Optimal Contracts with a Risk-Taking Agent*

Daniel Barron, George Georgiadis, and Jeroen Swinkels†

March 22, 2018

Abstract

Consider an agent who can costlessly add mean-preserving noise to his output. To deter such risk-taking, the principal optimally offers a contract that makes the agent's utility concave in output. If the agent is risk-neutral and protected by limited liability, this concavity constraint binds and so linear contracts maximize profit. If the agent is risk averse, the concavity constraint might bind for some outputs but not others. We characterize the unique profit-maximizing contract and show how deterring risk-taking affects the insurance-incentive tradeoff. Our logic extends to costly risk-taking and to dynamic settings where the agent can shift output over time.

*We are grateful to Simon Board, Ralph Boleslavsky, Luis Cabral, Odilon Câmara, Gabriel Carroll, Hugo Hopenhayn, Johannes Hörner, Eddie Lazear, Elliot Lipnowski, Erik Madsen, Niko Matouschek, John Matsusaka, Meg Meyer, Konstantin Milbradt, Dimitri Papanikolaou, Andrea Prat, Lars Stole, Doron Ravid, Luis Rayo, Patrick Rey, Balazs Szentes, Dimitri Vayanos, Glen Weyl, and seminar participants at Cambridge University, New York University, Northwestern University, the University of Chicago, the University of Southern California, University College London, the University of Miami, the University of Exeter, the University of Rochester, the University of Warwick, the 2016 ASU Theory Conference, the 2016 Transatlantic Theory Conference, CRETE 2017, and CSIO-IDEI 2017 for helpful comments and discussions.

†Kellogg School of Management, Northwestern University, Evanston, IL 60208, U.S.A.
D. Barron: d-barron@kellogg.northwestern.edu;
G. Georgiadis: g-georgiadis@kellogg.northwestern.edu;
J. Swinkels: j-swinkels@kellogg.northwestern.edu

1 Introduction

Contracts motivate employees, suppliers, and partners to exert effort. However, improperly designed incentives can instead encourage excessive risk-taking with dramatic consequences. For instance, following the 2008 financial crisis, Federal Reserve Chairman Ben Bernanke stated that "compensation practices at some banking organizations have led to misaligned incentives and excessive risk-taking, contributing to bank losses and financial instability" (Federal Reserve Press Release (10/22/2009)). Garicano and Rayo (2016) suggest that poorly designed incentives led the American International Group (AIG) to expose itself to massive tail risk in exchange for the appearance of stable earnings. Rajan (2011) echoes these concerns and suggests that misaligned incentives worsened the effects of the crisis.

Even without such disastrous outcomes, agents face opportunities to game their incentives by engaging in risk-taking in many other settings. Portfolio managers can choose riskier investments, as well as exert effort, to influence their average returns (Brown, Harlow, and Starks (1996); Chevalier and Ellison (1997); de Figueiredo, Rawley, and Shelef (2014)). Executives and entrepreneurs control both the expected profitability of their projects and the distribution over possible outcomes (Matta and Beamish (2008); Rahmandad, Henderson, and Repenning (2016); Vereshchagina and Hopenhayn (2009)). Salespeople can both invest to increase demand and adjust the timing of the resulting sales (Oyer (1998); Larkin (2014)).

This paper explores how a principal optimally motivates an agent who can engage in risk-taking in a canonical moral hazard setting. We argue that risk-taking renders convex incentives ineffective, so that the principal optimally offers a contract that makes the agent's utility concave in output. This result is the foundation of our analysis, which explores the implications of this concavity constraint and shows that it has potentially significant effects on the structure of contracts, profits, productivity, and welfare.

In our model, a principal contracts with a potentially liquidity-constrained agent. If the agent accepts the contract, then he exerts costly effort that produces a non-contractible intermediate output. The agent privately observes this output and can then manipulate it by costlessly adding mean-preserving noise, which in turn determines a contractible final output.

Building on the arguments of Jensen and Meckling (1976) and others, Section 3 shows that the agent will choose to take on additional risk whenever intermediate output is such that his utility under the contract is convex at that output. In so doing, the agent makes his expected utility concave in intermediate output. If the principal and agent are both weakly risk-averse, the principal finds it optimal to deter risk-taking entirely by offering an incentive scheme that directly makes the agent's utility concave in output. We refer to this additional constraint – that utility be weakly concave in output – as the **no-gaming constraint**. Wherever the no-gaming constraint binds, the optimal contract makes the agent's utility linear in output.

In Section 4, we consider a risk-neutral agent. In this setting, we show that the no-gaming constraint binds everywhere, so that a linear (technically, affine) contract is optimal, remains so regardless of the principal's attitude toward risk (even if she is risk-loving), and is uniquely optimal if the principal is risk averse. Intuitively, absent the no-gaming constraint, the principal would like to offer a convex contract, which would concentrate high pay on high outcomes and so inexpensively motivate the agent while respecting the liability constraint. Therefore, the no-gaming constraint binds and so the principal optimally offers a linear contract. We show that relative to any strictly concave contract, there is a linear contract that both better insures the principal and better motivates the agent.

Section 5 builds on this logic to explore optimal incentives with a risk-averse agent (and a risk-neutral principal). In this setting, the no-gaming constraint implies that the agent's *utility* must be concave in output. Similar to Section 4, the optimal contract makes the agent's utility linear wherever this constraint binds. Unlike that section, however, the no-gaming constraint does not necessarily bind everywhere.

Suppose that the limited liability constraint is slack. Without risk-taking, the optimal contract would equate the principal's marginal cost of paying the agent to the marginal benefit of relaxing the agent's participation and incentive constraints *at each output* (Mirrlees (1976); Holmström (1979)). However, doing so might violate the no-gaming constraint. Where it does, we show that the optimal contract is instead **ironed**, in the sense that it is linear in utility and sets *expected* marginal benefits equal to *expected* marginal costs over those regions. For instance, suppose that no-gaming binds for low output but not for high output, which turns out to be true under fairly weak conditions on preferences and production. In that case, the optimal contract makes the agent's utility linear in output below a threshold; above that threshold, utility is concave and equates marginal benefits to marginal costs output-by-output. In the extreme, if no-gaming is slack everywhere, then the standard contract characterized by Mirrlees (1976) and Holmström (1979) is optimal; if it binds everywhere, then the optimal contract makes the agent's utility linear in output.

If instead the participation constraint is slack (and so limited liability binds), then we show that the no-gaming constraint binds for any output that suggests less than the desired effort. As a result, and similar to the intuition outlined above, the profit-maximizing incentive scheme makes the agent's utility linear over these outputs.

These results are implications of a set of necessary and sufficient conditions we develop for a profit-maximizing contract. Since balancing benefits and costs output-by-output might violate concavity, we cannot characterize the profit-maximizing contract using the techniques of Mirrlees (1976) and Holmström (1979). Instead, we construct two simple perturbations of a candidate contract that preserve concavity while changing either the level or the slope of the agent's utility over appropriate *intervals* of output. Perhaps surprisingly, we prove that it is also sufficient to consider these two perturbations, so that a contract is profit-maximizing if, and only if, it cannot be improved by them.

Finally, Section 6 considers three extensions, all of which assume that both the principal and the agent are risk-neutral. First, we consider optimal contracts in a model where the agent engages in risk-taking before he observes intermediate output. We show that the agent's risk-taking concavifies his expected payoff conditional on his *effort*, rather than on intermediate output, and we identify mild conditions under which a linear contract is optimal. Second, we consider optimal contracts if the agent incurs a cost that is increasing in the variance of her risk-taking distribution. We show that our basic intuition extends to this setting, resulting in a unique optimal contract that is convex in output, but not so convex as to induce gaming. As gaming becomes cheap, the optimal contract becomes linear. Finally, we study a dynamic setting in which the principal offers a stationary contract that the agent can game by choosing *when* output is realized over an interval of time. Oyer (1998) and Larkin (2014) document how convex incentive schemes and long sales cycles can encourage such intertemporal gaming. We show that this setting is equivalent to our risk-taking model. Thus, a linear contract is optimal, since a convex contract would induce the agent to bunch sales over short time intervals and a strictly concave contract would provide subpar effort incentives.

Our analysis is inspired by Diamond (1998) and Garicano and Rayo (2016). Diamond (1998) is a seminal exploration of optimal contracts when the agent can both exert effort and make other choices that affect the output distribution. Section 6 of that paper argues that linear contracts are among the many optimal contracts in an example with risk-neutral parties, binary effort, and an agent who can choose any mean-preserving spread of output. Our Proposition 2 expands this result to settings with more general effort choices, output distributions, and principal utility functions.

Relative to Diamond (1998), we contribute in three ways. First, we show that the fundamental consequence of agent risk-taking is to constrain incentives to be *concave*, not necessarily linear. Linear contracts are instead a consequence of this concavity constraint binding everywhere, as it does if the agent is risk-neutral. However, as Section 5 demonstrates, the concavity constraint need not necessarily bind everywhere if the agent is risk-averse, in which case the optimal contract makes utility strictly concave in output. Second, our Proposition 2 identifies an additional conceptual advantage of linear contracts with a risk-neutral agent: relative to any strictly concave contract, they better insure the principal and so are *uniquely* optimal if the principal is even slightly risk-averse. Finally, our analysis of optimal contracts with a risk-averse agent in Section 5 is entirely new and shows how risk-taking affects contracts in a canonical moral hazard setting. Garicano and Rayo (2016) includes a model of risk-taking that is similar to ours, but it fixes an exogenous (non-concave) contract to focus on the social costs of excessive risk.

Our model of risk-taking is embedded in a canonical moral hazard problem. With a risk-neutral agent, our model builds on Innes (1990), Spulber and Poblete (2012), and other papers for which limited liability is the central contracting friction. With a risk-averse agent, we build on Mirrlees (1976) and Holmström (1979) if the limited liability constraint is slack, and Jewitt, Kadan, and Swinkels (2008) if it binds. Within the classic agency literature, our analysis is conceptually related to papers that study principal-agent relationships in which the agent both exerts effort and makes other decisions. Classic examples include Lambert (1986) on how agency problems in information-gathering can lead to inefficient investment in risky projects and Holmström and Ricart i Costa (1986) on project selection under career concerns. Malcomson (2009) presents a general model of such settings, but differs from our analysis in assuming that decisions are contractible. Other papers consider settings in which the principal also chooses actions other than the agent's wage contract, such as an endogenous performance measure; see for example, Halac and Prat (2016) and Georgiadis and Szentes (2018).

A growing literature studies agent risk-taking. Some papers in this literature assume that an agent chooses from a parametric class of risk-taking distributions in either static (Palomino and Prat (2003); Hellwig (2009)) or dynamic (Demarzo, Livdan, and Tchistyi (2014)) settings. We differ by allowing our agent to choose any mean-preserving spread of output, which means that our optimal contract must deter a more flexible form of gaming. Therefore, we join other papers that study non-parametric risk-taking, again in either static (Robson (1992); Diamond (1998); Hébert (2015)) or dynamic (Ray and Robson (2012); Makarov and Plantin (2015)) settings. We differ from these papers by identifying *concavity* as the key constraint on the optimal incentive scheme if the agent can costlessly take on risk and then characterizing optimal incentives given this constraint in the context of a canonical moral hazard problem.[1]

More broadly, our work is related to a long-standing literature which argues that optimal contracts must both induce effort and deter gaming. A seminal example is Holmström and Milgrom (1987), which displays a dynamic environment in which linear contracts are optimal. Recent papers, including Chassang (2013), Carroll (2015), and Antic (2016), take up this point by departing from a Bayesian framework and proving that simple contracts perform well under min-max or other non-Bayesian preferences. In contrast, our paper considers contracts that deter gaming in a setting that lies firmly within the Bayesian tradition.

While the contracting problems are quite different, Carroll's intuition is related to ours. In that paper, Nature selects a set of actions available to the agent in order to minimize the principal's expected payoffs. The key difference is in the *types* of gambles available to the agent. In Carroll's paper, Nature might allow the agent to take on additional risk to game a convex incentive scheme, in which case risk-taking behavior is similar to that in our paper and its predecessors. However, if the principal offers a concave incentive scheme, then Nature might also allow the agent to choose a distribution with *less* risk. In contrast, we allow the agent to add risk but not reduce it. This difference is most striking if the agent is risk-averse, in which case Carroll's optimal contract makes the agent's utility linear in output, while ours might make utility strictly concave.

2 Model

We consider a static game between a principal (P, "she") and an agent (A, "he"). The agent has limited liability, so he cannot pay more than $M \in \mathbb{R}$ to the principal. Let $[\underline{y}, \bar{y}] \equiv \mathcal{Y} \subseteq \mathbb{R}$ be the set of contractible outputs. The

[1] In Ray and Robson (2012), Condition R2 is a version of a concavity constraint. However, that paper analyzes how risk-taking by status-conscious customers affect the intergenerational wealth distribution, and in particular it studies neither moral hazard nor optimal contracts.

timing is as follows:

1. The principal offers an upper semicontinuous contract $s(y) : \mathcal{Y} \to [-M, \infty)$.[2]

2. The agent accepts or rejects the contract. If he rejects, the game ends, he receives u_0, and the principal receives 0.

3. If the agent accepts, he chooses effort $a \geq 0$.

4. Intermediate output x is realized according to $F(\cdot|a) \in \Delta(\mathcal{Y})$.

5. The agent chooses a distribution $G_x \in \Delta(\mathcal{Y})$ subject to $\mathbb{E}_{G_x}[y] = x$.

6. Final output y is realized according to G_x, and the agent is paid $s(y)$.

The principal's and agent's payoffs are equal to $\pi\left(y - s(y)\right)$ and $u\left(s\left(y\right)\right) - c(a)$, respectively.

We assume that $\pi(\cdot)$ and $u(\cdot)$ are strictly increasing and weakly concave, with $u(\cdot)$ onto, and that $c(\cdot)$ is infinitely differentiable, strictly increasing, and strictly convex. We also assume that $F(\cdot)$ has full support for all $a \geq 0$, satisfies $\mathbb{E}_{F(\cdot|a)}[x] = a$, and is infinitely differentiable with a density $f(\cdot)$ that is strictly MLRP-increasing in a, with $\frac{f_a(\cdot|a)}{f(\cdot|a)}$ uniformly bounded for all a.

This game is similar to a canonical moral hazard problem, with the twist that the agent can engage in risk-taking by choosing a mean-preserving spread G_x of intermediate output x. Let

$$\mathcal{G} = \left\{G : \mathcal{Y} \to \Delta(\mathcal{Y}) \mid \mathbb{E}_{G_x}[y] = x \text{ for all } x \in \mathcal{Y}\right\}.$$

Without loss, we can treat the agent as choosing a and $G \in \mathcal{G}$ simultaneously.

Intermediate output can take different forms in different settings. For instance, CEOs typically have advance information about whether or not they will hit their earnings targets in a given quarter, and they can cut maintenance or R&D expenditures if they are likely to fall short, taking on tail risk for

[2]One can show that the restriction to upper semicontinuous contracts is without loss: if the agent has an optimal action given a contract $s(\cdot)$, then there exists an upper semicontinuous contract that induces the same equilibrium payoffs and distribution over final output.

the appearance of higher earnings (Rahmandad, Henderson, and Repenning (2016)). Similarly, portfolio managers are typically compensated based on their annual returns and can adjust the riskiness of their investments over the course of the year in order to game those incentives (Chevalier and Ellison (1997)).[3]

After the agent observes x but before y is realized, we have a setting with both a hidden type and a hidden action. In such problems, it is often useful to ask the agent to report his type, in this case x. By punishing differences between this report and y, the principal may be able to dissuade some or all gambling. We restrict attention to situations where such intermediate reports are not useful. The simplest way to do so is to assume that the timing of x is random, and gambling is instantaneous.[4] We think this is the economically correct modeling assumption in many settings. Indeed, the spirit of the model is that the agent can misbehave in a particular direction, and it seems unlikely that the principal can catalog the precise moments and ways in which this might occur.

3 Risk-taking and optimal incentives

This section explores how the agent's ability to engage in risk-taking constrains the contract offered by the principal.

We find it convenient to rewrite the principal's problem in terms of the utility $v(y) \equiv u(s(y))$ that the agent receives for each output y. If we define $\underline{u} \equiv u(-M)$, then an optimal contract solves the following constrained

[3] An alternative assumption is that the agent engages in risk-taking before uncertainty is resolved, which may be more natural in some applications. Section 6.1 explores this alternative.

[4] Allowing reports would change the agent's gaming incentives but not completely eliminate them. Online Appendix E.1 has an analysis with risk-neutral parties and shows that linear contracts are optimal even if such reports are allowed.

maximization problem:

$$\max_{a,\, G\in\mathcal{G}, v(\cdot)} \quad \mathbb{E}_{F(\cdot|a)}\left[\mathbb{E}_{G_x}\left[\pi\left(y-u^{-1}\left(v\left(y\right)\right)\right)\right]\right] \tag{Obj$_F$}$$

$$\text{s.t.} \quad a, G \in \arg\max_{\tilde{a},\, \tilde{G}\in\mathcal{G}} \left\{\mathbb{E}_{F(\cdot|\tilde{a})}\left[\mathbb{E}_{\tilde{G}_x}\left[v\left(y\right)\right]\right] - c(\tilde{a})\right\} \tag{IC$_F$}$$

$$\mathbb{E}_{F(\cdot|a)}\left[\mathbb{E}_{G_x}\left[v\left(y\right)\right]\right] - c(a) \geq u_0 \tag{IR$_F$}$$

$$v\left(y\right) \geq \underline{u} \quad \text{for all } y. \tag{LL$_F$}$$

The main result of this section is Proposition 1, which characterizes how the threat of gaming affects the *incentive schemes* $v(\cdot)$ that the principal can offer. The principal optimally offers a contract that deters risk-taking entirely, but doing so constrains her to incentive schemes that are weakly concave in output. Define G^D so that for each $x \in \mathcal{Y}$, G^D_x is degenerate at x.

Proposition 1. *Suppose $(a, G, v(\cdot))$ satisfies (IC$_F$)-(LL$_F$). Then there exists a weakly concave $\hat{v}(\cdot)$ such that $(a, G^D, \hat{v}(\cdot))$ satisfies (IC$_F$)-(LL$_F$) and gives the principal a weakly higher expected payoff.*

The proof is in Appendix A. For an arbitrary incentive scheme $v(\cdot)$, define $v^c(\cdot) : \mathcal{Y} \to \mathbb{R}$ as its concave closure,

$$v^c(x) = \sup_{w,z\in\mathcal{Y}, p\in[0,1] \text{ s.t. } (1-p)w+pz=x} \left\{(1-p)v(w) + pv(z)\right\}. \tag{1}$$

At any outcome x such that the agent does not earn $v^c(x)$, he can engage in risk-taking to earn that amount in expectation (but no more). But then the principal can do at least as well by directly offering a concave contract, and if either the agent or the principal is strictly risk-averse, then offering a concave contract is strictly more profitable than inducing risk-taking.

Given Proposition 1, we can write the optimal contracting problem as one without risk-taking but with a no-gaming constraint that requires the agent's utility to be concave in output, with the caveat that our resulting solution is one of many if (but only if) both parties are risk-neutral:

$$\max_{a,v(\cdot)} \quad \mathbb{E}_{F(\cdot|a)}\left[\pi\left(y - u^{-1}(v(y))\right)\right] \tag{Obj}$$

$$\text{s.t.} \quad a \in \arg\max_{\tilde{a}} \left\{\mathbb{E}_{F(\cdot|\tilde{a})}\left[v(y)\right] - c(\tilde{a})\right\} \tag{IC}$$

$$\mathbb{E}_{F(\cdot|a)}\left[v(y)\right] - c(a) \geq u_0 \tag{IR}$$

$$v(y) \geq \underline{u} \text{ for all } y \in \mathcal{Y} \tag{LL}$$

$$v(\cdot) \text{ weakly concave.} \tag{NG}$$

For a fixed effort $a \geq 0$, we say that $v(\cdot)$ *implements* a if it satisfies (IC)-(NG) for a, and it does so *at maximum profit* if it maximizes (Obj) subject to (IC)-(NG). An *optimal* $v(\cdot)$ implements the optimal effort level $a^* \geq 0$ at maximum profit.

Mathematically, the set of concave contracts is well-behaved. Consequently, we can show that for any $a \geq 0$, a contract that implements a at maximum profit exists, and is unique if either $\pi(\cdot)$ or $u(\cdot)$ is strictly concave.

Lemma 1. *Fix $a \geq 0$ and suppose that $\underline{u} > -\infty$. Then there exists a contract that implements a at maximum profit, and does so uniquely if either $\pi(\cdot)$ or $u(\cdot)$ is strictly concave.*

This result, which follows from the Theorem of the Maximum, is an implication of Proposition 9 in Online Appendix D.[5] Existence is guaranteed by (NG); for example, without this constraint, no profit-maximizing contract would exist with a risk-neutral agent.[6] If at least one player is strictly risk-averse, then Jensen's Inequality implies that a convex combination of two different contracts that implement a also implements a and gives the principal a strictly higher payoff, which proves uniqueness.

[5] All online appendices may be found at https://sites.google.com/site/danielbarronecon/

[6] With a risk-neutral agent, the principal wants to pay the agent only after arbitrarily high output realizations, since those outputs are most indicative of high effort. See, e.g., Innes (1990).

4 Optimal Contracts for a Risk-Neutral Agent

Suppose the agent is risk-neutral, so $u(y) = y$, $v(\cdot) = s(\cdot)$, and $\underline{u} = -M$. In this setting, the key friction is the agent's liability constraint, which might prevent the principal from simply "selling the firm" to the agent.

For any effort level a, define

$$s_a^L(y) = c'(a)(y - \underline{y}) - w,$$

where $w = \min\left\{M, c'(a)(a - \underline{y}) - c(a) - u_0\right\}$. Intuitively, $s_a^L(y)$ is the least costly linear contract that implements a. Note that for a linear contract, (IC) can be replaced by its first-order condition because expected output is linear in effort and the cost of effort is convex.

Define the *first-best effort* $a^{FB} \in \mathbb{R}_+$ as the unique effort that maximizes $y - c(y)$ and so satisfies $c'\left(a^{FB}\right) = 1$. We prove that a linear contract that implements no more than first-best effort is optimal.

Proposition 2. *Let $u(s) \equiv s$. If a^* is optimal, then $a^* \leq a^{FB}$ and $s_{a^*}^L(\cdot)$ is optimal.*

The proofs for all results in this section can be found in Appendix A. To see the intuition, consider $s_{a^{FB}}^L(\cdot)$, which both implements a^{FB} and provides full insurance to the principal. If $s_{a^{FB}}^L(\cdot)$ satisfies (IR) with equality, then it is clearly optimal.

Suppose instead that (IR) is slack for $s_{a^{FB}}^L(\cdot)$, in which case (LL) must bind. Suppose that $(a^*, s^*(\cdot))$ is optimal, and let $\hat{s}(\cdot)$ be the linear contract that agrees with $s^*(\cdot)$ at $\underline{y}$ and gives the agent the same utility as $s^*(\cdot)$ if he chooses effort optimally. As shown in Figure 1, $\hat{s}(\cdot)$ must single-cross $s^*(\cdot)$ from below, effectively moving payments from low to high outputs. Since $F(\cdot|a)$ satisfies MLRP, paying more for high output motivates more effort and so $\hat{s}(\cdot)$ implements some $\hat{a} \geq a^*$. If $\hat{a} \geq a^{FB}$, then $\hat{s}(\cdot) \geq s_{a^{FB}}^L(\cdot)$, and so the principal prefers $s_{a^{FB}}^L(\cdot)$ to $s^*(\cdot)$ because it induces first-best effort, perfectly insures the principal, and gives the agent less utility than $s^*(\cdot)$.

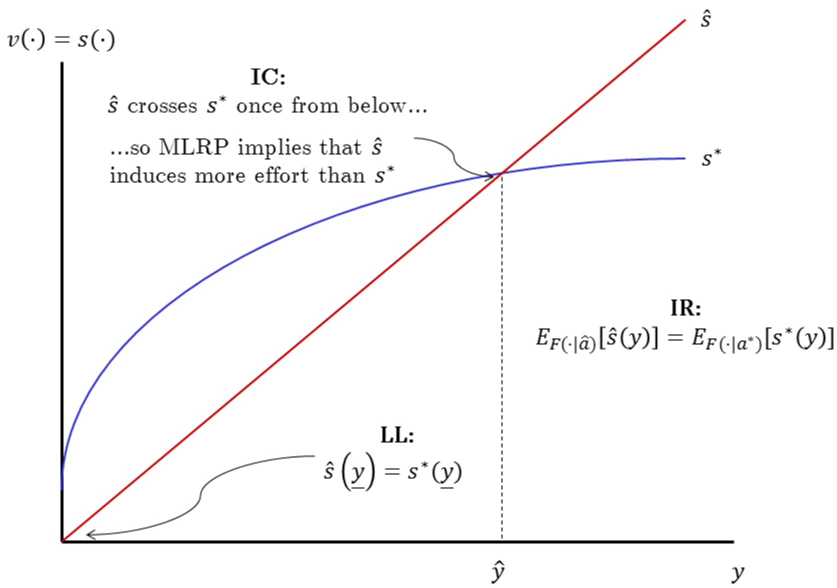

Figure 1: Intuition for the proof of Proposition 2.

If $\hat{a} < a^{FB}$, then $\hat{s}'(\cdot) < 1$ and so the principal's wealth under $\hat{s}(\cdot)$, $y - \hat{s}(y)$, is increasing in y. Consequently, the principal likes that $\hat{s}(\cdot)$ induces more effort than $s^*(\cdot)$. Moreover, $\hat{s}(y) > s^*(y)$ exactly when output is high and so her marginal utility of wealth is low, and so $\hat{s}(\cdot)$ also insures the principal better than $s^*(\cdot)$. So the principal prefers $\hat{s}(\cdot)$ to $s^*(\cdot)$, and *a fortiori* prefers $s^L_{\hat{a}}(\cdot)$, which lies weakly below $\hat{s}(\cdot)$. We conclude that any optimal contract $s^*(\cdot)$ must satisfy $s^*(\cdot) \equiv s^L_{a^*}(\cdot)$.

Lemma 1 implies that $s^L_{a^*}(\cdot)$ is uniquely optimal if the principal is risk-averse. If she is risk-neutral, then $s^L_{a^*}(\cdot)$ is optimal but not uniquely so; in particular, any contract with a concave closure equal to $s^L_{a^*}(\cdot)$ would give identical expected payoffs.

For any $a > 0$, the agent's promised utility under $s^L_a(\cdot)$ depends on $\underline{y}$, the *worst* possible outcome over which the agent can gamble. In particular, such a $s^L_a(\cdot)$ starts at $\underline{y}$ and has a strictly positive slope, so that the agent's expected compensation $E_y[s^L_a(y)] = s^L_a(a)$ increases without bound as $\underline{y}$ decreases. That is, as the agent's ability to take on left-tail risk becomes arbitrarily severe, motivating effort while deterring risk-taking becomes arbitrarily costly to the principal. Consequently, the optimal effort level converges to 0 as $\underline{y}$ becomes arbitrarily negative. Moreover, if the principal is risk-neutral, then we can show that effort is strictly increasing in $\underline{y}$: as the agent's ability to take left-tail risks becomes more severe, the principal responds by inducing lower effort. See Appendix E.3 for details.

In some applications, the principal might have risk-seeking preferences over output, for instance because she also faces convex incentives. For example, Rajan (2011) argues that, anticipating the possibility of bailouts, shareholders of financial institutions might have had an incentive to encourage risk-taking prior to the 2008 financial crisis. We can model such settings by allowing $\pi(\cdot)$ to be any strictly increasing and continuous function. Proposition 1 does not directly apply in this case because the principal might strictly prefer the agent to take on additional risk following some realizations of x. Nevertheless, we can modify the argument from Proposition 2 to show that a linear contract is optimal.

Corollary 1. *Let $u(s) \equiv s$ and let $\pi(\cdot)$ be an arbitrary continuous and strictly increasing function that has a strictly increasing concave closure $\pi^c(\cdot)$. If a^* is optimal, then $a^* \leq a^{FB}$ and $s^L_{a^*}(\cdot)$ is optimal.*

To see the proof of Corollary 1, note that the principal's expected payoff cannot exceed $\pi^c(\cdot)$ for reasons similar to Proposition 1. Therefore, the contract that maximizes $E_{F(\cdot|a)}\left[\pi^c(x - s(x))\right]$ subject to (IC)-(NG) provides an upper bound on the principal's payoff. But Proposition 2 asserts that $s^L_{a^*}(\cdot)$ is optimal in this problem because $\pi^c(\cdot)$ is concave. Given $s^L_{a^*}(\cdot)$, the agent is indifferent among distributions $G \in \mathcal{G}$, so he is willing to choose G such that the principal's expected payoff equals $\pi^c(\cdot)$.

5 Optimal contracts if the agent is risk averse

This section characterizes the unique contract that implements a given $a > 0$ at maximum profit in a setting with a risk-averse agent and a risk-neutral principal. Section 5.1 illustrates how this characterization sheds light on profit-maximizing incentives in the presence of risk-taking. These results are implications of our necessary and sufficient conditions for a profit-maximizing contract, developed in Section 5.2.

We impose two simplifying assumptions to make the analysis tractable. First, letting $\underline{w}$ denote the infimum of the domain of $u(\cdot)$, we assume that $\lim_{w \downarrow \underline{w}} u'(w) = \infty$ and $\lim_{w \uparrow \infty} u'(w) = 0$. Second, we replace (IC) with the weaker condition that local incentives are slack,

$$\frac{d}{da}\left\{\mathbb{E}_{F(\cdot|a)}\left[v\left(y\right)\right] - c(a)\right\} \geq 0. \tag{IC-FOC}$$

Given (NG), replacing (IC) with (IC-FOC) entails no loss if $F(\cdot|\cdot)$ satisfies weak regularity conditions; see, for example, Jewitt (1988) and Chade and Swinkels (2016). For a fixed effort $a \geq 0$, define the principal's problem

$$\max_{v(\cdot)}\left\{\text{(Obj) subject to (IC-FOC), (IR), (LL), and (NG)}\right\}. \tag{P}$$

For $a \geq 0$ and $y \in \mathcal{Y}$, define the likelihood function

$$l(y|a) = \frac{f_a(y|a)}{f(y|a)}.$$

Define $\rho(\cdot)$ as the function that maps $\frac{1}{u'(\cdot)}$ into $u(\cdot)$; that is, for every $w \in (\underline{w}, \infty)$, $\rho\left(\frac{1}{u'(w)}\right) = u(w)$, with $\rho(z) \equiv -\infty$ for all $z \leq 0$.[7] Then $\rho^{-1}(v(y))$ equals the marginal cost to the principal of giving the agent extra utility at y.

If $\underline{u} > -\infty$, then Lemma 1 implies that a unique solution to (P) exists. If $\underline{u} = -\infty$, then Online Appendix D shows that a unique solution exists so long as $u'(\cdot)$ is not excessively convex. Our results in this section apply in either setting. Unless otherwise noted, proofs for this section may be found in Appendix B.

[7] This function is well-defined because $u'(\cdot)$ and $u(\cdot)$ are strictly monotonic. It is continuous because $\lim_{w \downarrow \underline{w}} u'(w) = \infty$ and $\lim_{w \downarrow \underline{w}} u(w) = -\infty$.

5.1 Implications of the No-Gaming Constraint

This section illustrates how risk-taking affects the trade-off between insuring and motivating the agent that lies at the heart of this moral hazard problem. For a broad class of settings, we show that optimal incentives are linear in output where (NG) binds and otherwise equate the marginal costs and benefits of incentive pay.

Given the program (P), let λ and μ be the shadow values on (IR) and (IC-FOC), respectively. For a fixed $a \geq 0$ and an incentive scheme $v(\cdot)$ that implements a, define

$$n(y) \equiv \rho^{-1}(v(y)) - \lambda - \mu l(y|a) \tag{2}$$

as the **net cost** of increasing $v(\cdot)$ at y, taking into account how that increase affects (IR) and (IC-FOC). In particular, increasing $v(y)$ increases the principal's cost at rate $\rho^{-1}(v(y))f(y|a)$, relaxes (IR) at rate $f(y|a)$, which has implicit value λ, and relaxes (IC-FOC) at rate $f_a(y|a)$, which has implicit value μ. Taking the difference between these costs and benefits and dividing by $f(y|a)$ yields $n(y)$.

Suppose that (LL) is slack. Absent (NG), the optimal contract would set $n(y) = 0$ output-by-output and so $v(\cdot) = \rho(\lambda + \mu l(\cdot|a))$. Indeed, this incentive scheme (with the appropriate λ and μ) is the *Holmström-Mirrlees contract* characterized in Mirrlees (1976) and Holmström (1979). However, setting $n(y) = 0$ at each y might violate (NG).

Nevertheless, profit-maximizing contracts build on this basic logic. Intuitively, if setting $n(y) = 0$ at some output y would violate (NG), then this constraint binds, and so the optimal contract is locally linear in utility. These linear segments are "ironed" in the sense that they set net cost equal to 0 *in expectation*, even if they do not do so point-by-point. Outside of these ironed regions, (NG) is slack and so $n(y) = 0$ at each output.

We demonstrate this intuition if $\rho(\lambda + \mu l(\cdot|a))$ is first convex and then concave, which we argue is a natural case to consider.

Lemma 2. *Suppose $u(\cdot)$ and $F(\cdot|a)$ are analytic and $con(\rho') + con(l_y) > -1$.*[8] *Then for any λ and μ, there exists a y_I such that $\rho(\lambda + \mu l(\cdot|a))$ is convex on $[\underline{y}, y_I)$ and concave on $(y_I, \bar{y}]$.*

The proof of Lemma 2 may be found in Appendix E.2. The requirement that $con(\rho') + con(l_y) > -1$ is relatively mild, and holds, for example, if l_y is strictly log-concave and $u(w) = \log w$, or more generally, for a wide range of utilities that satisfy Hyperbolic Absolute Risk Aversion (HARA).[9]

The following Proposition characterizes the optimal contract if $\rho(\lambda+\mu l(\cdot|a))$ is first convex and then concave, and (LL) is slack.

Proposition 3. *Fix $a \geq 0$ and $\pi(y) \equiv y$. Let $v^*(\cdot)$ solve (P), let λ and μ be the shadow values on (IR) and (IC-FOC), respectively, and suppose that $v^*(\underline{y}) > \underline{u}$.*[10] *Suppose there exists y_I such that $\rho(\lambda+\mu l(\cdot|a))$ is convex on $[\underline{y}, y_I)$ and concave on $(y_I, \bar{y}]$. Then $v^*(\cdot)$ satisfies (IR) and (IC-FOC) with equality, and there exist $\hat{y} \geq y_I$, $\underline{v} \in \mathbb{R}$, and $\alpha \in \mathbb{R}_+$ such that $v^*(\cdot)$ is continuous,*

$$v^*(y) = \begin{cases} \underline{v} + \alpha(y - \underline{y}) & \text{if } y < \hat{y} \\ \rho(\lambda + \mu l(y|a)) & \text{otherwise,} \end{cases}$$

and such that $\int_{\underline{y}}^{\hat{y}} n(y) f(y) dy = 0$. If $y_I = \underline{y}$, then $\hat{y} = \underline{y}$.

We interpret Proposition 3 here and defer a discussion of the proof to Section 5.2. Under the condition that $\rho(\lambda + \mu l(\cdot|a))$ is first convex and then concave and (LL) is slack, the profit-maximizing contract $v^*(\cdot)$ is linear in utility for low output and otherwise sets $n(y) = 0$ output-by-output. Moreover,

[8] For any interval $X \subseteq \mathbb{R}$ and analytic function $h : X \rightarrow \mathbb{R}_+$, $con(h) = \inf_X \left\{1 - (hh'')/(h')^2\right\}$. Intuitively, $con(h)$ is the largest value t for which h^t/t is concave. See Prekopa (1973) and Borell (1975) for details.

[9] Recall that HARA utility satisfies $u(w) = \frac{1-\gamma}{\gamma}\left(\frac{\alpha w}{1-\gamma} + \beta\right)^\gamma$. Then for *any* λ and μ, $\rho(\lambda + \mu l(\cdot|a))$ is first convex and then concave for any well-defined HARA utility if either $\gamma < 0$ or $\gamma \in \left(\frac{1}{2}, 1\right)$.

[10] Appendix D gives conditions under which an optimal contract exists even if $\underline{u} = -\infty$. Under those conditions, this existence proof also shows that (LL) is slack if $\underline{u} > -\infty$ is sufficiently negative.

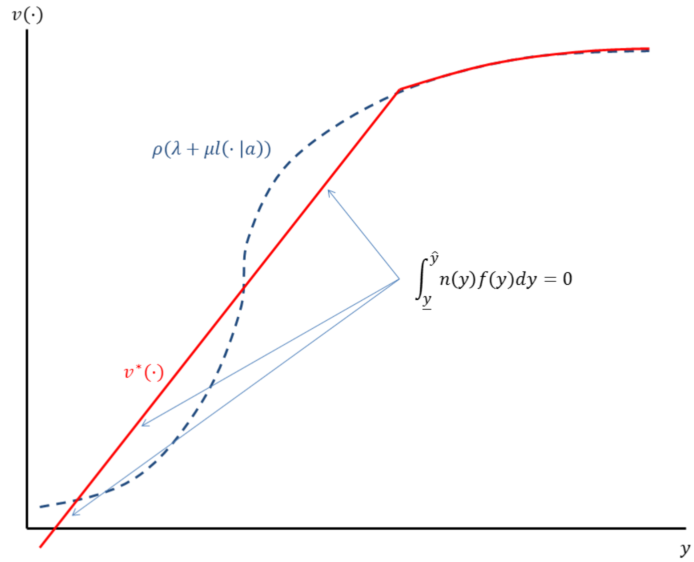

Figure 2: Illustration of $\rho(\lambda + \mu l(\cdot|a))$ and the profit-maximizing $v^*(\cdot)$.

on the linear region of $v^*(\cdot)$, expected net costs equal 0. See Figure 2 for an illustration.

In the extremes, if $\rho(\lambda + \mu l(\cdot|a))$ is convex everywhere, then the profit-maximizing contract is linear,[11] while the profit-maximizing contract equals $\rho(\lambda + \mu l(\cdot|a))$ if the latter is concave. Intuitively, $\rho(\lambda + \mu l(\cdot|a))$ is likely to be convex if the principal would like to "insure against downside risk" by offering low-powered incentives for low output and "motivate with upside risk" by giving steeper incentives for high output. For instance, $\rho(\cdot)$ tends to be more convex if *prudence*, which measures how rapidly the agent becomes less risk-averse as his compensation increases, is large relative to *relative risk aversion*.[12] Conversely, $\rho(\lambda + \mu l(\cdot|a))$ is likely to be concave if the principal would like to "motivate with downside risk" and "insure against upside risk."

Proposition 3 focuses on the case where (LL) is slack, but (NG) has a

[11] This case obtains if, for example, $l(\cdot|a)$ is convex and $\rho(\cdot)$ is convex on the range of $\lambda + \mu l(\cdot|a)$. Note that $\rho(\cdot)$ cannot be convex over its entire domain, because $\rho(0) = -\infty$.

[12] In particular, recalling that prudence is $-\frac{u'''(\cdot)}{u''(\cdot)}$ and relative risk aversion is $-\frac{u''(\cdot)}{u'(\cdot)}$, it can be shown that $\rho(\cdot)$ is convex whenever the ratio of prudence to relative risk aversion exceeds 3.

similar effect if (IR) is slack so that (LL) binds. In that case, the principal would like to pay the agent as little as possible for any y with $l(y|a) < 0$, since paying for low output both increases the agent's rent and tightens (IC-FOC) (Jewitt, Kadan, and Swinkels (2008)). However, rewarding the agent for high output while holding him to his liability constraint following low output violates (NG), which therefore binds following low output.

Proposition 4. *Fix $a \geq 0$ and $\pi(y) \equiv y$. Let $v^*(\cdot)$ solve (P), and suppose that (IR) is slack under $v^*(\cdot)$. Define y_0 such that $l\left(y_0|a\right) = 0$. Then $v^*(\cdot)$ is linear on $[\underline{y}, y_0]$.*

If (IR) is slack and $v^*(\cdot)$ is strictly concave for $y < y_0$, then making it "flatter" on $[\underline{y}, y_0]$ by taking a convex combination of it with the linear segment that connects $v(\underline{y})$ and $v(y_0)$ improves the agent's incentives, and decreases the principal's expected payment. So the profit-maximizing $v^*(\cdot)$ is linear on $[\underline{y}, y_0]$, though it might be strictly concave for higher output.

Before turning to our characterization, it is worth emphasizing that the effects of risk-taking extend beyond those outputs for which (NG) binds. In particular, so long as (NG) binds somewhere, risk-taking potentially distorts both λ and μ away from their levels absent (NG), and so can influence optimal incentives even at outputs where (NG) is slack. That is, λ and μ both shape, and are shaped by, the profit-maximizing incentive scheme.

5.2 A Characterization

This section develops the necessary and sufficient conditions for a profit-maximizing contract that underpin the results in Section 5.1. Since setting $n(y) = 0$ output-by-output might violate (NG), we instead identify perturbations that respect (NG) and affect an interval of an incentive scheme. Then we prove that an incentive scheme is profit-maximizing if and only if it cannot be improved by these perturbations.

We begin by defining several features of $v(\cdot)$ that will be useful for our construction.

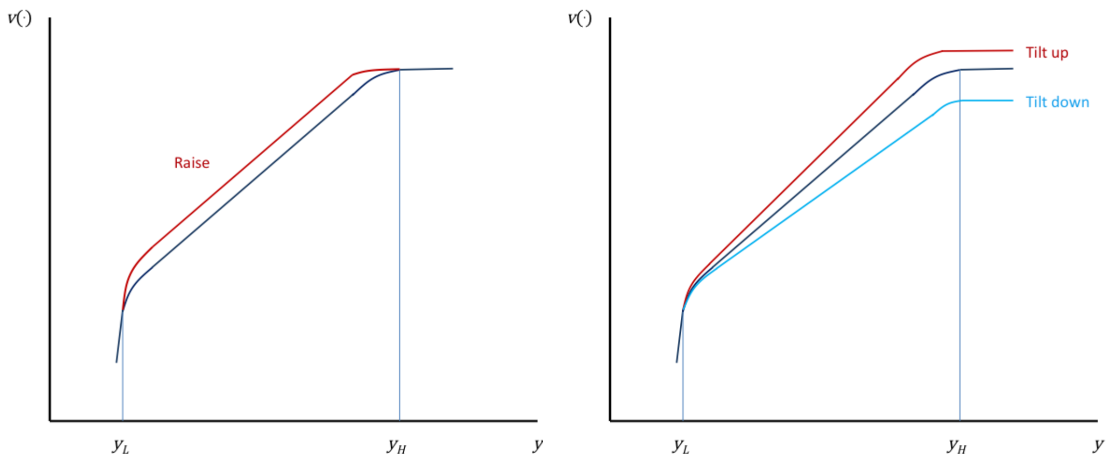

Figure 3: ***Raise*** and ***tilt***. These perturbations require care around y_L and y_H to ensure that concavity is preserved. For this reason, we need both y_L and y_H to be free for ***raise***. For ***tilt up***, we need y_L to be free, while y_H must be free for ***tilt down***.

Definition 1. *Given $v(\cdot)$:*

1. *An interval $[y_L, y_H]$ is a **linear segment** if $v(\cdot)$ is linear on $[y_L, y_H]$ but not on any strictly larger interval. Point y is **free** if it is not on the interior of any linear segment.*

2. *A free $y \in (\underline{y}, \bar{y})$ is a **kink point** of $v(\cdot)$ if two linear segments meet at y, and a **point of normal concavity** otherwise.*

Consider the following two perturbations, formally defined in Online Appendix B and illustrated in Figure 3. *Raise* increases the *level* of $v(\cdot)$ by a constant over an interval, while *tilt* increases the *slope* of $v(\cdot)$ by a constant over an interval. *Raising* an interval typically introduces non-concavities into $v(\cdot)$ at both endpoints of the interval. *Tilting* it a positive amount may introduce a non-concavity at the lower end of the interval, and *tilting* it a negative amount may introduce a non-concavity at the upper end of the interval. Online Appendix B shows that for small perturbations, we can repair these non-concavities on an arbitrarily small interval so long as the relevant endpoints are free.

While *raise* and *tilt* affect both (IR) and (IC-FOC), Online Appendix B shows that since $F(\cdot|a)$ satisfies MLRP, raise and tilt have non-collinear effects on (IR) and (IC-FOC), which means that we can construct combinations of the two perturbations to affect each of these constraints separately. Therefore, so long as there exists at least one free point $\hat{y} < \bar{y}$ such that $v(\hat{y}) > \underline{u}$, we can use these perturbations on $[\hat{y}, \bar{y}]$ to establish the shadow values λ and μ of relaxing (IR) and (IC-FOC).[13]

If an incentive scheme $v(\cdot)$ is profit-maximizing, then it cannot be improved by either *raise* or *tilt* on any valid interval. That is, *raising* $v(\cdot)$ on an interval $[y_L, y_H]$ with both endpoints free must have non-negative expected net cost:

$$\int_{y_L}^{y_H} n(y) f(y|a) dy \geq 0. \tag{3}$$

If $v(y_L) > \underline{u}$, then we can similarly perturb $v(\cdot)$ on $[y_L, y_H]$ by *raising* it a negative amount, so (3) must hold with equality.

Similarly, if y_L is free, then *tilting* $v(\cdot)$ on $[y_L, y_H]$ must have non-negative expected net cost:

$$\int_{y_L}^{y_H} n\left(y\right)\left(y - y_L\right) f\left(y|a\right) dy + \left(y_H - y_L\right) \int_{y_H}^{\bar{y}} n(y) f(y|a) \geq 0, \tag{4}$$

where the first term represents the fact that *tilt* increases the slope of $v(\cdot)$ from y_L to y_H and the second represents the resulting higher level of $v(\cdot)$ from y_H to $\bar{y}$. If y_H is free, then applying negative *tilt* yields the reverse inequality:

$$\int_{y_L}^{y_H} n\left(y\right)\left(y - y_L\right) f\left(y|a\right) dy + \left(y_H - y_L\right) \int_{y_H}^{\bar{y}} n(y) f(y|a) \leq 0. \tag{5}$$

Our characterization combines these perturbations with the usual complementary slackness condition that $\lambda = 0$ if (IR) is slack (so that (LL) binds).

Definition 2. *A contract $v(\cdot)$ is Generalized Holmström-Mirrlees (GHM) if (IC-FOC) holds with equality, (IR), (LL), and (NG) are satisfied, there exist*

[13] If no such point exists, then $v(\cdot)$ is linear and $v(\underline{y}) = \underline{u}$.

$\lambda \geq 0$ and $\mu > 0$ such that

$$\lambda \left(\int_{\underline{y}}^{\bar{y}} v(y) f\left(y|a\right) dy - c(a) - u_0 \right) = 0,$$

and for any $y_L < y_H$,

1. *if y_L and y_H are free, then (3) holds, and holds with equality if $v\left(y_L\right) > \underline{u}$;*
2. *if y_L is free, then (4) holds;*
3. *if y_H is free, then (5) holds.*

Our main result in this section characterizes the unique incentive scheme that implements any $a > 0$ at maximum profit.

Proposition 5. *Suppose $u(\cdot)$ is strictly concave and $\pi(y) \equiv y$. Then for any $a > 0$, $v(\cdot)$ implements a at maximum profit if and only if it is GHM.*

The necessity of GHM follows from the arguments above. To establish sufficiency, we first show that if any $\tilde{v}(\cdot)$ implements a at higher profit than $v(\cdot)$, then there exists a *local* perturbation that improves $v(\cdot)$. Then we show that among local perturbations, it suffices to consider *tilt* and *raise* on valid intervals. This result follows because any perturbation that respects concavity can be approximated arbitrarily closely by a combination of valid tilts and raises. Therefore, if any perturbation improves the principal's profitability, then so must some individual tilt or raise.

One implication of Proposition 5 is that net cost equals 0 for any output where both (LL) and (NG) are slack.

Corollary 2. *Suppose $u(\cdot)$ is strictly concave and $\pi(y) \equiv y$. For any $a > 0$, let $v(\cdot)$ solve (P) and suppose $y \in (\underline{y}, \bar{y})$ is free. Then $n(y) \leq 0$, and $n(y) = 0$ if y is a point of normal concavity.*

At any point of normal concavity y, we can find two free points that are arbitrarily close to y.[14] Proposition 5 implies that (3) holds with equality between these points; taking a limit as these points approach y yields $n(y) = 0$. If

[14]See Claim 1 in Appendix B.

y is a kink point, then we cannot perturb $v(\cdot)$ around y and preserve concavity. However, there is a sense in which (NG) binds on the linear segments on either side of y: Lemma 3 in Online Appendix B proves that absent (NG), the principal would want to increase payments near the ends of a linear segment and decrease them somewhere in the middle of that segment. Therefore, $n(y) \leq 0$ at the endpoints of any linear segment, which includes any kink point.

Together, Proposition 5 and Corollary 2 imply Proposition 3. If (LL) is slack, then (3) holds with equality over any valid interval. Therefore, for each y, the profit-maximizing incentive scheme either sets $n(y) = 0$ or is linear, with *expected* net cost equal to 0 over each linear segment. This is the sense in which our profit-maximizing contract "irons" $\rho(\lambda + \mu l(\cdot|a))$. Moreover, since (NG) binds on any linear segment, $n(\cdot)$ must be negative at the endpoints of that segment and positive somewhere in the middle. So a GHM contract $v(\cdot)$ can have two linear segments only if $\rho(\lambda + \mu l(\cdot|a))$ has a strictly concave region followed by a weakly convex region, which is assumed away in Proposition 3 and ruled out by the condition in Lemma 2.

6 Extensions and Reinterpretations

This section considers three extensions, all of which assume that both the principal and the agent are risk-neutral. Section 6.1 alters the timing so that the agent gambles before observing intermediate output. Section 6.2 changes the agent's utility so that he must incur a cost to gamble. Section 6.3 reinterprets the baseline model as a dynamic setting in which, rather than gambling, the agent can choose *when* output is realized in order to game a stationary contract. Proofs for this section may be found in Online Appendix C.

6.1 Risk-Taking Before Intermediate Output is Realized

If the agent engages in risk-taking before observing intermediate output, then he gambles to concavify his expected utility given *effort*. This section considers this alternative timing and gives conditions under which linear contracts are optimal.

Consider the following timing:

1. The principal offers a contract $s(y) : \mathcal{Y} \to [-M, \infty)$.
2. The agent accepts or rejects the contract. If he rejects, the game ends, he receives u_0, and the principal receives 0.
3. The agent chooses an effort $a \geq 0$ and a distribution $G(\cdot) \in \Delta(\mathcal{Y})$ subject to the constraint $\mathbb{E}_G [x|a] = a$.[15]
4. The outcome of the gamble $x \sim G(\cdot)$ is realized, and final output is realized according to $y \sim F(\cdot|x)$. We assume that $F(\cdot|x)$ has full support, with $\mathbb{E}_{F(\cdot|x)}[y] = x$ and a density $f(\cdot|x)$ that satisfies strict MLRP in x.

The principal and agent earn $y - s(y)$ and $s(y) - c(a)$, respectively, where $c(\cdot)$ is strictly convex.

To interpret risk-taking in this model, suppose the principal is an investor and the agent is an entrepreneur who chooses which project to pursue. Effort improves the expected returns of the project, while $G(\cdot)$ captures other features of the project. For example, a degenerate distribution over x might represent a project that would be modestly profitable irrespective of economic conditions, while a more variable distribution over x corresponds to a project that would be a dramatic success in a growing economy but an utter failure otherwise. Then $F(\cdot|x)$ represents uncertainty that remains after x is realized. We assume that the entrepreneur has access to a rich enough set of projects that $G(\cdot)$ can be any mean-preserving spread of a.[16]

[15] With some notational inconvenience, one can extend this argument to more general mappings from a to $\mathbb{E}_G[x|a]$.

[16] If $\int_{\underline{y}}^{z} F_{xx}(y|x)dy \geq 0$ for all $z \in \mathcal{Y}$ and x, then a riskier $G(\cdot)$ leads to a riskier distribution over final output (in each case, in the sense of second-order stochastic dominance). Note that this is not the only way to model ex-ante risk-taking. For example, we could have modeled the agent as choosing a random, additively separable noise term that affects output. We do not believe that optimal contracts would be linear in that alternative framework.

Given $s(\cdot)$ and x, the agent's expected payoff equals

$$V_s(x) \equiv \int_{\underline{y}}^{\bar{y}} s(y) f(y|x) dy. \tag{6}$$

Define $V_s^c(\cdot)$ as the concave closure of $V_s(\cdot)$ as in (1). Analogous to Proposition 1, the agent will optimally choose G such that $E_{G(\cdot)}[V_s(x)] = V_s^c(a)$. Since $\mathbb{E}_{G(\cdot)}\left[\mathbb{E}_{F(\cdot|x)}[y]\right] = a$ for any $G(\cdot)$, the principal's problem is

$$\begin{aligned} \max_{a,\, s(\cdot)} \;& a - V_s^c(a) \\ \text{s.t. } & a \in \arg\max_{\tilde{a}} \{V_s^c(\tilde{a}) - c(\tilde{a})\} \\ & V_s^c(a) - c(a) \geq u_0 \\ & s(\cdot) \geq -M. \end{aligned} \tag{7}$$

We prove that a linear contract solves this problem.

Proposition 6. *If $a^* \geq 0$ is optimal in the program (7), then $a^* \leq a^{FB}$ and $s_{a^*}^L(\cdot)$ is optimal.*

To see the argument, relax the optimal contracting problem by assuming that the principal can choose $V_s^c(\cdot)$ directly, subject only to the constraints that $V_s^c(\cdot)$ is concave and $V_s^c(\cdot) \geq -M$. This relaxed problem is very similar to (Obj)-(NG), except that $V_s^c(\cdot)$ is a function of effort rather than of intermediate output. Nevertheless, as in the proof of Proposition 2, a linear $V_s^c(\cdot)$ is optimal. But $V_s^c(\cdot)$ is linear if $V_s(\cdot)$ is linear, and $V_s(\cdot)$ is linear if $s(\cdot)$ is linear because $\mathbb{E}_{F(\cdot|x)}[y] = x$. Hence, $s_{a^*}^L(\cdot)$ induces the optimal $V_s^c(\cdot)$ from the relaxed problem and so is optimal.

6.2 Costly Risk-Taking

Consider the model from Section 2, and suppose that the agent must pay a private cost $\mathbb{E}_{G_x}[d(y)] - d(x)$ to implement distribution G_x following the realization of x, where $d(\cdot)$ is smooth, strictly increasing, and strictly convex, with $d(\underline{y}) = 0$. For example, this cost function equals the variance of G_x if

$d(y) = y^2$. More generally, $d(\cdot)$ captures the idea that the agent must incur a higher cost to take on more dispersed risk. The principal's and agent's payoffs are $y - s(y)$ and $s(y) - c(a) - d(y) + d(x)$, respectively.[17]

For any contract $s(\cdot)$, define

$$\tilde{v}(y) \equiv s(y) - d(y) \text{ and } \tilde{c}(a) \equiv c(a) - \mathbb{E}_{F(\cdot|a)}[d(x)],$$

so that conditional on effort, the agent's payoff equals $\tilde{v}(y) - \tilde{c}(a)$. Then the principal's payoff equals $\tilde{\pi}(y) - \tilde{v}(y)$, where $\tilde{\pi}(y) \equiv y - d(y)$ is strictly concave.

As in Section 3, the agent chooses G_x so that his expected payoff equals $\tilde{v}^c(x)$. Since $\tilde{\pi}(\cdot)$ is strictly concave, the principal prefers to deter risk-taking by offering a contract that makes the agent's payoff $\tilde{v}(\cdot)$ concave. Consequently, we can modify the proof of Proposition 2 to show that the principal's optimal contract makes $\tilde{v}(\cdot)$ linear. Therefore, the optimal $s(\cdot)$ is convex and equals the sum of a linear component and $d(\cdot)$.

Proposition 7. *Assume $\tilde{c}(\cdot)$ is strictly increasing and strictly convex. For optimal effort $a^* \geq 0$, define $s^*(y) = \tilde{c}'(a^*)(y - \underline{y}) + d(y) - \tilde{w}$, where $\tilde{w} = \min\left\{M, \tilde{c}'(a^*)(a - \underline{y}) - \tilde{c}(a^*) - u_0\right\}$. Then $s^*(\cdot)$ is optimal.*

This result follows a similar logic to Proposition 2, where the optimal $s^*(\cdot)$ ensures that $\tilde{v}(\cdot)$ is linear. Intuitively, $s^*(\cdot)$ is the most convex contract that deters the agent from gambling. Note that the optimal contract from Proposition 2 is strictly more expensive than $s^*(\cdot)$, since the principal can offer somewhat convex incentives if the agent finds risk-taking costly.

6.3 Manipulating the Timing of Output[18]

This section considers a model in which the principal offers a stationary contract that the agent can game by shifting output across time, rather than by engaging in risk-taking. We show that this model is identical to the setting in Section 4.

[17]We are grateful to Doron Ravid for suggesting this formulation of the cost function.
[18]We are grateful to Lars Stole for suggesting this interpretation of the model.

Consider a continuous-time game between an agent and a principal on the time interval $[0,1]$. Both parties are risk-neutral and do not discount time. At $t=0$:

1. The principal offers a stationary contract $s(y) : \mathcal{Y} \to [-M, \infty)$.
2. The agent accepts or rejects. If he rejects, he earns u_0 and the principal earns 0.
3. The agent chooses an effort $a \geq 0$.
4. Total output x is realized according to $F(\cdot|a) \in \Delta(\mathcal{Y})$.
5. The agent chooses a mapping from time t to output at time t, $y_x : [0,1] \to \mathcal{Y}$, subject to $\int_0^1 y_x(t)dt = x$.
6. The agent is paid $\int_0^1 s(y_x(t))dt$.

The principal's and agent's payoffs are $\int_0^1 [y_t - s(y_t)]\, dt$ and $\int_0^1 s(y_t)dt - c(a)$, respectively. Let $F(\cdot|\cdot)$ and $c(\cdot)$ satisfy the conditions from Section 2.

Crucially, the principal must offer a stationary contract $s(\cdot)$. Without this restriction, the principal could eliminate gaming incentives entirely, for instance by paying only for output realized at a specific time. While stationarity is a significant restriction, we believe it is realistic in many settings: as documented by Oyer (1998) and Larkin (2014), contracts tend to be stationary over some period of time (such as a quarter or a year).

This problem is equivalent to one in which, rather than choosing the realized output $y_x(t)$ at each time t, the agent instead decides *what fraction* of time in $t \in [0,1]$ to spend producing each possible output $y \in \mathcal{Y}$. In particular, define $G_x(y)$ as the fraction of time for which $y_x(t) \leq y$.[19] Then $G_x(\cdot)$ is a distribution that satisfies $\mathbb{E}_{G_x}[y] = x$, and the agent's and principal's payoffs are $\mathbb{E}_{G_x}[s(y)] - c(a)$ and $\mathbb{E}_{G_x}[y - s(y)]$, respectively. That is, intertemporal gaming plays exactly the same role as gambling in our baseline model.

[19] Formally, $G_x(y) = \mathcal{L}(\{t | y_x(t) \leq y\})$, where $\mathcal{L}(\cdot)$ denotes the Lebesgue measure.

Proposition 8. *The optimal contracting problem in this setting coincides with (Obj$_F$)-(LL$_F$) with $u(y) \equiv y$ and $\pi(y) \equiv y$. Hence, if $a^* \geq 0$ is optimal, then $a^* \leq a^{FB}$ and $s^L_{a^*}(\cdot)$ is optimal.*

Intuitively, the agent will adjust his realized output so that his total payoff equals the concave closure of $s(\cdot)$. He does so by smoothing output over time if $s(\cdot)$ is concave, or bunching it in a short interval if $s(\cdot)$ is convex. This behavior is consistent with Oyer (1998) and Larkin (2014), which find that salespeople facing convex incentives concentrate their sales. Conversely, Brav et al. (2005) find that CEOs and CFOs pursue smooth earnings to avoid the severe penalties that come from falling short of market expectations.

7 Concluding Remarks

Risk-taking fundamentally constrains how a principal motivates her agents. This paper argues that risk-taking blunts convex incentives, which has significant effects on optimal incentive provision. Apart for Corollary 1, the agent does not engage in risk-taking under our optimal contract. Therefore, our analysis focuses on the *incentive* costs of risk-taking, rather than any *direct* costs that risk-taking has on society.

Nevertheless, our framework provides a natural starting point to consider why contracts might not deter risk-taking. Corollary 1 suggests one reason: the principal might be risk-seeking, for instance because her own incentives are non-concave. A second reason is implicit in our assumption that the principal can commit to an incentive scheme. Commitment might be difficult in some settings, for instance because output can serve as the basis for future compensation (Chevalier and Ellison (1997); Makarov and Plantin (2015)). More generally, an agent's competitive context shapes the incentives they face, which in turn determine the kinds of risks they optimally pursue; see Fang and Noe (2015) for a step in this direction. A further analysis of how competition shapes incentives could shed more light on risk-taking behavior in both financial and product markets.

References

Antic, N., 2016. Contracting with unknown technologies. Working Paper.

Aumann, R.J. and Perles, M., 1965. A variational problem arising in economics. *Journal of Mathematical Analysis and Applications*, 11, pp.488-503.

Beesack, P.R., 1957. A note on an integral inequality. *Proceedings of the American Mathematical Society*, 8(5), pp.875-879.

Brav, A., Graham, J.R., Harvey, C.R. and Michaely, R., 2005. Payout policy in the 21st century. *Journal of Financial Economics*, 77(3), pp.483-527.

Borell, C., 1975. Convex Set Functions in d-Space. *Periodica Mathematica Hungarica* 6, pp. 111-136.

Brown, K.C., Harlow, W.V. and Starks, L.T., 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance*, 51(1), pp.85-110.

Carroll, G., 2015. Robustness and linear contracts. *American Economic Review*, 105(2), pp. 536-563.

Chade, H. and Swinkels, J., 2016. The No-Upward-Crossing Condition and the Moral Hazard Problem.

Chassang, S., 2013. Calibrated incentive contracts. *Econometrica*, 81(5), pp. 1935-1971.

Chevalier, J. and Ellison, G., 1997. Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 105(6), pp. 1167-1200.

DeMarzo, P.M., Livdan, D. and Tchistyi, A., 2013. Risking other people's money: Gambling, limited liability, and optimal incentives. Working Paper.

Diamond, P., 1998. Managerial incentives: on the near linearity of optimal compensation. *Journal of Political Economy*, 106(5), pp. 931-957.

Edmans, A. and Gabaix, X., 2011. The effect of risk on the CEO market. *Review of Financial Studies*, 24(8), pp. 2822-2863.

de Figueiredo Jr, R.J., Rawley, E. and Shelef, O., 2014. Bad Bets: Excessive Risk-Taking, Convex Incentives, and Performance.

Fang, D. and Noe, T.H., 2015. Skewing the Odds: Strategic Risk Taking in Contests.

Federal Reserve Press Release. *Date:* October 22, 2009.

Garicano, L. and Rayo, L., 2016. Why Organizations Fail: Models and Cases. *Journal of Economic Literature*, 54(1), pp. 137-192.

Georgiadis, G. and Szentes, B., 2018. Optimal Monitoring Design.

Grossman, S. and Hart, O.D., 1983. An Analysis of the Principal-Agent Problem. *Econometrica,* 51(1), pp. 7-45.

Halac, M. and Prat, A., 2016. Managerial attention and worker performance. *American Economic Review* 106(10), pp. 3104-3132.

Hart, O.D. and Holmström, B., 1986. *The Theory of Contracts*. Department of Economics, Massachusetts Institute of Technology.

Hébert, B., 2015. Moral hazard and the optimality of debt. Working Paper.

Hellwig, M., 2009. A reconsideration of the Jensen-Meckling model of outside finance. *Journal of Financial Intermediation,* 18, pp. 495-525.

Holmström, B., 1979. Moral hazard and observability. *Bell Journal of Economics*, pp. 74-91.

Holmström, B. and Milgrom, P., 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, pp. 303-328.

Holmström, B. and Ricart i Costa, J., 1986. Managerial incentives and capital management. *The Quarterly Journal of Economics* 101(4), pp. 835-860.

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

Jensen, M.C. and Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), pp. 305-360.

Jewitt, I., 1988. Justifying the first-order approach to principal-agent problems. *Econometrica*, pp. 1177-1190.

Jewitt, I., Kadan, O., and Swinkels, J., 2008. Moral hazard with bounded payments. *Journal of Economic Theory,* pp. 59-82.

Kamenica, E. and Gentzkow, M., 2011. Bayesian persuasion. *American Economic Review*, 101(6), pp. 2590-2615.

Lambert, R., 1986. Executive effort and selection of risky projects. *Rand Journal of Economics* 17 (1), pp. 77-88.

Larkin, I., 2014. The cost of high-powered incentives: Employee gaming in enterprise software sales. *Journal of Labor Economics*, 32(2), pp. 199-227.

Malcomson, J., 2009. Principal and expert agent. *The B.E. Journal of Theoretical Economics* 9(1), pp. 1-36.

Makarov, I. and Plantin, G., 2015. Rewarding trading skills without inducing gambling. *Journal of Finance*, 70(3), pp. 925-962.

Matta, E. and Beamish, P.W., 2008. The accentuated CEO career horizon problem: Evidence from international acquisitions. *Strategic Management Journal*, 29(7), pp. 683-700.

Mirrlees, J.A., 1976. The optimal structure of incentives and authority within an organization. *Bell Journal of Economics*, pp. 105-131.

Oyer, P., 1998. Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality. *Quarterly Journal of Economics*, pp. 149-185.

Palomino, F. and Prat, A., 2003. Risk taking and optimal contracts for money managers. *RAND Journal of Economics*, pp. 113-137.

Prekopa, A., 1973. On Logarithmic Concave Measures and Functions. *Acta Sci. Math. (Szeged)* 34, pp. 335-343.

Rajan, R.G., 2011. Fault lines: How hidden fractures still threaten the world economy._Princeton University Press.

Rahmandad, H., Henderson, R., and Repenning, N.P., 2015. Making the numbers? "Short Termism" & the puzzle of only occasional disaster. *Management Science.*

Ray, D., and Robson, A., 2012. Status, intertemporal choice, and risk-taking. *Econometrica* 80(4), pp. 1505-1531.

Robson, A., 1992. Status, the distribution of wealth, private and social attitudes to risk. *Econometrica 60(4),* pp. 837-857.

Spulber, D., and Poblete, J., 2012. The form of incentive contracts: agency with moral hazard, risk neutrality, and limited liability. *Rand Journal of Economics* 43(2), pp. 215-234.

Vereshchagina, G. and Hopenhayn, H.A., 2009. Risk taking by entrepreneurs. *American Economic Review*, 99(5), pp. 1808-1830.

A Proofs for Sections 3 and 4

For notational convenience, we use the indefinite integral to indicate an integral on $[\underline{y}, \bar{y}]$ in all of the appendices.

A.1 Proof of Proposition 1

Fix $a \geq 0$, and let $v(\cdot)$ implement a at maximum profit. We first claim that following each realization x, the agent's payoff equals $v^c(x)$ and the principal's payoff is no larger than $\pi(x - \hat{v}^c(x))$.

Fix $x \in \mathcal{Y}$. Since v is upper semicontinuous, there exists $p \in [0,1]$ and $z_1, z_2 \in \mathcal{Y}$ such that $pz_1 + (1-p)z_2 = x$ and $pv(z_1) + (1-p)v(z_2) = v^c(x)$. Since the agent can choose $\tilde{G}_x$ to assign probability p to z_1 and $1-p$ to z_2, his expected equilibrium payoff satisfies $E_{G_x}[v(y)] \geq v^c(x)$. But v^c is concave and $v^c(y) \geq v(y)$ for any $y \in \mathcal{Y}$, so by Jensen's Inequality $E_{G_x}[v(y)] \leq E_{G_x}[v^c(y)] \leq v^c(E_{G_x}[y]) = v^c(x)$. So $E_{G_x}[v(y)] = v^c(x)$, and hence the contract $v^c(x)$ satisfies (IC_F)-(LL_F) for effort a and the degenerate distribution G.

Next, consider the principal's expected payoff. Since $\pi(\cdot)$ is concave, applying Jensen's Inequality and the previous result yields

$$\begin{aligned} E_{F(\cdot|a)}\left[E_{G_x}\left[\pi(y - u^{-1}(v(y)))\right]\right] &\leq E_{F(\cdot|a)}\left[\pi\left(E_{G_x}\left[y - u^{-1}(v(y))\right]\right)\right] \\ &\leq E_{F(\cdot|a)}\left[\pi\left(x - u^{-1}(v^c(x))\right)\right], \end{aligned}$$

where the first inequality is strict if π is strictly concave and the second is strict if u is strictly concave (so that $-u^{-1}$ is also strictly concave). Therefore, the principal weakly prefers the contract $v^c(x)$, and strictly so if either $\pi(\cdot)$ or $u(\cdot)$ is strictly concave. ■

A.2 Proof of Lemma 1

Existence follows from Proposition 9 in Appendix D. To prove uniqueness, suppose at least one of $\pi(\cdot)$ or $u(\cdot)$ is strictly concave, and suppose that two contracts $v(\cdot)$ and $\tilde{v}(\cdot)$ both implement $a \geq 0$ at maximum profit, with $v(x) \neq \tilde{v}(x)$

for some $x \in \mathcal{Y}$. Since $v(\cdot)$ and $\tilde{v}(\cdot)$ are upper semi-continuous and concave, they must differ on an interval of positive length. But then the contract $v^*(\cdot) \equiv \frac{1}{2}(v(\cdot) + \tilde{v}(\cdot))$ satisfies (IC_F)-(LL_F) for effort a, and the principal's payoff under v^* is

$$\mathbb{E}_{F(\cdot|a)}\left[\pi(y - u^{-1}(v^*(y)))\right] \geq \mathbb{E}_{F(\cdot|a)}\left[\pi\left(y - \tfrac{1}{2}\left(u^{-1}(v(y)) + u^{-1}(\tilde{v}(y))\right)\right)\right] \geq \\ \tfrac{1}{2}\mathbb{E}_{F(\cdot|a)}\left[\pi(y - u^{-1}(v(y)))\right] + \tfrac{1}{2}\mathbb{E}_{F(\cdot|a)}\left[\pi(y - u^{-1}(\tilde{v}(y)))\right],$$

by Jensen's Inequality, where at least one of the inequalities is strict.■

A.3 Proof of Proposition 2

For any contract s, write $U(s) = \max_a \left\{\mathbb{E}_{F(\cdot|a)}[s(y)] - c(a)\right\}$. Fix an optimal pair (a^*, s^*) where $s^*(\cdot)$ implements a^*. Recall that for each a, s_a^L is the lowest-cost linear contract that implements a, and that $s_{a^{FB}}^L$ has slope 1.

Assume first that $U(s^*) \geq U(s_{a^{FB}}^L)$. Then

$$\begin{aligned}
\mathbb{E}_{F(\cdot|a^*)}\left[\pi(y - s^*(y))\right] &\leq \pi\left(\mathbb{E}_{F(\cdot|a^*)}\left[y - s^*(y)\right]\right) \\
&= \pi\left(a^* - \mathbb{E}_{F(\cdot|a^*)}\left[s^*(y)\right]\right) \\
&= \pi\left(a^* - c(a^*) - \left(\mathbb{E}_{F(\cdot|a^*)}\left[s^*(y)\right] - c(a^*)\right)\right) \\
&\leq \pi\left(a^{FB} - c(a^{FB}) - \left(\mathbb{E}_{F(\cdot|a^{FB})}\left[s_{a^{FB}}^L(y)\right] - c(a^{FB})\right)\right) \\
&= \pi\left(\mathbb{E}_{F(\cdot|a^{FB})}\left[y - s_{a^{FB}}^L(y)\right]\right) \\
&= \mathbb{E}_{F(\cdot|a^{FB})}\left[\pi(y - s_{a^{FB}}^L(y))\right].
\end{aligned}$$

The first inequality is Jensen's, and is strict unless either $y - s^*(y)$ is constant or the principal is risk neutral. The second inequality uses $U(s^*) \geq U(s_{a^{FB}}^L)$ and $a^* - c(a^*) \leq a^{FB} - c(a^{FB})$, and is strict unless $a^* = a^{FB}$ and $U(s^*) = U(s_{a^{FB}}^L)$. The final equality uses that $y - s_{a^{FB}}^L(y)$ is a constant. For (a^*, s^*) to be optimal, these inequalities must hold with equality, so $a^* = a^{FB}$, $s_{a^{FB}}^L(\cdot)$ is optimal, and moreover $s^* = s_{a^{FB}}^L$ if the principal is risk averse.

Assume instead that $U(s_{a^{FB}}^L) > U(s^*)$. Then, since $U(s^*) \geq u_0$, it follows that $s_{a^{FB}}^L(\underline{y}) = -M$. For each a, let $\hat{s}_a(\cdot)$ be the linear contract $\hat{s}_a(y) = s^*(\underline{y}) + c'(a)(y - \underline{y})$ that equals $s^*(\underline{y})$ at $\underline{y}$ and implements a. Note that $\hat{s}_{a^{FB}}(y) \geq$

$s^L_{a^{FB}}(y)$ for any y, so $U(\hat{s}_{a^{FB}}) \geq U(s^L_{a^{FB}}) > U(s^*)$.

We claim that $U(\hat{s}_{a^*}) \leq U(s^*)$. To see this, define $\hat{u}$ so that

$$\int (\hat{s}_{a^*}(y) - (s^*(y) + \hat{u}))f(x|a^*)dx = 0 \tag{8}$$

and suppose to the contrary that $\hat{u} > 0$. Then, since $\hat{s}_{a^*}(\underline{y}) < s^*(\underline{y}) + \hat{u}$, and since $\hat{s}_{a^*}(\cdot)$ is linear and $s^*(\cdot) + \hat{u}$ is concave, there exists $\tilde{y} > \underline{y}$ such that $\hat{s}_{a^*}(\cdot) - (s^*(\cdot) + \hat{u})$ is strictly negative below $\tilde{y}$ and strictly positive above $\tilde{y}$. Hence, since $\frac{f_a(\cdot|a^*)}{f(\cdot|a^*)}$ is strictly increasing, by Beesack's inequality,[20] (8) implies that

$$\begin{aligned} 0 &< \int (\hat{s}_{a^*}(y) - (s^*(y) + \hat{u})) \frac{f_a(y|a^*)}{f(y|a^*)} f(y|a^*)dx \\ &= \int (\hat{s}_{a^*}(y) - s^*(y)) f_a(y|a^*)dy \end{aligned}$$

where the equality uses that $\int f_a(y|a^*)dy = 0$. This contradicts that both $\hat{s}_{a^*}$ and s^* implement a^*, and so $U(\hat{s}_{a^*}) \leq U(s^*)$.

Since $U(\hat{s}_a)$ is continuous in a and $U(\hat{s}_{a^{FB}}) > U(s^*) \geq U(s_{a^*})$, there exists $\hat{a} \in [a^*, a^{FB})$ such that $U(\hat{s}_{\hat{a}}) = U(s^*)$. Since $s^L_{\hat{a}}$ is weakly below $\hat{s}_{\hat{a}}$,

$$\begin{aligned} \mathbb{E}_{F(\cdot|a^*)}\left[s^L_{\hat{a}}(y)\right] &\leq \mathbb{E}_{F(\cdot|a^*)}\left[\hat{s}_{\hat{a}}(y)\right] \\ &= \mathbb{E}_{F(\cdot|\hat{a})}\left[\hat{s}_{\hat{a}}(y)\right] - \int_{a^*}^{\hat{a}} \left(\frac{\partial}{\partial a}\mathbb{E}_{F(\cdot|a)}\left[\hat{s}_{\hat{a}}(y)\right]\right) da \\ &= \mathbb{E}_{F(\cdot|\hat{a})}\left[\hat{s}_{\hat{a}}(y)\right] - c'(\hat{a})(\hat{a} - a^*) \\ &= U(\hat{s}_{\hat{a}}) + c(\hat{a}) - c'(\hat{a})(\hat{a} - a^*) \\ &\leq U(\hat{s}_{\hat{a}}) + c(a^*) \\ &= U(s^*) + c(a^*) \\ &= \mathbb{E}_{F(\cdot|a^*)}\left[s^*(y)\right]. \end{aligned}$$

[20] The relevant version of Beesack's inequality states that if a function $h(\cdot)$ single-crosses 0 from below and satisfies $\int h(x)dx = 0$, then for any increasing function $g(\cdot)$, $\int h(x)g(x)dx \geq 0$, and strictly so if $g(\cdot)$ is strictly increasing and $h(\cdot)$ is not everywhere 0. See Beesack (1957), available online at https://www.jstor.org/stable/2033682.

Here, the second equality uses that $\mathbb{E}_{F(\cdot|a)}[\hat{s}_{\hat{a}}(y)]$ is linear in a and that $\hat{s}_{\hat{a}}(\cdot)$ implements $\hat{a}$, and the second inequality uses that $c(\cdot)$ is convex.

Choose $\hat{y}$ so that $s^L_{\hat{a}}(\cdot)$ crosses the concave contract $s^*(\cdot)$ from below at $\hat{y}$, where if $s^L_{\hat{a}}(y) < s^*(y)$ for all y, then $\hat{y} = \bar{y}$. Since $\hat{a} < a^{FB}$, and hence $s^L_{\hat{a}}(\cdot)$ has slope strictly less than 1, it follows that for all $y < \hat{y}$ and $t > s^L_{\hat{a}}(y)$,

$$\pi'(y - t) \geq \pi'(y - s^L_{\hat{a}}(y)) \geq \pi'(\hat{y} - s^L_{\hat{a}}(\hat{y})),$$

and strictly so if $\pi(\cdot)$ is not linear. Similarly, for all $y > \hat{y}$ and $t < s^L_{\hat{a}}(y)$,

$$\pi'(y - t) \leq \pi'(y - s^L_{\hat{a}}(y)) \leq \pi'(\hat{y} - s^L_{\hat{a}}(\hat{y})),$$

and strictly so if $\pi(\cdot)$ is not linear. That is, the marginal cost to the principal of paying the agent is no less than $\pi'(\hat{y} - s^L_{\hat{a}}(\hat{y}))$ for $y < \hat{y}$, and no more than this amount for $y > \hat{y}$.

But then, since $\mathbb{E}_{F(\cdot|a^*)}\left[s^L_{\hat{a}}(y)\right] \leq \mathbb{E}_{F(\cdot|a^*)}[s^*(y)]$ and $s^L_{\hat{a}}(y) < s^*(y)$ if and only if $y < \hat{y}$,

$$\mathbb{E}_{F(\cdot|a^*)}\left[\pi(y - s^L_{\hat{a}}(y))\right] \geq \mathbb{E}_{F(\cdot|a^*)}\left[\pi(y - s^*(y))\right],$$

and strictly so unless the principal is risk neutral, or $s^L_{\hat{a}}(\cdot)$ and $s^*(\cdot)$ agree. Finally, since the slope of $s^L_{\hat{a}}(\cdot)$ is strictly less than 1 and $\hat{a} \geq a^*$,

$$\mathbb{E}_{F(\cdot|\hat{a})}\left[\pi(y - s^L_{\hat{a}}(y))\right] \geq \mathbb{E}_{F(\cdot|a^*)}\left[\pi(y - s^L_{\hat{a}}(y))\right],$$

and strictly so unless $\hat{a} = a^*$.

To conclude the proof, note that since (a^*, s^*) is optimal, each of these inequalities is an equality, and hence $a^* = \hat{a} \leq a^{FB}$. If the principal is risk averse, then $s^* = s^L_{\hat{a}}$ as well. If the principal is risk neutral, then $s^L_{\hat{a}}(\cdot)$ is optimal but not uniquely so.■

A.4 Proof of Corollary 1

Fix $a > 0$ and consider the problem (Obj$_F$)-(LL$_F$) with an arbitrary $\pi(\cdot)$ and $u(s) \equiv s$. Define $\mathbb{E}_{G_x}[\pi(y)] = \pi^c(x)$, where $\pi^c(\cdot)$ denotes the concave closure of $\pi(\cdot)$.

Modify (Obj)-(NG) so that the principal's utility equals $\pi^c(\cdot)$. Since $\pi^c(y) \geq \pi(y)$ for any y, so the principal's payoff in this modified problem must be weakly larger than under the original problem. But $\pi^c(\cdot)$ is concave and $s_a^L(\underline{y}) = -M$, so Proposition 2 implies that $s_a^L(\cdot)$ implements a at maximum profit in this modified problem. So the principal's expected payoff equals $\mathbb{E}_{F(\cdot|a)}\left[\pi^c(x - s_a^L(x))\right]$ in this modified problem.

Now, consider the contract $s_a^L(x)$ in the original problem (Obj)-(NG). For any distribution $G_x \in \Delta(\mathcal{Y})$ such that $\mathbb{E}_{G_x}[y] = x$, $\mathbb{E}_{G_x}\left[y - s_a^L(y)\right] = x - s_a^L(x)$ because s_a^L is linear. Therefore, as in Proposition 1, there exists some G_x^P such that $\mathbb{E}_{G_x^P}\left[\pi(y - s_a^L(y))\right] = \pi^c(x - s_a^L(x))$. Furthermore, conditional on x, the agent's expected payoff satisfies $\mathbb{E}_{G_x}\left[s_a^L(y) - c(a)\right] = s_a^L(x) - c(a)$ for any G_x with $\mathbb{E}_{G_x}[y] = x$. So $s_a^L(\cdot)$ satisfies (IC$_F$)-(LL$_F$) for $a > 0$ and $G_x = G_x^P$ for each $x \in \mathcal{Y}$. The principal's expected payoff if she offers s_a^L equals $\mathbb{E}_{F(\cdot|a)}\left[\pi^c(x - s_a^L(x))\right]$, her payoff from the modified problem. So s_a^L *a fortiori* implements a at maximum profit for any $a \geq 0$. ■