of the first-best

evident from the figure that it is always possible to find such a security: The problem that $\underline{\lambda}_{F^{fb}}$ might be larger than $\overline{\lambda}_{F^{fb}}$ that arises in the right panel in Figure 1 does not arise for a first-best distribution, implying (3).[15]

Because a first-best distribution F^{fb} is incentive-feasible, it can be implemented if, in addition, the investor's individual rationality constraint can be satisfied. By (3), this is the case if and only if $\underline{\lambda}_{F^{fb}} \leq \lambda^{IR}_{F^{fb}}$. This inequality says that the smallest lump sum $\underline{\lambda}_{F^{fb}}$ that the entrepreneur has to retain to satisfy her limited liability constraint $S(x) \leq x$, is sufficiently small so that the investor's IR constraint can be satisfied.

Now observe that if F^{fb} is implementable, then F^{fb} is optimal for the entrepreneur. The reason is that she can then choose the lump sum so as to just compensate the entrepreneur for her capital costs K and in this way extract the entire first-best project value. The next proposition summarizes.

Proposition 2 *1. A first-best distribution F^{fb} is implementable if and only if $\underline{\lambda}_{F^{fb}} \leq \lambda^{IR}_{F^{fb}}$, or, equivalently, $\Pi_I(F^{fb}) - K \geq 0$.*

2. If a first-best distribution F^{fb} is implementable, then $F^* = F^{fb}$ and $\lambda^* = \Pi_I(F^{fb}) - K - c_{F^{fb}}(\underline{x})$

[15] Formally, because $x - c_{F^{fb}}(x)$ is maximized by any point in the support of F^{fb} by (FB_{supp}), it follows

$$\underline{\lambda}_{F^{fb}} = -c_{F^{fb}}(\underline{x}) \leq \underline{x} - c_{F^{fb}}(\underline{x}) \leq \min_{x \in supp(F^{fb})} (x - c_{F^{fb}}(x)) = \overline{\lambda}_{F^{fb}}. \tag{4}$$

is a solution to P, and the entrepreneur's ex ante profit is the first-best project value V^{fb}.

Next, I show that the degree of the convexity of the cost function drives whether the first-best is implementable or not. To do so, I denote by

$$\Pi_E^{net}(F) = \Pi_E(F) - C(F)$$

the entrepreneur's *net moral hazard rent*. The condition $\Pi_I(F^{fb}) \geq K$ for the implementability of the first-best can then be expressed as saying that the difference between the first-best surplus and the entrepreneur's net moral hazard rent is larger than K:

$$\int x dF^{fb} - C(F^{fb}) - \Pi_E^{net}(F^{fb}) \geq K. \tag{5}$$

Since the first-best surplus is larger than K by assumption, the first-best is therefore implementable if the net moral hazard rent is not too large.

Next, I show that the first-best is indeed implementable in the special but important case that costs are linear. C is linear if $C(F) = \int \varphi(x)\, dF$ for some function φ.[16]

Proposition 3 *Let C be linear. Then the net moral hazard rent* $\Pi_E^{net}(F)$ *is zero for all* F*. In particular, the first-best is implementable.*

The result follows from the fact that for linear costs, the entrepreneur's minimal payout, $\Pi_E(F)$, is equal to costs $C(F)$ for all distributions F.[17] In other words, any implementable distribution can be implemented without affording the entrepreneur a net moral hazard rent.

Intuitively, recall that in models with limited liability and risk-neutral agents, the principal must generally leave the agent a rent to induce costly effort. At an optimal contract, the agent is compensated at her marginal cost of effort—the contract adjusts pay so that each incremental increase in effort is rewarded just enough to offset its marginal cost. When the cost function is linear, marginal and average effort costs coincide. In that case, compensating the agent for marginal costs exactly covers total effort costs, leaving no rent. As a consequence, the first-best can be implemented without agency costs.

[16]Costs being linear means that the entrepreneur can generate a stochastic return distribution F only through a "mixed strategy" that randomizes over deterministic returns x, each costing $\varphi(x)$, according to F, and the costs of doing so is "expected costs". (When costs C are convex, such a mixed strategy is more costly than choosing F directly.)

[17]A similar observation appears in slightly different context in Georgiadis et al. (2024) and Krähmer (2025, 2026).

The previous reasoning suggests that the more convex the cost function, the higher the higher the net hazard rent the entrepreneur accumulates. The next result confirms this. I say that $\tilde{C}$ is more convex than C if the difference $\tilde{C} - C$ is convex.

Proposition 4 *Let $\tilde{C}$ be more convex than C, then $\widetilde{\Pi}_E^{net}(F) \geq \Pi_E^{net}(F)$ for all F.*

Note that in the statement, F is arbitrary but fixed. Because a first-best distribution F^{fb} depends on the cost function, the result does not say that the net moral hazard rent, evaluated at the first-best, increases in the degree of the convexity of the cost function. Therefore, Proposition 4 identifies merely one force that contributes to the severity of the agency problem. For a more complete comparative statics of the left-hand side of (5) additional conditions need to be identified which is beyond the scope of this paper.

5 Second-Best

I now turn to the entrepreneur's problem when the first-best cannot be implemented and identify the resulting distortions. The fact that the first-best is incentive-feasible by Lemma 2 implies that if the first-best is not implementable, it is because the investor's IR constraint is violated even when the lump sum takes the smallest possible value $\underline{\lambda}_{F^{fb}}$ that is consistent with the entrepreneur's LL constraint $S(x) \leq x$. This suggests to consider the relaxed problem where the lump sum λ is only required to satisfy the entrepreneur's limited liability constraint $S(x) \leq x$ and the IR constraint, that is,

$$\underline{\lambda}_F \leq \lambda \leq \lambda_F^{IR}. \tag{6}$$

Since the entrepreneur's objective in P is increasing in λ, it is optimal to set $\lambda = \lambda_F^{IR}$. If one plugs this into the objective and observes that the constraint (6) is equivalent to the condition $\int x\, dF - \Pi_E(F) \geq K$, one arrives at the relaxed problem

$$R: \qquad \max_F \int x\, dF - C(F) - K \quad s.t. \quad \int x\, dF - \Pi_E(F) - K \geq 0. \tag{7}$$

The relaxed problem has the intuitive interpretation that the entrepreneur maximizes the total project value subject to the constraint that the minimal payout to the investor covers his capital costs.

If K is large, the feasible set in problem R is empty, and the project is not funded. To rule out this rather uninteresting case, I from now on assume:

$$\sup_F \int x\, dF - \Pi_E(F) > K.$$

I shall solve problem R for various cost structures. Costs are increasing (resp. decreasing) in risk if $C(F) \geq C(G)$ (resp. $\leq$) whenever F is a mean preserving spread of G. I say C is mean-based if it depends only on the mean of the distribution, in which case it is both increasing and decreasing in risk.[18] I also consider costs which depend only on K-many generalized moments and are defined as follows. Define the moment vector

$$\Phi_F = (\Phi_{1,F}, \ldots, \Phi_{K,F}), \quad \Phi_{k,F} = \int \varphi_k(x) dF,$$

where for all k, $\varphi_k : X \to \mathbb{R}$ is an increasing function with $\varphi_k(\underline{x}) = 0$. Then C is moment-based if $C(F) = \Gamma(\Phi_F)$ for a strictly convex, (coordinate-wise) increasing, and twice differentiable function $\Gamma : \mathbb{R}^K \to \mathbb{R}$.[19,20]

5.1 C is increasing in risk

In this section, I assume that C is increasing in risk. I begin by characterizing the first-best distribution.

Proposition 5 *If C is increasing in risk, the degenerate distribution $F^{fb} = \delta_{x^{fb}}$ where x^{fb} maximizes $x - C(\delta_x)$ is a first-best distribution. Moreover, the first-best is implementable if and only if $x^{fb} - \Pi_E(\delta_{x^{fb}}) - K \geq 0$.*

The first-best problem can be solved in two steps. (A similar approach will be used to solve for the second-best below.) First, one solves for the optimal distribution among the set of distributions

[18] Costs that are increasing in risk capture mature firms whose potential for large innovations is relatively small, whereas costs that are decreasing in risk capture young, high-growth firms for whom a breakthrough innovation is easier to achieve than reliable returns.

[19] Setting $\varphi_k(\underline{x}) = 0$ is without loss. Because $C(\delta_{\underline{x}}) = 0$ means that $\Gamma(\varphi_1(\underline{x}), \ldots, \varphi_K(\underline{x})) = 0$, one obtains an equivalent model with functions $\tilde{\Gamma}$ and $\tilde{\varphi}_k$ and with the property that $\tilde{\varphi}_k(\underline{x}) = 0$ by setting: $\tilde{\varphi}_k(x) = \varphi_k(x) - \varphi(\underline{x})$ and $\tilde{\Gamma}(\Phi_1, \ldots, \Phi_K) = \Gamma(\Phi_1 + \varphi_1(\underline{x}), \ldots, \Phi_K + \varphi_K(\underline{x}))$.

[20] Convexity of Γ ensures convexity of C. Assuming strict convexity and differentiability simplifies some arguments. Note that when Γ is linear, costs C are linear so that Proposition 3 applies.

with a given mean $\int x\, dF = M \in [\underline{x}, \overline{x}]$:

$$P_M^{fb}: \qquad \max_{F:\int x\, dF=M} \int x\, dF - C(F). \tag{8}$$

Because the mean of the distribution is fixed and C is increasing in risk, the distribution δ_M which is the minimally risky distribution among all distributions with mean M is a solution. In the second step, one maximizes over M to find the first-best mean: $\max_{M\in X} M - C(\delta_M)$. Together with the implementability condition in part 1. of Proposition 2, the statement follows.[21]

Next, I characterize the second-best. To do so, I proceed as above and first look for the optimal distribution among the set of distributions with a given mean:

$$R_M: \qquad \max_{F:\int x\, dF=M} \int x\, dF - C(F) - K \quad s.t. \quad \int x\, dF - \Pi_E(F) - K \geq 0. \tag{9}$$

For fixed M, decreasing the risk of the distribution now not only increases the objective, but also affects the constraint via $\Pi_E(F)$. This effect is in general not clear-cut. But if $\Pi_E(F)$ is decreasing in risk, then lowering risk relaxes the constraint and increases the objective, and consequently a degenerate distribution is optimal. In other words, if both C and $\Pi_E(F)$ are increasing in risk, the degenerate distribution δ_M solves R_M. Maximizing over M then yields:

Proposition 6 *If C and Π_E are increasing in risk, then a solution to the relaxed problem R is the degenerate distribution δ_{x^*} where x^* maximizes*

$$x - C(\delta_x) \quad s.t. \quad x - \Pi_E(\delta_x) - K \geq 0. \tag{10}$$

Moreover, the solution to R is also a solution to the original problem P.

The fact that the solution to R solves P can be easily verified. Next, I show that Π_E is increasing in risk if C is moment-based with a supermodular function Γ. In this case, the first- and the second-best can also be ranked.

Proposition 7 *Let C be increasing and moment-based, and let Γ be supermodular. Then Π_E is increasing in risk. Moreover, $x^* \leq x^{fb}$.*

[21] In the special case that C is mean-based and can be written as $C(F) = \Gamma(\int x\, dF)$, the first-best objective depends only on the mean. Therefore, any distribution with mean $M^{fb} \in \arg\max_M M - \Gamma(M)$ is a first-best distribution.

To see that Π_E is increasing in risk, I use the well-known fact that a Gateaux-differerentiable function is increasing in risk if and only if its Gateaux derivative is convex in x for all F.[22] To illustrate the argument, suppose that C depends on one moment only, that is $K = 1$ (supermodularity is then vacuous). Then its Gateaux derivative is $c_F(x) = \Gamma'(\Phi_F)\varphi(x)$, and thus convex in x for all F if φ is convex. Moreover, if C is moment-based, then

$$\Pi_E(F) = \int c_F(x)\, d(F - \delta_{\underline{x}}) = \Gamma'(\Phi_F)\Phi_F, \tag{11}$$

and the Gateaux derivative of Π_E is

$$(\Pi_E)_F(x) = [\Gamma''(\Phi_F)\Phi_F + \Gamma'(\Phi_F)]\varphi(x).$$

Since Γ is increasing and convex, the term in the square bracket is positive so that convexity of $(\Pi_E)_F(x)$ for all F follows from convexity of φ.

The fact that $x^* \leq x^{fb}$ is driven by two features of moment-based costs. First, $C(\delta_x)$ is convex in x. Second, the maximizer of the constraint in problem R is smaller than the maximizer of the objective. These two features do not need to hold in general, however. For example, it does not need to be the case that $C(\delta_x)$ is convex in x even though C is convex on the whole domain of cdf's.

5.2 C is decreasing in risk

In this section, I assume that C is decreasing in risk. This is to some extent the mirror image of the case discussed in the previous section, and an important role will now be played by distributions that are maximally risky in the sense that they put the entire mass on the most extreme return realizations.

I denote by T_f the distribution whose support contains at most the boundary points $\underline{x}$ and $\overline{x}$ and which places mass $f \in [0,1]$ on $\overline{x}$. Define $\kappa(f) = C(T_f)$ and $\pi_E(f) = \Pi_E(T_f)$. Moreover, I impose throughout this section the assumption that implementing zero effort is not valuable: $\underline{x} < K$. I begin with the characterization of the first-best.

[22]This characterization is due to Cerreia-Vioglio et al. (2017). The research program to characterize the risk properties of non-expected utility functionals in terms of their "local utility functions" goes back to Machina (1982) who considered Frechet-differentiable functionals. See also Hong and Nishimura (1992) for a partial characterization for Gateaux-differentiable functionals.

Proposition 8 *If C is decreasing in risk, a first-best distribution is a distribution $T_{f^{fb}}$ where f^{fb} maximizes $(1-f)\underline{x}+f\overline{x}-\kappa(f)$. Moreover, the first-best $T_{f^{fb}}$ is implementable if and only if it places mass 1 on the point $x=\overline{x}$ (that is, $f^{fb}=1$) and*

$$\overline{x}-\pi_E(1)-K\geq 0. \tag{12}$$

The intuition can be seen again from considering subproblem P_M^{fb} where the mean of the distribution is fixed. Because C is now decreasing in risk, the distribution T_f which is the maximally risky distribution among all distributions with mean M is a solution. In the second step, one maximizes over f to find the first-best value f^{fb}.[23]

The first-best can be implemented only if it assigns probability 1 to the largest return realization $\overline{x}$. Otherwise, the investor's IR constraint can never be satisfied. To see this intuitively, recall from Lemma 2 and Figure 2 that if the first-best is implementable, then the security $S(x)$ that implements it is constant on the support. Therefore, because $S(x)\leq x$ by LL, it follows that if a first-best distribution puts positive probability on the lowest return realization $\underline{x}$, then the security pays at most $\underline{x}$ for all return realizations in the support. Because $\underline{x}<K$ by assumption, the investor cannot recoup his capital cost under a security that pays less than K.

Next, I characterize the second-best. Consider subproblem R_M where the mean of the distribution is fixed. Because costs are now decreasing in risk, increasing the risk of the distribution increases the objective. If, in addition, $\Pi_E(F)$ is decreasing in risk, then increasing the risk of the distribution also relaxes the constraint. Hence, the distribution T_f with mean M solves problem R_M. Maximizing over M then yields:

Proposition 9 *If C and Π_E are decreasing in risk, then the solution to the relaxed problem R is the two-point distribution T_{f^*} where f^* maximizes*

$$(1-f)\underline{x}+f\overline{x}-\kappa(f) \quad s.t. \quad (1-f)\underline{x}+f\overline{x}-\pi_E(f)-K\geq 0. \tag{13}$$

Moreover, the solution to R is also a solution to the original problem P, and it holds: $f^*\leq f^{fb}$.

The fact that the solution to R solves P can be easily verified. Analogously to Proposition 7, Π_E is decreasing in risk if C is moment-based with a supermodular function Γ. The result by

[23]As above, if C is mean-based any distribution with mean $M^{fb}\in\arg\max_M M-\Gamma(M)$ is a first-best distribution.

Cerreia-Vioglio (2017) quoted above now says that a Gateaux-differerentiable function is decreasing in risk if its Gateaux-derivative is concave in x for all F. Therefore, the arguments to show Proposition 7 equally apply.

5.3 Moment-based costs

In this section, I assume that costs are moment-based, but not necessarily increasing or decreasing in risk.

I begin with the case that costs depend on a single moment. I say that C is single-moment-based if $C(F) = \Gamma(\Phi_F)$ with $\Phi_F = \int \varphi(x)\, dF$ for a strictly convex, increasing and twice differentiable function $\Gamma : \mathbb{R} \to \mathbb{R}$, and an increasing function $\varphi : X \to \mathbb{R}$. Let

$$\breve{\varphi}(x) = \sup\{h : X \to \mathbb{R} \mid h \text{ convex}, h(x) \leq \varphi(x) \text{ for all } x \in X\}$$

be the lower convex envelope of φ. I begin by characterizing the first-best.

Proposition 10 *Let C be single-moment-based, then there is a first-best distribution with at most two points in its support. The first-best value V^{fb} is the value of the maximization problem* $\max_{M\in[\underline{x},\overline{x}]} M - \Gamma(\breve{\varphi}(M))$*, and the (unique) maximizer is the first-best mean $M^{fb} = \int x\, dF^{fb}$.*

To see the result, consider first problem P_M^{fb} which now writes:

$$\max_{F:\int x\, dF=M} \int x\, dF - \Gamma(\Phi_F).$$

Because the mean is fixed, and Γ is increasing, the solution is the distribution with mean M that minimizes the moment: $\min_{F:\int x\, dF=M} \Phi_F$.

This is a linear problem with one linear constraint, and it is well-known that there is a solution to this problem that has at most two points in its support (see Winkler, 1988). Moreover, standard convexification arguments imply that the value of this problem is $\min_{F:\int x\, dF=M} \Phi_F = \breve{\varphi}(M)$.[24] In the second step, one maximizes over the possible means M to find the first-best value: $\max_M M - \Gamma(\breve{\varphi}(M))$.

[24] I am grateful to Siwen Liu for pointing this out to me.

Next, I characterize the second-best. Recall from (11) that with single-moment-based costs, also the entrepreneur's minimal payout is moment-based:

$$\Pi_E(F) = \Gamma'(\Phi_F)\Phi_F.$$

Problem R_M thus writes:

$$R_M: \quad \max_{F:\int x\, dF = M} \int x\, dF - \Gamma(\Phi_F) \quad s.t. \quad \int x\, dF - \Gamma'(\Phi_F)\Phi_F - K \geq 0.$$

Due to convexity of Γ, the function $\Gamma'(\Phi_F)\Phi_F$ is increasing in Φ_F. Therefore, for given mean M, decreasing the moment Φ_F both increases the objective and relaxes the constraint. Hence, the solution is the distribution with mean M that minimizes the moment. As discussed in the previous paragraph, there is a solution with at most two points in its support, and $\breve{\varphi}(M)$ minimizes the moment, given M. I conclude:

Proposition 11 *Let C be single-moment-based.*

1. *Then there is a solution F^* to problem R with at most two points in its support, and the value of R is the value of the maximization problem*

$$\max_M M - \Gamma(\breve{\varphi}(M)) \quad s.t. \quad M - \Gamma'(\breve{\varphi}(M))\breve{\varphi}(M) - K \geq 0. \tag{14}$$

   and its maximizer M^* is the mean of F^*.

2. *Moreover, $M^* \leq M^{fb}$, and there are a first-best distribution F^{fb} and a solution F^* to R, each with at most two points in their support, such that F^{fb} first-order stochastically dominates F^*.*

3. *If F^* is a degenerate distribution, then it also solves the original problem P. Otherwise, F^* solves the original problem if and only if for both points x_1, x_2 in the support, it holds: $x_i - \Gamma'(\breve{\varphi}(M^*)\varphi(x_i) \geq 0$. A sufficient primitive condition for this is that the slope of φ is bounded by $1/\Gamma'(\varphi(\overline{x}))$.*

That $M^* \leq M^{fb}$ can be shown as in Proposition 7 because due to $\breve{\varphi}$ being convex, problem (14) is effectively the same as when costs are increasing in risk and moment-based.

The second statement in part 2. can best be seen for the case that F^{fb} and F^* are unique. Recall that F^{fb} and F^* are obtained by convexifying φ. The interval $[\underline{x}, \overline{x}]$ can be divided into disjoint segments in each of which the convexification $\breve{\varphi}$ corresponds to an affine function (at a point where $\breve{\varphi}$ is strictly convex, the segment is just this point). The support of F^i consists of the lower and upper points of the segment in which M^i is located, $i \in \{fb, *\}$. Now, because $M^* \leq M^{fb}$, there are only two possibilities. Either M^* and M^{fb} are located in the same segment. In this case, F^{fb} and F^* have the same support. Or, M^* is located in a "lower" segment than M^{fb}. In this case, all points in the support of F^* are smaller than all points in the support of F^{fb}. In either case, F^{fb} first order stochastically F^*. (Part 3. is technical and shows that the constraints that were neglected in the relaxed problem are met under the stated conditions.)

Next, I turn to (general) moment-based cost functions. The next proposition characterizes a second-best solution.

Proposition 12 *Let costs be moment-based. Then there is a discrete distribution F^* with at most $K+1$ mass points that is a solution to R. F^* is also a solution to the original problem if and only if for all points $x_1, \ldots, x_{K+1}$ in the support, it holds: $x_k - \Gamma'(\Phi_{F^*}) \cdot \varphi(x_k) \geq 0$.*

The result rests on the fact that the Gateaux derivative $c_F(x) = \Gamma'(\Phi_F) \cdot \varphi(x)$ is the dot-product of the gradient Γ' of Γ and the vector $\varphi(x)$. Therefore, also the entrepreneur's minimal payout is moment-based,

$$\Pi_E(F) = \int c_F(x)\, d(F - \delta_{\underline{x}}) = \Gamma'(\Phi_F) \cdot \Phi_F,$$

and problem R writes:

$$\max_F \int x\, dF - \Gamma'(\Phi_F) \quad s.t. \quad \int x\, dF - \Gamma'(\Phi_F) \cdot \Phi_F - K \geq 0.$$

For given $\Phi_k \in [\varphi_k(\underline{x}), \varphi_k(\overline{x})]$, consider now the subproblem where the moments are fixed: $\Phi_{k,F} = \Phi_k, k = 1, \ldots, K$. Both the objective as well as the constraint then depend only on the mean $\int x\, dF$ of the distribution, and both are increasing in the mean. Therefore, the problem is equivalent to maximizing the mean subject to the K-many moments constraints $\Phi_{k,F} = \Phi_k, k = 1, \ldots, K$. Because this is a linear problem with K-many linear constraints, there is a solution in the set of extreme points of the constraint set. It is well known (Winkler, 1988) that the set of extreme points is the set of discrete distributions with at most $K + 1$-many mass points.

The comparison of the second-best solution with the first-best is difficult. While similar arguments as in the previous paragraph imply that there is a discrete first-best distribution with finitely many mass points, an argument in the spirit of the convexification argument that I used in the single-moment case is not available here. I leave this question for future research.

6 Conclusion

In this paper, I have applied the smooth flexible moral hazard approach by Georgiadis et al. (2024) to a security design problem in the spirit of Innes (1990). I characterize when the first-best is implementable and derive second-best solutions for various cost structures. I show that the agency costs are driven by the convexity of the cost function, that optimal distributions are discrete and finite, and that second-best solutions are first-order stochastically dominated by the first-best and, in this sense, display inefficiently little risk.

A prominent question in the literature is whether the theory can explain real world securities such as debt or equity as the outcome of optimal design. Such predictions are difficult to make within the flexible moral hazard framework presented in this paper. The reason is that optimal securities are only pinned down on the support of optimal distributions, but leave many degrees of freedom off the support. The optimal distributions I identify have finite support within an interval of possible returns, and can thus be specified fairly arbitrarily for almost all returns. In other words, even if a standard security is optimal, many other securities are optimal, too. This is different in the approaches by Innes (1990) or Hébert (2018) where optimal distributions have full support either by assumption or because non-full support distributions are infinitely costly. Thus, similar to Hellwig (2009), my paper offers a note of caution about the predictive value of the framework to explain real world securities.

In this paper, I focus on moral hazard. A large literature in security design considers settings with adverse selection where either the entrepreneur or the investor has private information (for a recent contribution, see Gershkov et al. (2025) and the references therein). While most work in the literature considers adverse selection and moral hazard in isolation, it is an interesting and relevant avenue for future research to incorporate both. Recent work in other contexts (Krähmer, 2025) demonstrates that the flexible moral hazard approach is sufficiently tractable to do so.

7 Proofs

Proof of Proposition 1 By Georgiadis et al. (2024), and ss explained in the main text, the moral hazard constraint (MH) is equivalent to (MH_{supp}) and (MH_{all}).

Now, let F be implementable. The left inequality of (LL') follows from (MH_{all}) and the left inequality $0 \leq x$ of (LL). The right inequality of (LL') follows from (MH_{supp}) and the right inequality $x \leq S(x)$ of LL. Finally, (IR') follows from (MH_{supp}) and IR.

Conversely, let (LL') and (IR') hold, define the security $S(x) = x - c_F(x) - \lambda$ for all $x \in supp(F)$, and $S(x) = x$ for all $x \notin supp(F)$. Then S satisfies (MH_{supp}) by construction. Condition (MH_{all}) holds because $\lambda \geq \underline{\lambda}_F$ by (LL'). Moreover, S trivially satisfies LL for $x \notin supp(F)$, and for $x \in supp(F)$ because $\lambda \leq \overline{\lambda}_F$ by (LL'). Finally, S satisfies IR by (IR'). QED

Proof of Lemma 1 The claim follows from Georgiadis et al. (2024). QED

Proof of Proposition 2 As to 1. Let F^{fb} be a first-best distribution. By (FB_{supp}), it follows that

$$\overline{\lambda}_{F^{fb}} = \lambda^{fb}, \quad \lambda^{IR}_{F^{fb}} = \lambda^{fb} - K, \tag{15}$$

so that $\lambda^{IR}_{F^{fb}} < \overline{\lambda}_{F^{fb}}$. By Proposition 1, F^{fb} is therefore implementable if and only if $\underline{\lambda}_{F^{fb}} \leq \lambda^{IR}_{F^{fb}}$. Moreover, using the definition in (1), it is straightforward to verify that this inequality is equivalent to $\Pi_I(F^{fb}) \geq K$.

As to 2. Let F^{fb} be implementable. By part 1. thus implies that $\underline{\lambda}_{F^{fb}} \leq \lambda^{IR}_{F^{fb}}$. Accordingly, F^{fb} can be implemented by choosing any $\lambda \in [\underline{\lambda}_{F^{fb}}, \lambda^{IR}_{F^{fb}}]$. For the particular choice

$$\lambda = \lambda^{IR}_{F^{fb}} = \int x - c_{F^{fb}}(x)\, dF^{fb} - K = \Pi_I(F^{fb}) - K - c_{F^{fb}}(x) \tag{16}$$

the entrepreneur's expected payoff is

$$\int c_{F^{fb}}(x)\, dF^{fb} - C(F^{fb}) + \lambda^{IR}_{F^{fb}} = \int x\, dF^{fb} - C(F^{fb}) - K. \tag{17}$$

Thus, the entrepreneur gets the first-best project value V^{fb}, and she can clearly not do better than implementing F^{fb} in this way. QED

Proof of Proposition 3 If C is linear, the Gateaux-derivative is $c_F(x) = \varphi(x)$. Moreover, $C(\delta_{\underline{x}}) =$

$\varphi(\underline{x}) = 0$ by assumption. Therefore,

$$\underline{\lambda}_{F^{fb}} = -\varphi(\underline{x}) = 0, \quad \lambda^{IR}_{F^{fb}} = \int x \, dF - C(F^{fb}) - K. \tag{18}$$

Hence, since $\lambda^{IR}_{F^{fb}}$ is equal to the first-best project value, it is positive by assumption. Therefore, F^{fb} is implementable by part 1. of Proposition 2. QED

Proof of Proposition 4 Let $\kappa(F) = \tilde{C}(F) - C(F)$. Then

$$\widetilde{\Pi}^{net}_E(F) - \Pi^{net}_E(F) = \int \kappa_F(x) d(F - \delta_{\underline{x}}) - \kappa(F). \tag{19}$$

Since κ is convex by assumption, it follows from Proposition 2 in Krähmer (2026) that the expression on the right-hand side is positive. QED

Proof of Proposition 5 The argument is in the main text. QED

Proof of Proposition 6 All arguments are in the main text except for the claim that the solution to R is a solution to the original problem P. To see this claim, recall that when deriving problem R, λ was optimally set equal to λ^{IR}_F. Because in the relaxed problem, the constraint $\lambda \leq \overline{\lambda}_F$, is dropped it follows that a solution F^* to the relaxed problem is also a solution to the original problem if and only if $\lambda^{IR}_{F^*} \leq \overline{\lambda}_{F^*}$, that is,

$$\int x - c_{F^*}(x) \, dF^* - K \leq \min_{x \in supp(F^*)} (x - c_{F^*}(x)). \tag{20}$$

Now observe that for the degenerate distribution $F^* = \delta_{x^*}$, condition (20) writes

$$x^* - c_{F^*}(x^*) - K \leq x^* - c_{F^*}(x^*), \tag{21}$$

and is thus satisfied. QED

Proof of Proposition 7 I first show that Π_E is increasing in risk. As argued in the main text, to do so it is sufficient to show that Π_E has a convex Gateaux derivative. To see this, note first that since C is moment-based, its Gateaux derivative

$$c_F(x) = \sum_k \frac{\partial \Gamma(\Phi_F)}{\partial \Phi_k} \varphi_k(x) \tag{22}$$

is the dot-product of the gradient of Γ and the vector $\varphi(x) = (\varphi_1(x), \dots, \varphi_K(x))$. Note that because C is increasing in risk, its Gateaux derivative $c_F(x)$ is convex in x. Together with the fact that Γ is coordinate-wise monotone, this implies that $\varphi_k(x)$ is convex in x for all k.

Plugging c_F into the expression for Π_E yields (recall that $\varphi_k(\underline{x}) = 0$):

$$\Pi_E(F) = \int c_F(x)\, d(F - \delta_{\underline{x}}) = \sum_k \frac{\partial \Gamma(\Phi_F)}{\partial \Phi_k} \Phi_{k,F}(x). \tag{23}$$

Thus, by the chain rule, the Gateaux derivative of Π_E is:

$$\begin{aligned} (\Pi_E)_F(x) &= \sum_{k,\ell} \frac{\partial^2 \Gamma(\Phi_F)}{\partial \Phi_\ell \partial \Phi_k} \varphi_\ell(x) \Phi_{k,F} + \sum_k \frac{\partial \Gamma(\Phi_F)}{\partial \Phi_k} \varphi_k(x) &(24)\\ &= \sum_\ell \left[\sum_k \frac{\partial^2 \Gamma(\Phi_F)}{\partial \Phi_\ell \partial \Phi_k} \Phi_{k,F} + \frac{\partial \Gamma(\Phi_F)}{\partial \Phi_\ell} \right] \varphi_\ell(x). &(25) \end{aligned}$$

To see that this is convex in x, observe that all terms in the square brackets are positive because Γ is convex, supermodular, and coordinate-wise increasing, and that φ_ℓ is convex as mentioned above. This shows that Π_E is increasing in risk.

It remains to show that $x^* \le x^{fb}$. This is trivial if $x^{fb} = \overline{x}$. Hence, let $x^{fb} < \overline{x}$ be interior. Denote by $\alpha(x) = x - C(\delta_x)$ the objective function and by $\beta(x) = x - \Pi_E(\delta_x) - K$ the constraint in problem (10). Since costs are moment-based,

$$\alpha(x) = x - \Gamma(\varphi_1(x), \dots, \varphi_K(x)), \quad \beta(x) = x - \sum_k \frac{\partial \Gamma(\varphi_1(x), \dots, \varphi_K(x))}{\partial \Phi_k} \varphi_k(x) - K. \tag{26}$$

(The expression for β follows from plugging δ_x in (23).) Note that because Γ is strictly convex and φ_k is convex for all k (as shown in the first part of the proof), the first-best value x^{fb} uniquely maximizes α, and it follows that $\alpha'(x) < 0$ for all $x > x^{fb}$. Moreover, note that

$$\beta'(x) = \alpha'(x) - \sum_{k,\ell} \frac{\partial^2 \Gamma(\varphi_1(x), \dots, \varphi_K(x))}{\partial \Phi_\ell \partial \Phi_k} \varphi_k(x) \varphi_\ell'(x) \tag{27}$$

Thus, $\alpha'(x) \geq \beta'(x)$, since Γ is convex and supermodular, and φ_ℓ is increasing. Therefore, it follows that also $\beta'(x) < 0$ for all $x > x^{fb}$. Now, suppose, contrary to the claim, that $x^{fb} < x^*$. Then lowering x^* would increase the objective and relax the constraint, a contradiction to the optimality of x^*. QED

Proof of Proposition 8 The argument for why a two-point distribution T_f is optimal, is given in the main text. The problem stated in the proposition is the first-best problem in the class of these distributions.

For the rest of the proof, I use the following property of cost functions that are decreasing in risk:

$$\kappa'(f) = c_{T_f}(\overline{x}) - c_{T_f}(\underline{x}), \quad and \quad \pi_E(f) = f\kappa'(f). \tag{28}$$

To see this, note that $T_{f+\epsilon} = T_f + \frac{\epsilon}{1-f}(T_1 - T_f)$, which implies

$$\begin{aligned}
\kappa'(f) &= \frac{dC(T_f)}{df} && (29)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[C(T_{f+\epsilon}) - C(T_f)\right] && (30)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[C(T_f + \frac{\epsilon}{1-f}(T_1 - T_f)) - C(T_f)\right] && (31)\\
&= \frac{1}{1-f}\lim_{\epsilon\to 0}\frac{1}{\frac{\epsilon}{1-f}}\left[C(T_f + \frac{\epsilon}{1-f}(T_1 - T_f)) - C(T_f)\right] && (32)\\
&= \frac{1}{1-f}\int c_{T_f}(x)\, d(T_1 - T_f) && (33)\\
&= c_{T_f}(\overline{x}) - c_{T_f}(\underline{x}). && (34)
\end{aligned}$$

Moreover,

$$\pi_E(f) = \Pi_E(T_f) = \int c_{T_f}(x)\, dT_f - c_{T_f}(\underline{x}) = f[c_{T_f}(\overline{x}) - c_{T_f}(\underline{x})] = f\kappa'(f), \tag{35}$$

as desired.

I next prove the “only if”-part of the implementability statement. Let $T_{f^{fb}}$ be implementable. I show first that $f^{fb} = 1$. To the contrary, assume $T_{f^{fb}}$ had positive mass $1 - f^{fb} > 0$ on $x = \underline{x}$. I derive a contradiction to the first-best implementabiltiy condition $\Pi_I(T_{f^{fb}}) - K \geq 0$ stated in

Proposition 2. Indeed, if there is positive mass on $\underline{x}$, then (FB_{supp}) and (FB_{all}) imply

$$\underline{x} - c_{T_{f^{fb}}}(\underline{x}) = \lambda^{fb} \quad and \quad \overline{x} - c_{T_{f^{fb}}}(\overline{x}) \leq \lambda^{fb}. \tag{36}$$

Thus,

$$\begin{aligned}
\Pi_I(T_{f^{fb}}) - K &= \int x - c_{T_{f^{fb}}}(x)\, dT_{f^{fb}} + c_{T_{f^{fb}}}(\underline{x}) - K & (37)\\
&= (1-f^{fb})(\underline{x} - c_{T_{f^{fb}}}(\underline{x})) + f^{fb}(\overline{x} - c_{T_{f^{fb}}}(\overline{x})) + c_{T_{f^{fb}}}(\underline{x}) - K & (38)\\
&\leq \lambda^{fb} + c_{T_{f^{fb}}}(\underline{x})) - K & (39)\\
&= \underline{x} - K. & (40)
\end{aligned}$$

Because $\underline{x} - K < 0$ by assumption, this is the desired contradiction.

To complete the proof of the "only if"-part, I show that $\overline{x} - \pi_E(1) - K \geq 0$. Indeed, because $T_{f^{fb}}$ is implementable, the inequality $\Pi_I(T_{f^{fb}}) - K \geq 0$ holds by Proposition 2. Since $f^{fb} = 1$ and because $\pi_E(f) = f[c_{T_f}(\overline{x}) - c_{T_f}(\underline{x})]$ by (28), the said inequality simplifies to $\overline{x} - \pi_E(1) - K \geq 0$, as desired.

To see the "if"-part, let $f^{fb} = 1$ and $\overline{x} - \pi_E(1) - K \geq 0$. I show that $T_{f^{fb}} = T_1$ is implementable by verifying the condition $\Pi_I(T_1) - K \geq 0$ from Proposition 2. Indeed, since $\pi_E(f) = f[c_{T_f}(\overline{x}) - c_{T_f}(\underline{x})]$ by (28), one has

$$\begin{aligned}
\Pi_I(T_1) - K &= \int x - c_{T_1}(x)\, dT_1 + c_{T_1}(\underline{x}) - K & (41)\\
&= \overline{x} - c_{T_1}(\overline{x})) + c_{T_1}(\underline{x}) - K & (42)\\
&= \overline{x} - \pi_E(1) - K, & (43)
\end{aligned}$$

which is positive by assumption. Thus, $T_{f^{fb}} = T_1$ is implementable by Proposition 2. QED

Proof of Proposition 9 All arguments are in the main text except for the claims that (a) the solution to R is a solution to the original problem P, and that (b) $f^* \leq f^{fb}$.

As to (a). With the arguments in the proof of Proposition 6, it follows that the solution F^* to the relaxed problem is also a solution to the original problem if and only if

$$\int x - c_{F^*}(x)\, dF^* - K \leq \min_{x \in supp(F^*)} (x - c_{F^*}(x)). \tag{44}$$

For the two-point distribution $F^* = T_{f^*}$, condition (44) writes

$$(1-f^*)[\underline{x}-c_{T_{f^*}}(\underline{x})]+f^*[\overline{x}-c_{T_{f^*}}(\overline{x})]-K \leq \min\{\underline{x}-c_{T_{f^*}}(\underline{x}),\ \overline{x}-c_{T_{f^*}}(\overline{x})\}. \tag{45}$$

I now distinguish two cases. First, suppose T_{f^*} coincides with a first-best distribution. Recall that in this case, $x-c_{T_{f^*}}(x)$ is maximized—and in particular takes on the same value—on the support of T_{f^*} by (FB_{supp}) and (FB_{all}). Thus, (45) is equivalent to $-K \leq 0$ and therefore holds.

Second, suppose T_{f^*} is not a first-best distribution. Then the constraint in (13) is binding, and, by (28), reads

$$(1-f^*)\underline{x}+f^*\overline{x}-f^*[c_{T_{f^*}}(\overline{x})-c_{T_{f^*}}(\underline{x})]-K=0. \tag{46}$$

Therefore, the left hand side of (45) is equal to $-c_{T_{f^*}}(\underline{x})$. To determine the right hand side, (46) can be re-arranged to

$$\underline{x}-c_{T_{f^*}}(\underline{x}))+\frac{K-\underline{x}}{f}=\overline{x}-c_{T_{f^*}}(\overline{x}). \tag{47}$$

Since $\underline{x}<K$ by assumption, this implies that

$$\min\{\underline{x}-c_{T_{f^*}}(\underline{x})),\overline{x}-c_{T_{f^*}}(\overline{x})\}=\underline{x}-c_{T_{f^*}}(\underline{x})). \tag{48}$$

Therefore, the right hand side of (45) is equal to $\underline{x}-c_{T_{f^*}}(\underline{x}))$. Because the right hand side is equal to $-c_{T_{f^*}}(\underline{x}))$, and $0 \leq \overline{x}$, (45) is shown, and this establishes claim (a).

As To (b). To see that $f^* \leq f^{fb}$, I show below that κ is convex and that $\kappa'(f) \leq \pi_E'(f)$ for all f. The claim then follows from the same formal arguments that establish that $x^* \leq x^{fb}$ in the proof of Proposition 7, where now $\alpha(f)=(1-f)\underline{x}+f\overline{x}-\kappa(f)$, and $\beta(f)=(1-f)\underline{x}+f\overline{x}-\pi_E(f)-K$.

To see that κ is convex, recall that $\kappa(f)=C(T_f)$. Now, observe that for a convex combination $\tau f+(1-\tau)g$, $\tau \in [0,1]$, it holds:

$$T_{\tau f+(1-\tau)g}=\tau T_f+(1-\tau)T_g. \tag{49}$$

Therefore, the convexity of κ follows from the convexity of C. To see that $\kappa'(f) \leq \pi_E'(f)$ for all

f, recall that $\pi_E(f) = f\kappa'(f)$, and thus

$$\pi_E'(f) = \kappa'(f) + \kappa''(f)f \geq \kappa'(f), \tag{50}$$

as desired. QED

Proof of Proposition 10 The argument is in the main text. QED

Proof of Proposition 11 The arguments for part 1. of the statement are in the main text.

As to part 2. note first that problem 14 is structurally identical to problem (10) for moment-based costs, because $\breve{\varphi}$ is convex by definition. The proof that $M^* \leq M^{fb}$ is thus identical to the proof that $x^* \leq x^{fb}$ in Proposition 7.

To see the second statement of part 2. recall that Proposition 10 and part 1. of the statement imply that there is a first-distribution $\tilde{F}^{fb}$ a solution $\tilde{F}^*$ to R with support points $\tilde{x}_1^{fb} \leq M^{fb} \leq \tilde{x}_2^{fb}$ and $\tilde{x}_1^* \leq M^* \leq \tilde{x}_2^*$.

Observe that if $\tilde{x}_2^* \leq \tilde{x}_1^{fb}$ then all points in the support $\tilde{F}^{fb}$ are larger than all points in the support $\tilde{F}^*$, and hence $\tilde{F}^{fb}$ evidently first order stochastically dominates $\tilde{F}^*$. Therefore, I from now on assume that $\tilde{x}_2^* > \tilde{x}_1^{fb}$. I distinguish the following cases.

Case 1: $\tilde{x}_2^{fb} \leq \tilde{x}_2^*$. I show that there is a solution F^* to R with $x_1^* = \tilde{x}_1^*$ and $x_2^* = \tilde{x}_2^{fb}$. Because $M^* \leq M^{fb}$, this implies that $\tilde{F}^{fb}$ first order stochastically dominates F^*.

To see the claim, note that because $\tilde{x}_1^* \leq M^* \leq M^{fb} \leq \tilde{x}_2^{fb}$, there is a convex combination of $\tilde{x}_1^*$ and $\tilde{x}_2^{fb}$ that is equal to M^*. Therefore there is a two-point distribution with support $\{\tilde{x}_1^*, \tilde{x}_2^{fb}\}$ and mean M^*. Let this distribution be F^*. Moreover, because $\tilde{x}_2^{fb} \leq \tilde{x}_2^*$ by assumption, F^* is a mean preserving contraction (MPC) of F^{fb}. I show now that F^* is indeed a solution to R. To do so, recall from the discussion preceding the statement of the proposition that $\tilde{F}^i$ minimizes $\Phi_F = \int \varphi \; dF$ subject to the mean of F being M^i, $i \in \{fb, *\}$. By standard convexification arguments, this also implies that φ coincides with its lower convex envelope on the respective supports. In particular,

$$\varphi(\tilde{x}_1^*) = \breve{\varphi}(\tilde{x}_1^*), \quad \varphi(\tilde{x}_2^{fb}) = \breve{\varphi}(\tilde{x}_2^{fb}), \quad \varphi(\tilde{x}_2^*) = \breve{\varphi}(\tilde{x}_2^*). \tag{51}$$

Therefore,

$$\Phi_{F^*} = \int \varphi \; dF^* = \int \breve{\varphi} \; dF^* \leq \int \breve{\varphi} \; d\tilde{F}^* = \int \varphi \; d\tilde{F}^* = \Phi_{\tilde{F}^*} = \min_{\int x \; dF = M^*} \Phi_F, \tag{52}$$

where the inequality follows from the facts that F^* is an MPC of $\tilde{F}^*$ and $\breve{\varphi}$ is convex. But this implies that F^* is (also) a solution to R, as desired.

Case 2: $\tilde{x}_2^{fb} > \tilde{x}_2^*$.

Case 2(a): $M^{fb} \leq \tilde{x}_2^*$: I show that there is a first-best distribution F^{fb} that first order stochastically dominates F^*. To see this, let F^{fb} be the two-point distribution with support $\{\tilde{x}^{fb}, \tilde{x}_2^*\}$ and mean M^{fb} (because $M^{fb} \leq \tilde{x}_2^*$, probabilities can be assigned so that the mean is indeed M^{fb}. Note that F^{fb} first order stochastically dominates $\tilde{F}^*$ and is an MPC of $\tilde{F}^{fb}$. Analogous steps as in Case 1 can now be used to show that F^{fb} is a first-best distribution. Therefore, F^{fb} and $\tilde{F}^*$ satisfy the properties stated in the proposition.

Case 2(b): $M^{fb} > \tilde{x}_2^*$: I show that there is a first-best distribution F^{fb} that first order stochastically dominates F^*. To see this, let F^{fb} be the two-point distribution with support $\{\tilde{x}_2^*, \tilde{x}^{(2)},\}$ and mean M^{fb} (because $M^{fb} > \tilde{x}_2^*$, probabilities can be assigned so that the mean is indeed M^{fb}. Note that F^{fb} first order stochastically dominates $\tilde{F}^*$ and is an MPC of $\tilde{F}^{fb}$. (To see the latter, recall that $x_1^{fb} < \tilde{x}_2^*$ by assumption.) Analogous steps as in Case 1 can now be used to show that F^{fb} is a first-best distribution. Therefore, F^{fb} and F^* satisfy the properties stated in the proposition. QED

As to 3. I show that the solution to R is a solution to the original problem P under the stated sufficient conditions. Indeed, the same arguments as in the proof of Proposition 6 apply that show that a solution F^* to the relaxed problem is also a solution to the original problem if and only if

$$\int x - c_{F^*}(x)\, dF^* - K \leq \min_{x \in supp(F^*)} (x - c_{F^*}(x)). \tag{53}$$

If F^* is a degenerate distribution, the inequality is evidently always true. For the case that F^* coincides with a first-best distribution, recall that $x - c_{F^*}(x)$ is maximized—and in particular takes on the same value—on the support of F^* by (FB_{supp}) and (FB_{all}). Thus, (53) is equivalent to $-K \leq 0$ and therefore holds.

Thus, suppose that F^* has two points $\{x_1, x_2\}$ in its support with $f^* = Prob(x_2)$, and is not first-best. Condition (53) writes

$$\begin{aligned} & (1-f^*)[x_1 - \Gamma'(\Phi_{F^*})\varphi(x_1)] + f^*[x_2 - \Gamma'(\Phi_{F^*})\varphi(x_2)] - K \\ \leq \quad & \min\{x_1 - \Gamma'(\Phi_{F^*})\varphi(x_1),\ x_2 - \Gamma'(\Phi_{F^*})\varphi(x_2)\}, \end{aligned} \tag{54}$$

where I have used that for moment-based costs, $c_F(x) = \Gamma'(\Phi_F)\varphi(x)$. and $\Phi_{F^*} = \breve{\varphi}(M^*)$ at an optimum.

Because F^* is not first-best, the constraint in (14) is binding:

$$M^* - \Gamma'(\Phi_{F^*})\Phi_{F^*} - K = \int x - \Gamma'(\Phi_{F^*})\varphi(x)\, dF^* - K = 0. \tag{55}$$

Therefore, the left hand side of inequality (54) is equal to zero. Accordingly, inequality (54) holds if and only

$$x_i - \Gamma'(\breve{\varphi}(M^*)\varphi(x_i) \geq 0, \quad i = 1, 2. \tag{56}$$

To complete the proof, I show that a sufficient condition for (56) is

$$\varphi'(x) \leq \frac{1}{\Gamma'(\varphi(\overline{x}))} \quad \forall x \in X. \tag{57}$$

Indeed, recall that $\varphi(x) = 0$. Thus a sufficient condition for (56) is that $x - \Gamma'(\breve{\varphi}(M^*)\varphi(x)$ is increasing in x, that is,

$$1 - \Gamma'(\breve{\varphi}(M^*)\varphi'(x) \geq 0 \quad \forall x \in X. \tag{58}$$

This condition is implied by (57) because Γ' is increasing by assumption and because $\breve{\varphi}(M^*) \leq \varphi(M^*) \leq \varphi(\overline{x})$ by definition of the lower convex envelope and because φ is increasing by assumption. And this completes the proof. QED

Proof of Proposition 12 All arguments are in the main text except for when the solution to R is a solution to the original problem P. The argument for this is analogous to the corresponding argument in the proof of Proposition 11. QED

References

Allen, Franklin, and Adelina Barbalau (2024). Security design: A review. Journal of Financial Intermediation 60, 101113.

Biais, Bruno, and Catherine Casamatta (1999). Optimal leverage and aggregate investment. Journal of Finance 54.4, 1291-1323.

Castro-Pires, Henrique, Deniz Kattwinkel, and Jan Knöpfle. Who and how? Adverse selection with flexible moral hazard. Mimeo. 2025.

Cerreia-Vioglio, Simone, Fabio Maccheroni, and Massimo Marinacci (2017). Stochastic dominance analysis without the independence axiom. Management Science 63(4), 1097-1109.

Georgiadis, George, Doron Ravid, and Balazs Szentes (2024). Flexible moral hazard problems. Econometrica 92(2), 387-409.

Gershkov, Alex, Benny Moldovanu, Philipp Strack, and Mengxi Zhang (2025). Optimal security design for risk-averse investors. American Economic Review, 115 (6): 2050-92.

Grossman, Sanford J. and Oliver D. Hart (1983). An analysis of the principal-agent problem. Econometrica 51(1), 7-49.

Hébert, Benjamin (2018). Moral hazard and the optimality of debt. Review of Economic Studies 85(4), 2214-2252.

Hellwig, Martin F. (2009). A reconsideration of the Jensen-Meckling model of outside finance. Journal of Financial Intermediation 18(4), 495-525.

Hong, Chew Soo, and Naoko Nishimura. Differentiability, comparative statics, and non-expected utility preferences. Journal of Economic Theory 56.2 (1992): 294-312.

Innes, Robert D. (1990). Limited liability and incentive contracting with ex-ante action choices. Journal of Economic Theory 52(1), 45-67.

Jensen, Michael C. and William H. Meckling (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3, 305-360.

Kocherlakota, Narayana R. Optimal labor income taxation: a flexible moral hazard approach. Review of Economic Studies, forthcoming. (2025).

Krähmer, Daniel (2025). Optimal monopoly regulation with flexible investments, mimeo.

Krähmer, Daniel (2026). The hold-up problem with flexible unobservable investments, mimeo.

Liu, Siwen. Flexible Moral Hazard Problems with Adverse Selection. Mimeo. 2025.

Machina, Mark J. “Expected utility” analysis without the independence axiom. Econometrica (1982): 277-323.

Mattsson, Lars-Göran, and Jörgen W. Weibull. An analytically solvable principal-agent model. Games and Economic Behavior 140 (2023): 33-49.

Winkler, Gerhard (1988). Extreme points of moment sets. Mathematics of Operations Research 13(4), 581-587.

Yang, Kai Hao, and Alexander K. Zentefis (2024). Monotone function intervals: theory and applications. American Economic Review 114(8), 2239-2270.