Security Design with Flexible Moral Hazard and Limited Liability

Daniel Krähmer*

June 7, 2025

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 effort choice by the entrepreneur (moral hazard). Effort 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 effort cost functions that are increasing or decreasing in risk or depend only on moments of the distribution. Securities that implement optimal distributions are not unique, and I identify cases where both debt and equity is optimal.

Keywords: Security Design, Moral Hazard, Limited Liability, Flexible Effort

JEL: D21, D82, D86, G32

*Universität Bonn (kraehmer@hcm.uni-bonn.de); I thank 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).

I study security design in the spirit of Innes (1990). To receive funding for a project, a risk-neutral entrepreneur issues a security that promises to pay a risk-neutral investor a portion of the realized project return. There is two-sided limited liability: The entrepreneur's liability is limited to the project return, and the investor is not liable for losses incurred by the entrepreneur. The distribution of the project return depends on an unobservable effort choice by the entrepreneur, that is, there is moral hazard. Moreover, in order to provide capital for the project, the investor requires the expected security payout cover his cost of capital.

In Innes (1990) and much of the subsequent literature, the entrepreneur chooses among a set of one-dimensional effort parameters that each give rise to a stochastic return distribution. My paper contributes to the literature by assuming that, in contrast, the entrepreneur's effort is flexible: She can choose any probability distribution of returns subject to a cost. Such a security design problem with flexible effort choice is studied in Hébert (2018) who considers a setting where return distributions are discrete and finite, and more importantly, the effort cost function has the form of a divergence.[1] In contrast, I assume that the space of possible return realizations is a continuous interval and consider effort cost functions that are assumed only to be monotone, convex, and smooth. Monotonicity, which means that costs increase if the return distribution increases in the sense of first order stochastic dominance sense, and convexity, which captures increasing marginal costs, are natural assumptions in the effort context considered here. Smoothness means that the cost function admits a Gateaux-derivative[2], which makes the analysis amenable to the flexible moral hazard approach developed in Georgiadis et al. (2024).[3]

My approach is to 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).[4] This approach highlights the basic trade-off that underlies the security design prob-

[1]In Hébert's (2018) basic model, there is no funding constraint, but the conflict of interest between the entrepreneur and the investor arises because they discount future cash flows differently. Hébert's (2018) shows that his results carry over to the environment with a funding constraint that I consider here.

[2]Recall that 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 evaluated at a given distribution and a given type measures the marginal cost of increasing the probability mass on this type.

[3]In Georgiadis et al. (2024), a principal offers a wage contract to an agent who covertly chooses a return distribution and is protected by limited liability. The security design problem considered here differs in two points: First, it is the party who offers the contract (the entrepreneur) who also chooses effort. 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. As I will show, the second point implies that not all distributions are implementable by some contract, unlike in Georgiadis et al. (2024). The first point implies that the optimization problem is different.

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

lem: 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.

My analysis shows that the magnitude of agency costs ultimately comes down to the degree of the convexity of the effort cost function. Indeed, when effort costs are linear, agency costs are shown to be zero, and thus the optimal design is efficient. Linear costs correspond to the case in which to generate a stochastic return distribution the entrepreneur must randomize over deterministic ones. The intuitive reason why agency costs are zero in this case, is that at an optimum the entrepreneur increases her effort up to the point where marginal benefits are equal to marginal costs of effort. When costs are linear, marginal effort costs are constant (globally). Moreover, marginal effort benefits are also constant (by definition). Thus, at an optimum, total costs equal total benefits, resulting in zero overall utility from effort for the entrepreneur, hence zero agency costs.

For the case that the first-best is not implementable, I solve the security design problem for various classes of cost functions, and I discuss economic distortions induced by the agency problem. In the wake of the seminal work by Jensen and Meckling (1976), the literature has extensively discussed to what extent the agency problem induces lax effort and/or excessive risk taking by the agent, the former often considered an issue with equity financing, the latter with debt. Because I study optimal design, the distortions I identify are driven by the cost function, not by the choice of financial instrument.[5]

I first consider cost functions which, for a given mean of the distribution, are monotone in its risk (in the sense of a mean preserving spread).[6] When costs are increasing in risk, the first-best distribution displays no risk at all and is deterministic. The same is true for the second-best distribution if, in addition, the moral hazard rent, which only depends on the cost function, is also increasing in risk. The reason is that in this case, keeping the expected return fixed, both actual

[5]On this point, see also Hellwig (2009) where the entrepreneur can separately choose between the mean and (a measure of) the risk of the return distribution.

[6]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.

costs and agency costs are decreasing in risk. The expected return under the first- and second-best can, in general, not be ranked without additional assumptions. But for the case that costs depend only on a moment of the distribution, the expected return in the second-best is downward distorted.

When costs are decreasing in risk, both first- and the second-best distributions are "maximally risky" in the sense that they put probability mass only on the lowest and highest possible return. But now the first-best expected return is always larger than the second-best, implying the that the first-best distribution first (and thus second) order stochastically dominates the second-best.

To shed light on the case when costs are neither increasing nor decreasing in risk, I focus on cost functions that depend on finitely many generalized moments of the distribution. Using extreme point arguments based on Winkler (1988), I show that the resulting first- and second-best distributions are then discrete with a finite support. Ranking the first- and second best stochastically is generally difficult, especially when the cost function depends on more than one moment. However, when the cost function do depend on one moment only, the second-best mean is always smaller than the first-best mean.

An important 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.[7] In particular, in Innes' (1990) parametric and in Hébert's (2018) divergence cost based approach, (risky) debt arises as an optimal security when securities are required to be monotone.[8] One intuition is that, within the constraints of the entrepreneur's limited liability, debt renders the entrepreneur approximately the residual claimant, thus aligning her incentives best with the first-best.

By contrast, in my flexible 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). For example, I show that if costs are decreasing in risk, the optimal distribution can be implemented by both debt and equity. The reason is that in my setting, optimal securities are only pinned down on the support of optimal distributions but leave many (in fact, infinitely many) degrees of freedom off the support. The optimal distributions I identify have finite support within the continuous interval of possible returns, and whether the security makes the entrepreneur the residual claimant off the support is

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

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

immaterial. 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.[9,10]

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

1 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 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.

[9]Within the setting of Innes (1990) with one-dimensional effort, Yang and Zefentis (2024) show that a contingent debt security, where the face value of the debt is contingent on the project return, is optimal when the conditional return distributions, conditional on effort, do not satisfy the monotone likelihood ratio property and securities must be monotone.

[10]Requiring monotonicity does not substantially change my conclusions. In fact, I show that monotonicity of an optimal security can often be obtained "for free".

The entrepreneur's problem is

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

Note that limited liability is required for all 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.

1. C is continuous, monotone, and convex.[11]

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

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

3. The cost of the "smallest" distribution, which places mass 1 on $\underline{x}$, is normalized to 0: $C(\delta_{\underline{x}}) = 0$.[12]

Continuity is a technical condition that ensures the existence of various maximizers below. Monotonicity means that larger return distributions are more costly, that is, $C(F) \geq C(G)$ if F first order stochastically dominates G. Convexity captures increasing marginal effort costs. For example, consider a weighted average of a "low-return" and a "high-return" distribution, that is, the latter first order stochastically dominates the former. Convexity then implies that marginally increasing the weight on the high-return distribution gets more costly the higher the weight. Both monotonicity and convexity are natural assumptions in the effort provision context considered here.

[11] Continuity refers to the weak topology.

[12] 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 Cerreia-Vioglio et al., 2017).

A first-best distribution $F^{(1)}$ maximizes the total expected project value $\int x\,dF - C(F) - K$. Let

$$V^{(1)} = \max_F \int x\,dF - C(F) - K$$

be the first-best project value. To make the problem non-trivial, I assume that $V^{(1)}$ is positive.

2 Implementability

I start the analysis by studying which distributions are implementable by some security. F is implementable if there is a security S so that all constraints in the entrepreneur's problem are satisfied. The next proposition characterizes implementability. To state it, let

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

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

$$\lambda \leq \hat{\lambda}_F, \tag{IR'}$$

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

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})).

In other words, (MH_{supp}) and (MH_{all}) mean that for a security to implement a given F, the security is—up to the constant λ— pinned down by the marginal effort costs c_F on the support of F, and bounded from below on the set of possible returns X. The economic interpretation is that equation (MH_{supp}) defines a security that 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 (IR) and (LL). The lump sum must not be too large (so that IR and the investor's limited liability is not violated) and not be too small (so that the entrepreneur's limited liability is not violated). Condition (IR') in Proposition 1 specifies the range of λ's that are consistent with (IR), and condition (LL') specifies the range of λ's that are consistent with (LL).

It is instructive to illustrate the constraints (MH_{supp}), (MH_{all}), and (LL') graphically. The

[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 (see Georgiadis et al. (2024)).

left panel of Figure 1 depicts a distribution F that 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 implement 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 that 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 LL 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 LL 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 implementable. 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 LL constraint $S(\underline{x}) \leq \underline{x}$. For the security to implement F, it would need to pass trough the dotted points in order to satisfy the moral hazard constraint (MH_{supp}). But this is impossible without violating the LL constraint $S(x_2) \geq 0$. In other words, in the right panel, (LL') cannot be satisfied because $\underline{\lambda}_F > \overline{\lambda}_F$.

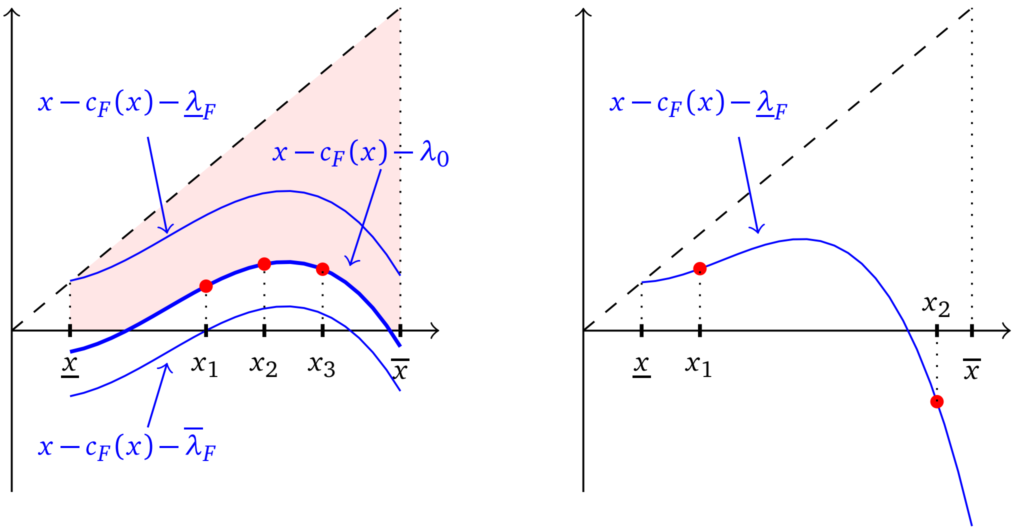


Figure 1: The figure illustrates when the implementability conditions (MH_{supp}), (MH_{all}), and (LL') can (left) and cannot (right) be satisfied

I conclude this section with a remark on when a distribution can be implemented by securities that play prominent roles in practice, such as monotone or debt securities. A security is monotone if S is increasing, and a debt security is of the form $S_d(x) = \min\{x, d\}$ for some "face value" of debt $d \in [\underline{x}, \overline{x}]$. Observe that the distribution in the left panel of Figure 1 *cannot* be implemented by a monotone security (let alone by a debt security). The reason is that a security S that implements F has to pass through the three dotted points and lie in the shaded area. Thus, $S(x_3) < S(x_2)$.

3 Optimal security design

I now turn to the problem of finding an optimal security and an optimal distribution. By (MH_{supp}), the security that implements a given F is pinned down on the support of F up to the lump sum λ. Therefore, expected payoffs from implementing F are pinned down up to the lump sum. In particular, the entrepreneur's expected payoff, gross of effort costs $C(F)$, is $\int x - S(x)\, dF = \int c_F(x)\, dF + \lambda$. In light of Proposition 1, the entrepreneur's problem P_0 can therefore be re-written as

$$P: \qquad \max_{F,\lambda} \int c_F(x)\, dF + \lambda - C(F) \quad s.t. \quad (IR'), (LL').$$

Before I solve P formally, a reminder of the basic tradeoff 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 her LL constraint $S(x) \leq x$. The combination of moral hazard and this LL 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 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 ideally like to maximize the overall expected project return by committing to choose a first-best distribution $F^{(1)}$. She could then just compensate the investor for her capital costs K and in this way obtain the first-best project value $V^{(1)}$. The key issue, however, is that the entrepreneur's minimal payout $\Pi_E(F^{(1)})$ that is required to implement the first-best under moral hazard and limited liability might be too large for the remaining payout $\Pi_I(F^{(1)})$ to allow the investor to recoup his capital costs K.

Conversely, this suggests that whenever $\Pi_I(F^{(1)}) \geq K$, it is optimal to implement the first-best. I now make these considerations more precise.

3.1 Implementability of the first-best

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

Lemma 1 *A distribution $F^{(1)}$ is first-best if and only if there is $\lambda^{(1)}$ so that*

$$x - c_{F^{(1)}}(x) - \lambda^{(1)} = 0 \quad \forall x \in supp(F^{(1)}), \tag{FB_{supp}}$$

$$x - c_{F^{(1)}}(x) - \lambda^{(1)} \geq 0 \quad \forall x \in X. \tag{FB_{all}}$$

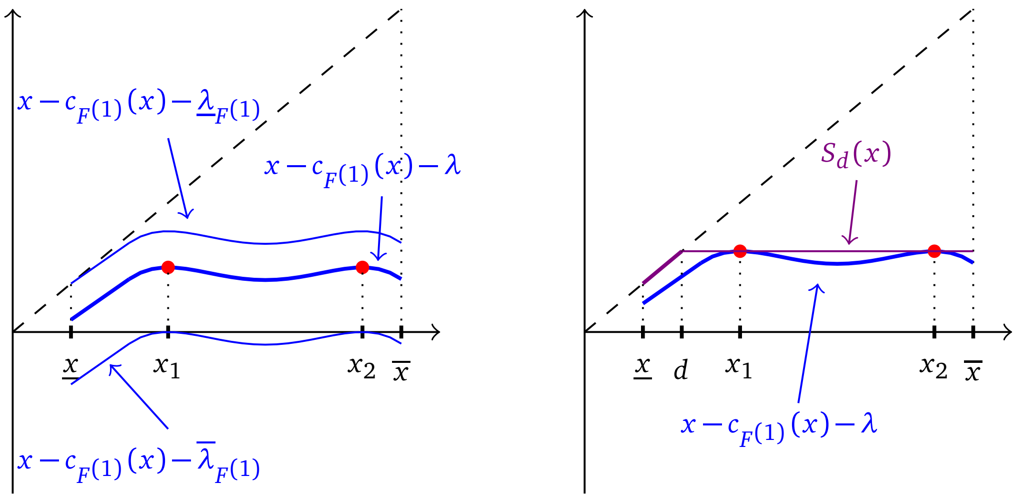
The conditions in the lemma are the first order conditions that equate the marginal (first-best) project return and the marginal costs of effort. They are analogous to the moral hazard conditions, but in the first-best the benefit from marginally increasing the probability mass assigned to the return x is equal to x (instead of the entrepreneur's payout $x - S(x)$). Crucially, the conditions (FB_{supp}) and (FB_{all}) imply that the function $x - c_{F^{(1)}}(x)$ is maximized on the support of $F^{(1)}$. This is illustrated in panel (a) of Figure 2 which displays a first-best distribution with support $\{x_1, x_2\}$.

I now use panel (a) of Figure 2 to shed light on when $F^{(1)}$ is implementable. As explained above, a security that satisfies (MH_{supp}), (MH_{all}), and (LL) has to pass through the dotted points, lie above the curve $x - c_{F^{(1)}}(x) - \lambda$, be positive and below the 45 degree line. It is evident from the figure that it is always possible to find such a security: The problem that $\underline{\lambda}_{F^{(1)}}$ might be larger than $\overline{\lambda}_{F^{(1)}}$ that arises in the right panel in Figure 1 does not arise for a first-best distribution. Hence:[15]

$$\underline{\lambda}_{F^{(1)}} \leq \overline{\lambda}_{F^{(1)}}. \tag{4}$$

[15]Formally, because $x - c_{F^{(1)}}(x)$ is maximized by any point in the support of $F^{(1)}$ by (FB_{supp}), it follows

$$\underline{\lambda}_{F^{(1)}} = -c_{F^{(1)}}(\underline{x}) \leq \underline{x} - c_{F^{(1)}}(\underline{x}) \leq \min_{x \in supp(F^{(1)})} (x - c_{F^{(1)}}(x)) = \overline{\lambda}_{F^{(1)}}. \tag{3}$$



(a) A first-best distribution

(b) Implementation of $F^{(1)}$ by a debt security

Figure 2: Implementation of the first-best

In other words, the constraint (LL') can always be satisfied for a first-best distribution $F^{(1)}$. Therefore, $F^{(1)}$ is implementable whenever (LL') can be jointly satisfied with the remaining constraint (IR'), which combines to $\underline{\lambda}_{F^{(1)}} < \hat{\lambda}_{F^{(1)}}$. This inequality says that the smallest lump sum $\underline{\lambda}_{F^{(1)}}$ that the entrepreneur has to receive to satisfy her LL constraint $S(x) \leq x$, is sufficiently small to satisfy the investor's IR constraint.

Now observe that if $F^{(1)}$ is implementable, then it 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 $F^{(1)}$ is implementable if and only if $\underline{\lambda}_{F^{(1)}} \leq \hat{\lambda}_{F^{(1)}}$, or, equivalently, $\Pi_I(F^{(1)}) - K \geq 0$.*

2. If a first-best distribution $F^{(1)}$ is implementable, then $F^* = F^{(1)}$ and $\lambda^* = \Pi_I(F^{(1)}) - K - c_{F^{(1)}}(\underline{x})$ is a solution to P, and the entrepreneur's ex ante profit is the first-best project value $V^{(1)}$.

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]

[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.)

Proposition 3 *Let C be linear. Then 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. Therefore, the investor's maximal payout, $\Pi_I(F)$, is equal to $\int x \; dF - C(F)$. In particular, for a first-best distribution this is larger than capital costs K by assumption so that $\Pi_I(F^{(1)}) - K \geq 0$ holds.

Intuitively, the entrepreneur optimally "increases" her effort up to the point where marginal benefits are equal to marginal costs of effort. When costs are linear, marginal effort costs are constant, that is, independent of F. Moreover, the entrepreneur's benefit from effort is her expected payout, and thus linear in F. Accordingly, marginal benefits of effort are constant, too. Thus, at an optimum, total costs equal total benefits, resulting in zero overall utility from effort for the entrepreneur, that is, a zero net moral hazard rent.

I conclude this section by observing that if a first-best distribution is implementable, it is implementable by a debt security. This can be seen in panel (b) of Figure 2. Recall that for a security to implement $F^{(1)}$, it has to pass through $x - c_{F^{(1)}}(x) - \lambda$ on the support of $F^{(1)}$ (the dotted points in the figure) and lie above it everywhere else. Because $x - c_{F^{(1)}}(x) - \lambda$ is maximal on the support of $F^{(1)}$ by (FB_{supp}) and (FB_{all}) (the dotted points are on same level) the debt security $S_d(x)$ where $d = \max_{x \in supp(F^{(1)})} x - c_{F^{(1)}}(x) - \lambda$ satisfies the implementability conditions.

Proposition 4 *If the first-best is implementable, it can be implemented by a debt security.*

Notice that the debt security that implements the first-best is "safe debt" in the sense that the return realizations where the entrepreneur becomes insolvent are off the support. In this sense, the proposition is a kind of converse to Jensen and Meckling's (1976) observation that safe debt induces efficient effort becomes it makes the entrepreneur the residual claimant.

4 Second-Best

I now turn to the entrepreneur's problem when the first-best cannot be implemented. Inequality (4) implies that there is always a security that satisfies the moral hazard and the limited liability constraints for a first-best distribution. Therefore, if the first-best is not implementable, it is

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

because the investor's IR constraint is violated even when the lump sum takes the smallest possible value $\underline{\lambda}_{F(1)}$ 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 LL constraint $S(x) \leq x$ and the IR constraint, that is,

$$\underline{\lambda}_F \leq \lambda \leq \hat{\lambda}_F. \tag{5}$$

Since the entrepreneur's objective in P is increasing in λ, it is optimal to set $\lambda = \hat{\lambda}_F$. If one plugs this into the objective and observes that the constraint (5) is equivalent to the condition $\Pi_I(F) \geq K$, one arrives at the relaxed problem

$$R: \qquad \max_F \int x \, dF - C(F) - K \quad s.t. \quad \int x \, dF - \Pi_E(F) - K \geq 0. \tag{6}$$

(In the formulation of the constraint, I use that $\Pi_I(F) = \int x \, dF - \Pi_E(F)$. This makes the arguments below more transparent.) 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. 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 that depend only on one generalized moment

$$\Phi_F = \int \varphi(x) \, dF$$

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

[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 $\varphi(\underline{x}) = 0$ is without loss. Because $C(\delta_{\underline{x}}) = 0$ means that $\Gamma(\varphi(\underline{x})) = 0$, one obtains an equivalent model with functions $\tilde{\Gamma}$ and $\tilde{\varphi}$ and with the property that $\tilde{\varphi}(\underline{x}) = 0$ by setting: $\tilde{\varphi}(x) = \varphi(x) - \varphi(\underline{x})$ and $\tilde{\Gamma}(\Phi) = \Gamma(\Phi + \varphi(\underline{x}))$.

[20]Convexity of Γ ensures convexity of C. Assuming strict convexity simplifies some arguments, but is not required for my results. Note that when Γ is linear, costs C are linear so that Proposition 3 applies.

4.1 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 $F^{(1)} = \delta_{x^{(1)}}$ where $x^{(1)}$ maximizes $x - C(\delta_x)$ is a first-best distribution. Moreover, the first-best is implementable if and only if $x^{(1)} - \Pi_E(\delta_{x^{(1)}}) - K \geq 0$.*

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 with a given mean $\int x\, dF = M$:

$$P_M^{(1)}: \qquad \max_{F: \int x\, dF = M} \int x\, dF - C(F). \tag{7}$$

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 $\int x\, dF = M$:

$$R_M: \qquad \max_{F: \int x\, dF = M} \int x\, dF - C(F) - K \quad s.t. \quad \int x\, dF - \Pi_E(F) - K \geq 0. \tag{8}$$

Fixing M and decreasing the risk of the distribution now not only increases the objective, but also affects the constraint. However, if $\Pi_E(F)$ decreases as the distribution becomes less risky, 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*

[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 $M^{(1)} \in \arg\max_M M - \Gamma(M)$ is a first-best distribution.

degenerate distribution δ_{x^*} where x^* maximizes

$$x - C(\delta_x) \quad s.t. \quad x - \Pi_E(\delta_x) - K \geq 0. \tag{9}$$

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. An example where C and Π_E are both increasing in risk is a moment-based cost function with a convex moment function φ. In this case, the first- and the second-best can also be ranked.

Proposition 7 *Let C be increasing and moment-based. Then Π_E is increasing in risk. Moreover, $x^* \leq x^{(1)}$, that is, the first-best first order dominates the second-best.*

To see that Π_E is increasing in risk, I use a result by Cerreia-Vioglio (2017) which says that a Gateaux-differerentiable function is increasing in risk if its Gateaux-derivative is convex in x for all F. If C is moment-based, 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

$$\Pi_E(F) = \int c_F(x)\, d(F - \delta_{\underline{x}}) = \Gamma'(\Phi_F)\Phi_F, \tag{10}$$

and the Gateaux derivative of Π_E is

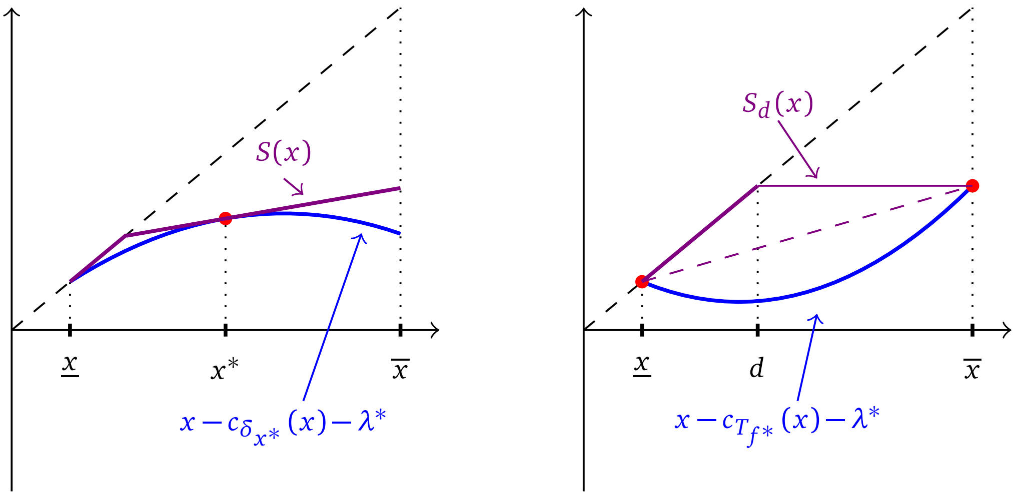
$$(\Pi_E)_F(x) = [\Gamma''(\Phi_F)\Phi_F + \Gamma'(\Phi_F)]\varphi(x).$$

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 $x^* \leq x^{(1)}$ is driven by two feature 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.

To conclude this section, I ask what type of security implements the second-best distribution δ_{x^*}. To see when the second-best can be implemented by a monotone security, note that when C is increasing in risk, the function $x - c_F(x) - \lambda$ is concave (because c_F is convex, as pointed out above). This observation together with the discussion at the end of Section 2 implies that δ_{x^*} is implementable by a monotone security if and only if $x - c_{\delta_{x^*}}(x)$ is upward sloping at the point $x = x^*$.

The left panel of Figure 3 illustrates this case. The piece-wise linear security S in the figure implements δ_{x^*} because it passes through the dotted point and is above the curve $x - c_{\delta_{x^*}}(x) - \lambda^*$ everywhere else. While the depicted security corresponds to a form of "piece-wise equity", it is evident that many other securities implement δ_{x^*}, too.



(a) C and Π_E increasing in risk

(b) C and Π_E decreasing in risk

Figure 3: Implementation of the second-best for the case that costs and the entrepreneur's minimal payout are increasing in risk (left) and decreasing in risk (right).

The next result shows that $x - c_{\delta_{x^*}}(x)$ is indeed upward sloping at the point $x = x^*$ when costs are moment-based.

Proposition 8 *Let C be increasing in risk and moment-based. Then there is a monotone security that implements the optimal distribution δ_{x^*}*

The argument behind the proposition uses that when costs are moment-based, the second best x^* is smaller than the first-best $x^{(1)}$. Together with the first order condition for the first-best, this can be used to show the result.

4.2 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 two-point distribution supported on $\{\underline{x}, \overline{x}\}$ that places mass f 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 now begin with the characterization of the first-best.

Proposition 9 *If C is decreasing in risk, a first-best distribution is the two-point distribution $T_{f^{(1)}}$ where $f^{(1)}$ maximizes $(1-f)\underline{x} + f\overline{x} - \kappa(f)$. Moreover, $T_{f^{(1)}}$ is implementable if and only if it places mass 1 on the point $x = \overline{x}$ (that is, $f^{(1)} = 1$) and*

$$\overline{x} - \pi_E(1) - K \geq 0. \tag{11}$$

The intuition can be seen again from considering subproblem $P_M^{(1)}$ 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 $f^{(1)}$.[22]

The first-best can be implemented only if it puts probability 1 on the largest return realization $\overline{x}$. Otherwise, the investor's IR constraint can never be satisfied. To see this intuitively, recall that if the first-best is implementable, then it can be implemented by a debt security with a debt level $d = \max_{x \in supp(F^{(1)})} x - c_{F^{(1)}}(x) - \lambda$ where λ is larger than $\lambda_{F^{(1)}} = -c_{F^{(1)}}(\underline{x})$ (if λ was smaller, the entrepreneur's LL constraint $S(x) \leq x$ would be violated at $\underline{x}$). Therefore, if a first-best distribution puts probability on the lowest return realization, the debt level is at most $d = \overline{x}$. Because $\underline{x} < K$ by assumption, the investor cannot recoup his capital cost under a debt 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

[22] As above, if C is mean-based any distribution with mean $M^{(1)} \in \arg\max_M M - \Gamma(M)$ is a first-best distribution.

distribution also relaxes the constraint. Hence, the distribution T_f with mean M solves problem R_M. Maximizing over M then yields:

Proposition 10 *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*

$$(1-f)\underline{x}+f\overline{x}-\kappa(f) \quad s.t. \quad (1-f)\underline{x}+f\overline{x}-\pi_E(f)-K\geq 0. \tag{12}$$

Moreover, the solution to R is also a solution to the original problem P, and it holds: $f^*\leq f^{(1)}$.

The fact that the solution to R solves P can be easily verified. An example where C and Π_E are both decreasing in risk is a moment-based cost function with a concave moment function φ. The result by 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. The arguments used after Proposition 6 equally apply to see that if C is decreasing in risk, so is Π_E.

I conclude this section by arguing that the optimal distribution from Proposition 10 can be implemented by a debt security. This is illustrated in the right panel of Figure 3. Because C is decreasing in risk, its Gateaux-derivative is concave, and thus the function $x-c_{T_{f^*}}(x)$ is convex. The support of the optimal distribution T_{f^*} is $\{\underline{x},\overline{x}\}$. Thus, any security that passes through the dotted points and is located above the curve $x-c_{T_{f^*}}(x)-\lambda^*$ implements T_{f^*}. Notice that for a security to implement T_{f^*}, $S(\overline{x})$ (the right dot) must be larger than $S(\overline{x})$ (the left dot). The reason is that otherwise the security would always pay out less than $\underline{x}$ to the investor, and since $\underline{x}<K$ by assumption, it would thus not cover capital costs. Therefore, since $x-c_{T_{f^*}}(x)$ is concave, the debt security S_d with $d=\overline{x}-c_{T_{f^*}}(\overline{x})-\lambda^*$ passes through the dotted points and is located above the curve $x-c_{T_{f^*}}(x)-\lambda^*$, and hence implements T_{f^*}.

Observe, however, that many other securities also implement T_{f^*}. A case in point is an affine security, corresponding to "equity", that connects the dotted points through a straight line.

Proposition 11 *Let C and Π_E be decreasing in risk. Then the optimal distribution T_{f^*} can be implemented by both debt and equity.*

Note that in contrast to Proposition 8, the statement does not require that costs are moment-based.

4.3 Moment-based costs

In this section, I assume that costs are moment-based, but not necessarily increasing or decreasing in risk. Recall that C is moment-based if $C(F) = \Gamma(\Phi_F)$ with $\Phi_F = \int \varphi(x)\, dF$ for a convex and increasing 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 12 *Let C be moment-based, then there is a first-best distribution with at most two points in its support, and the first-best mean $M^{(1)}$ maximizes $M - \Gamma(\breve{\varphi}(M))$.*

To see the result, consider first problem $P_M^{(1)}$ 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 $\min_{F: \int x\, dF = M} \Phi_F = \breve{\varphi}(M)$. In the second step, one maximizes over M to find the first-best mean: $\max_M M - \Gamma(\breve{\varphi}(M))$.

Next, I characterize the second-best. Recall from (13) that with moment-based costs, also the entrepreneur's minimal payout is moment-based:

$$\Pi_E(F) = \Gamma'(\Phi_F)\Phi_F. \tag{13}$$

Problem R_M thus writes:

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

Note that 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 13 *Let C be moment-based.*

1. *Then there is a solution F^* to problem R with at most two points in its support and its mean M^* is a solution to*

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

2. *Moreover, $M^* \leq M^{(1)}$, and there are a first-best distribution $F^{(1)}$ and a solution F^* to R, each with at most two points in their support, so that $F^{(1)}$ first order stochastically dominates F^*.*

3. *If F^* is a degenerate distribution, then it also solves the original problem P. Otherwise, it 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 $M^* \leq M^{(1)}$ 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 $F^{(1)}$ and F^* are unique. Recall that $F^{(1)}$ 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 are the lower and upper points of the segment in which M^i is located, $i \in \{(1), *\}$. Now, because $M^* \leq M^{(1)}$, there are only two possibilities. Either M^* and $M^{(1)}$ are located in the same segment. In this case, $F^{(1)}$ and F^* have the same support. Or, M^* is located in a "lower" segment than $M^{(1)}$. In this case, all points in the support of F^* are smaller than all points in the support of $F^{(1)}$. In either case, $F^{(1)}$ first order stochastically F^*. (Part 3. is technical.)

I now turn to the question whether the optimal distribution F^* from the previous proposition can be implemented by a monotone security. Note that at both points x_1, x_2 in the support of F^*, the function φ coincides with its lower convex envelope: $\varphi(x_i) = \breve{\varphi}(x_i)$. This follows from the convexification argument above: The distribution that minimizes the moment Φ_F convexifies φ. Therefore, the security $S(x)$ that corresponds to the upper concave envelope of $x - c_{F^*}(x) = x - \Gamma'(\Phi_{F^*})\varphi(x)$ up to a constant implements F^*. In particular, if $x - \Gamma'(\Phi_{F^*})\varphi(x)$ is increasing in x, so is S. A sufficient condition for this is that $\varphi'(x) \leq \Gamma'(\varphi(\overline{x}))$ for all x. This is the same condition that is stated at the end of Proposition 13. Figure 4 illustrates.

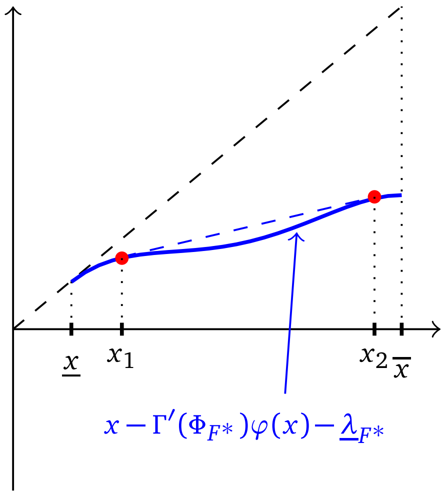


Figure 4: Moment-based costs

I conclude this section with a generalization to cost functions that depend not on one but on an arbitrary, finite number of moments.

Proposition 14 *Let $\varphi = (\varphi_1, \ldots, \varphi_K)$ be a vector of K increasing functions $\varphi_k : X \to \mathbb{R}$ and define the moment vector*

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

Suppose costs are moment-based, that is, $C(F) = \Gamma(\Phi_F)$, where $\Gamma : \mathbb{R}^K \to \mathbb{R}$ is convex.

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 $\Phi_F^k = \Phi^k, k = 1, \ldots, K$. 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.

5 Conclusion

This paper studies security design with moral hazard when the entrepreneur's effort choice is flexible and she can choose any distribution of returns subject to a cost. I characterize how agency costs depend on the cost function and show that the first-best is optimal when the cost function is linear. I characterize optimal second-best outcomes for cost functions that are monotone in risk or depend on the moments of the distribution only. I show that optimal distributions have finite support which implies that optimal securities are not unique, and seemingly very different financial instruments might be equivalent. For example, if costs are decreasing in risk, both debt and equity are optimal.

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, 2025a,b) demonstrates that the flexible moral hazard approach is sufficiently tractable to do so.

6 Proofs

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

$$\overline{\lambda}_{F^{(1)}} = \lambda^{(1)}, \quad \hat{\lambda}_{F^{(1)}} = \lambda^{(1)} - K, \tag{15}$$

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

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

$$\lambda = \hat{\lambda}_{F^{(1)}} = \int x - c_{F^{(1)}}(x)\, dF^{(1)} - K = \Pi_I(F^{(1)}) - K - c_{F^{(1)}}(x) \tag{16}$$

the entrepreneur's expected payoff is

$$\int c_{F^{(1)}}(x)\, dF^{(1)} - C(F^{(1)}) + \hat{\lambda}_{F^{(1)}} = \int x\, dF^{(1)} - C(F^{(1)}) - K. \tag{17}$$

Thus, the entrepreneur gets the first-best project value $V^{(1)}$, and she can clearly not do better than implementing $F^{(1)}$ 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,

$$\underline{\lambda}_{F^{(1)}} = -\varphi(\underline{x}) = 0, \quad \hat{\lambda}_{F^{(1)}} = \int x \, dF - C(F^{(1)}) - K. \tag{18}$$

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

Proof of Proposition 4 The argument is in the main text. 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 $\hat{\lambda}_F$. 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 $\hat{\lambda}_{F^*} \leq \overline{\lambda}_{F^*}$, that is,

$$\int x - c_{F^*}(x) \, dF^* - K \leq \min_{x \in supp(F^*)} (x - c_{F^*}(x)). \tag{19}$$

Now observe that for the degenerate distribution $F^* = \delta_{x^*}$, condition (19) writes

$$x^* - c_{F^*}(x^*) - K \leq x^* - c_{F^*}(x^*), \tag{20}$$

and is thus satisfied. QED

Proof of Proposition 7 The proof that Π_E is increasing in risk is in the main text. It remains to show that $x^* \leq x^{(1)}$. This is trivial if $x^{(1)} = \overline{x}$. Hence, let $x^{(1)} < \overline{x}$ be interior. Denote by $\alpha(x)$ the objective function and by $\beta(x)$ the constraint in problem (9). Since costs are moment-based,

$$\alpha(x) = x - \Gamma(\varphi(x)), \quad \beta(x) = x - \Gamma'(\varphi(x))\varphi(x) - K. \tag{21}$$

Note that because Γ is strictly convex and φ is convex (because costs are increasing in risk), the first-best value $x^{(1)}$ uniquely maximizes α, and it follows that $\alpha'(x) < 0$ for all $x > x^{(1)}$. Moreover, $\alpha'(x) \geq \beta'(x)$, since Γ is strictly convex. Therefore, it follows that also $\beta'(x) < 0$ for all $x > x^{(1)}$.

Now, suppose, contrary to the claim, that $x^{(1)} < x^*$. Then lowering x^* would increase the objective and relax the constraint, a contradiction to the optimality of x^*. QED

Proof of Proposition 8 As discussed before the statement of the proposition, it is sufficient to show that $x - c_{\delta_{x^*}}(x)$ is increasing at $x = x^*$. With moment-based costs, this is equivalent to

$$0 \leq \frac{d}{dx}(x - \Gamma'(\varphi(x^*))\varphi(x)) \,|_{x=x^*} = 1 - \Gamma'(\varphi(x^*))\varphi'(x^*). \tag{22}$$

To see this, note that the first-order condition for $x^{(1)}$ to maximize the first-best objective $\alpha(x) = x - \Gamma(\varphi(x))$ is

$$\begin{aligned} 0 &\leq \frac{d}{dx}(x - \Gamma(\varphi(x))) \,|_{x=x^{(1)}} && (23) \\ &= 1 - \Gamma'(\varphi(x^{(1)}))\varphi'(x^{(1)}) && (24) \\ &\leq 1 - \Gamma'(\varphi(x^*))\varphi'(x^*), && (25) \end{aligned}$$

where the last inequality follows from the fact that $x^* \leq x^{(1)}$ (by Proposition 7) and that Γ', φ and φ' are increasing by assumption. But this is (22). QED

Proof of Proposition 9 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:

$$\kappa'(f) = c_{T_f}(\overline{x}) - c_{T_f}(\underline{x}), \quad \mathit{and} \quad \pi_E(f) = f\kappa'(f). \tag{26}$$

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

$$\begin{aligned}
\kappa'(f) &= \frac{dC(T_f)}{df} && (27)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[C(T_{f+\epsilon}) - C(T_f)\right] && (28)\\
&= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[C(T_f + \frac{\epsilon}{1-f}(T_1 - T_f)) - C(T_f)\right] && (29)\\
&= \frac{1}{1-f}\lim_{\epsilon\to 0}\frac{1}{\frac{\epsilon}{1-f}}\left[C(T_f + \frac{\epsilon}{1-f}(T_1 - T_f)) - C(T_f)\right] && (30)\\
&= \frac{1}{1-f}\int c_{T_f}(x)\, d(T_1 - T_f) && (31)\\
&= c_{T_f}(\overline{x}) - c_{T_f}(\underline{x}). && (32)
\end{aligned}$$

Moreover,

$$\pi_E(f) = \Pi_E(T_f) = \int c_{T_f}(x)\, dT_f - c_{T_f}(\underline{x}) = f\,[c_{T_f}(\overline{x}) - c_{T_f}(\underline{x})] = f\,\kappa'(f), \qquad (33)$$

as desired.

I next prove the "only if"-part of the implementability statement. Let $T_{f^{(1)}}$ be implementable. I show first that $f^{(1)} = 1$. To the contrary, assume $T_{f^{(1)}}$ had positive mass $1 - f^{(1)} > 0$ on $x = \underline{x}$. I derive a contradiction to the first-best implementabiltiy condition $\Pi_I(T_{f^{(1)}}) - K \geq 0$ stated in Proposition 2. Indeed, if there is positive mass on $\underline{x}$, then (FB_{supp}) and (FB_{all}) imply

$$\underline{x} - c_{T_{f^{(1)}}}(\underline{x}) = \lambda^{(1)} \quad and \quad \overline{x} - c_{T_{f^{(1)}}}(\overline{x}) \leq \lambda^{(1)}. \qquad (34)$$

Thus,

$$\begin{aligned}
\Pi_I(T_{f^{(1)}}) - K &= \int x - c_{T_{f^{(1)}}}(x)\, dT_{f^{(1)}} + c_{T_{f^{(1)}}}(\underline{x}) - K && (35)\\
&= (1 - f^{(1)})(\underline{x} - c_{T_{f^{(1)}}}(\underline{x})) + f^{(1)}(\overline{x} - c_{T_{f^{(1)}}}(\overline{x})) + c_{T_{f^{(1)}}}(\underline{x}) - K && (36)\\
&\leq \lambda^{(1)} + c_{T_{f^{(1)}}}(\underline{x})) - K && (37)\\
&= \underline{x} - K. && (38)
\end{aligned}$$

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

$T_{f^{(1)}}$ is implementable, the inequality $\Pi_I(T_{f^{(1)}})-K \geq 0$ holds by Proposition 2. Since $f^{(1)}=1$ and because $\pi_E(f)=f[c_{T_f}(\overline{x})-c_{T_f}(\underline{x})]$ by (26), the said inequality simplifies to $\overline{x}-\pi_E(1)-K \geq 0$, as desired.

To see the "if"-part, let $f^{(1)}=1$ and $\overline{x}-\pi_E(1)-K \geq 0$. I show that $T_{f^{(1)}}=T_1$ 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 (26), one has

$$\Pi_I(T_1)-K = \int x-c_{T_1}(x)\,dT_1+c_{T_1}(\underline{x})-K \tag{39}$$

$$= \overline{x}-c_{T_1}(\overline{x}))+c_{T_1}(\underline{x})-K \tag{40}$$

$$= \overline{x}-\pi_E(1)-K, \tag{41}$$

which is positive by assumption. Thus, $T_{f^{(1)}}=T_1$ is implementable by Proposition 2. QED

Proof of Proposition 10 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) $f^* \leq f^{(1)}$.

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

$$\int x-c_{F^*}(x)\,dF^*-K \leq \min_{x\in supp(F^*)}(x-c_{F^*}(x)). \tag{42}$$

For the two-point distribution $F^*=T_{f^*}$, condition (42) writes

$$(1-f^*)[\underline{x}-c_{T_{f^*}}(\underline{x})]+f^*[\overline{x}-c_{T_{f^*}}(\overline{x})]-K \leq \min\{\underline{x}-c_{T_{f^*}}(\underline{x}),\, \overline{x}-c_{T_{f^*}}(\overline{x})\}. \tag{43}$$

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, (43) is equivalent to $-K \leq 0$ and therefore holds.

Second, suppose T_{f^*} is not a first-best distribution. Then the constraint in (12) is binding, and, by (26), reads

$$(1-f^*)\underline{x}+f^*\overline{x}-f^*[c_{T_{f^*}}(\overline{x})-c_{T_{f^*}}(\underline{x})]-K=0. \tag{44}$$

Therefore, the left hand side of (43) is equal to $-c_{T_{f^*}}(\underline{x})$. To determine the right hand side, (44) can be re-arranged to

$$\underline{x} - c_{T_{f^*}}(\underline{x})) + \frac{K - \underline{x}}{f} = \overline{x} - c_{T_{f^*}}(\overline{x}). \tag{45}$$

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

$$\min\{\underline{x} - c_{T_{f^*}}(\underline{x})), \overline{x} - c_{T_{f^*}}(\overline{x})\} = \underline{x} - c_{T_{f^*}}(\underline{x})). \tag{46}$$

Therefore, the right hand side of (43) is equal to $\underline{x} - c_{T_{f^*}}(\underline{x}))$. Because the right hand side is equal to $-c_{T_{f^*}}(\underline{x}))$, and $0 \le \overline{x}$, (43) is shown, and this establishes claim (a).

As To (b). To see that $f^* \le f^{(1)}$, I show below that κ is convex and that $\kappa'(f) \le \pi'_E(f)$ for all f. The claim then follows from the same formal arguments that establish that $x^* \le x^{(1)}$ in the proof of Proposition 8, 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:

$$T_{\tau f + (1-\tau)g} = \tau T_f + (1-\tau) T_g. \tag{47}$$

Therefore, the convexity of κ follows from the convexity of C. To see that $\kappa'(f) \le \pi'_E(f)$ for all f, recall that $\pi_E(f) = f\kappa'(f)$, and thus

$$\pi'_E(f) = \kappa'(f) + \kappa''(f)f \ge \kappa'(f), \tag{48}$$

as desired. QED

Proof of Proposition 11 As illustrated in the right panel of Figure 3, it is sufficient to show that

$$\underline{x} - c_{T_{f^*}}(\underline{x}) \le \overline{x} - c_{T_{f^*}}(\overline{x}). \tag{49}$$

Recall that $\kappa'(f) = [c_{T_f}(\overline{x}) - c_{T_f}(\underline{x})]$. Thus, (49) is equivalent to

$$\kappa'(f^*) \le \overline{x} - \underline{x}. \tag{50}$$

To see this, recall from the proof of Proposition 10 that κ is convex and that $f^* \leq f^{(1)}$ so that

$$\kappa'(f^*) \leq \kappa'(f^{(1)}). \tag{51}$$

Finally, observe that the first-order condition for $f^{(1)}$ to maximize the first-best objective is

$$\frac{d}{df}\left[(1-f)\underline{x} + f\overline{x} - \kappa(f)\right]_{f=f^{(1)}} \geq 0 \quad \Leftrightarrow \quad \overline{x} - \underline{x} \geq \kappa'(f^{(1)}). \tag{52}$$

Putting the two inequalities together, establishes (50) and completes the proof. QED

Proof of Proposition 12 The argument is in the main text. QED

Proof of Proposition 13 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 (9) for moment-based costs, because $\breve{\varphi}$ is convex by definition. The proof that $M^* \leq M^{(1)}$ is thus identical to the proof that $x^* \leq x^{(1)}$ in Proposition 7.

To see the second statement of part 2. recall that Proposition 12 and part 1. of the statement imply that there is a first-distribution $\tilde{F}^{(1)}$ a solution $\tilde{F}^*$ to R with support points $\tilde{x}_1^{(1)} \leq M^{(1)} \leq \tilde{x}_2^{(1)}$ and $\tilde{x}_1^* \leq M^* \leq \tilde{x}_2^*$.

Observe that if $\tilde{x}_2^* \leq \tilde{x}_1^{(1)}$ then all points in the support $\tilde{F}^{(1)}$ are larger than all points in the support $\tilde{F}^*$, and hence $\tilde{F}^{(1)}$ evidently first order stochastically dominates $\tilde{F}^*$. Therefore, I from now on assume that $\tilde{x}_2^* > \tilde{x}_1^{(1)}$. I distinguish the following cases.

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

To see the claim, note that because $\tilde{x}_1^* \leq M^* \leq M^{(1)} \leq \tilde{x}_2^{(1)}$, there is a convex combination of $\tilde{x}_1^*$ and $\tilde{x}_2^{(1)}$ that is equal to M^*. Therefore there is a two-point distribution with support $\{\tilde{x}_1^*, \tilde{x}_2^{(1)}\}$ and mean M^*. Let this distribution be F^*. Moreover, because $\tilde{x}_2^{(1)} \leq \tilde{x}_2^*$ by assumption, F^* is a mean preserving contraction (MPC) of $F^{(1)}$. 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, $i \in \{(1), *\}$. By standard convexification arguments, this also implies that φ coincides with its lower convex envelope on the respective supports. In particular,

$$\varphi(\tilde{x}_1^*) = \breve{\varphi}(\tilde{x}_1^*), \quad \varphi(\tilde{x}_2^{(1)}) = \breve{\varphi}(\tilde{x}_2^{(1)}), \quad \varphi(\tilde{x}_2^*) = \breve{\varphi}(\tilde{x}_2^*). \tag{53}$$

Therefore,

$$\Phi_{F^*} = \int \varphi \, dF^* = \int \breve{\varphi} \, dF^* \leq \int \breve{\varphi} \, d\tilde{F}^* = \int \varphi \, d\tilde{F}^* = \Phi_{\tilde{F}^*} = \min_{\int x \, dF = M^*} \Phi_F, \tag{54}$$

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: $\tilde{x}_2^{(1)} > \tilde{x}_2^*$.

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

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

$$\int x - c_{F^*}(x) \, dF^* - K \leq \min_{x \in supp(F^*)} (x - c_{F^*}(x)). \tag{55}$$

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, (55) 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 (55) writes

$$
\begin{aligned}
& (1-f^*)[x_1-\Gamma'(\Phi_{F^*})\varphi(x_1)]+f^*[x_2-\Gamma'(\Phi_{F^*})\varphi(x_2)]-K \\
\leq \; & \min\{x_1-\Gamma'(\Phi_{F^*})\varphi(x_1),\; x_2-\Gamma'(\Phi_{F^*})\varphi(x_2)\}, \qquad (56)
\end{aligned}
$$

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:

$$
M^*-\Gamma'(\Phi_{F^*})\Phi_{F^*}-K=\int x-\Gamma'(\Phi_{F^*})\varphi(x)\,dF^*-K=0. \qquad (57)
$$

Therefore, the left hand side of inequality (56) is equal to zero. Accordingly, inequality (56) holds if and only

$$
x_i-\Gamma'(\breve{\varphi}(M^*)\varphi(x_i)\geq 0, \quad i=1,2. \qquad (58)
$$

To complete the proof, I show that a sufficient condition for (58) is

$$
\varphi'(x)\leq\frac{1}{\Gamma'(\varphi(\overline{x}))} \quad \forall x\in X. \qquad (59)
$$

Indeed, recall that $\varphi(x)=0$. Thus a sufficient condition for (58) is that $x-\Gamma'(\breve{\varphi}(M^*)\varphi(x)$ is increasing in x, that is,

$$
1-\Gamma'(\breve{\varphi}(M^*)\varphi'(x)\geq 0 \quad \forall x\in X. \qquad (60)
$$

This condition is implied by (59) 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 14 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 13. QED

References

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

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.

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.

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

Krähmer, Daniel (2025b). The hold-up problem with flexible unobservable investments, mimeo.

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.