Optimal Contracts with a Risk-Taking Agent*

Daniel Barron, George Georgiadis, and Jeroen Swinkels†

June 21, 2017

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, optimal incentives are strikingly simple: linear contracts maximize profit. If the agent is risk averse, we characterize the unique profit-maximizing contract and show how deterring risk-taking affects the insurance-incentive tradeoff. We extend our model to analyze costly risk-taking and alternative timings, and reinterpret our model as a dynamic setting in which the agent can manipulate the timing of output.

*We are grateful to Simon Board, Ralph Boleslavsky, Luis Cabral, Gabriel Carroll, Hugo Hopenhayn, Johannes Hörner, Eddie Lazear, Erik Madsen, Niko Matouschek, 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, 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, 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, which can have 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 badly 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 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); Repenning and Henderson (2010); 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 extends a canonical moral hazard setting to explore how a principal motivates an agent who can engage in risk-taking. In our setting, a principal offers a contract to 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 the final contractible output.

Echoing Jensen and Meckling (1976) and others, we argue that the possibility of risk-taking constrains the types of incentives that can be used to motivate an agent, with potentially significant effects on profits, productivity, and welfare. In particular, Section 3 begins with the observation that the agent takes 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, this additional risk-taking is costly, so the principal finds it optimal to deter risk-taking entirely. She accomplishes this by offering an incentive scheme that directly makes the agent's utility concave in output. Motivated by this logic, we consider how the principal optimally motivates the agent under a **no-gaming constraint** that requires the agent's utility to be weakly concave in output.

In Section 4, we consider the case in which the agent is risk-neutral. We prove that linear (technically, affine) contracts are optimal, remain so regardless of the principal's attitude toward risk (even if she is risk-loving), and are uniquely optimal if the principal is risk averse. Intuitively, if we ignored the no-gaming constraint, the principal would like to offer the agent a convex contract to both respect the limited liability constraint and concentrate high pay on high outcomes, which are most indicative of high effort. However, such a contract would violate the no-gaming constraint and so induce risk-taking. Instead, the principal optimally offers the most convex contract that satisfies the no-gaming constraint, which is linear. We show that relative to any strictly concave contract, a linear contract both better insures the principal and better motivates the agent.

If the agent is risk averse, then the optimal incentive scheme must motivate the agent while reflecting the fact that exposing him to risk is costly. Section 5 focuses on this case, restricting attention for simplicity to a risk-neutral principal. Since the agent does not like risk, a convex contract does not necessarily induce inefficient risk-taking, so long as that contract makes the agent's *utility* concave in output. We give necessary and sufficient conditions that characterize the unique profit-maximizing contract that implements any given positive effort. Whenever the no-gaming constraint binds, this contract is convex enough to make the agent's utility linear in output, but this constraint might bind only following some outputs or not at all.

Analogous to a risk-neutral agent, the no-gaming constraint binds if, absent that constraint, the optimal contract would make the agent's utility non-concave. For instance, if the individual rationality constraint does not bind (so that the limited liability constraint must), then the optimal contract without risk-taking must entail a non-concave region over low output (Jewitt, Kadan, and Swinkels (2008)). Consequently, the optimal contract in our setting is linear in utility for low output. The no-gaming constraint is also likely to bind—and hence the profit-maximizing incentive scheme is likely to be linear in utility and strictly convex in money—if the principal would like to insure the agent against "downside risk" by offering flat incentives over low output. We develop conditions that capture this intuition and depend on both the likelihood ratio of output, which captures how higher pay at any given output affects effort incentives, and on the agent's relative attitudes towards upside and downside risk.

Our characterization builds on Mirrlees (1976) and Holmström (1979), who show that, absent the no-gaming constraint, the optimal contract balances outcome-by-outcome the marginal cost of giving the agent higher utility to the marginal benefit of doing so, as measured by the shadow values of the agent's participation and incentive constraints, and by the likelihood ratio at that outcome. However, doing so might violate the no-gaming constraint, and so we cannot follow the same logic. Instead, we construct two simple perturbations of an incentive scheme that increase either the level or the slope of the agent's utility over a suitable interval while preserving concavity. As in Mirrlees and Holmström, the profit-maximizing contract balances the benefit and cost of any such perturbation, but now does so in expectation, where both the expected marginal cost of giving the agent higher utility and the expected likelihood ratio are with respect to that perturbation. Perhaps surprisingly, we prove that it is also sufficient to consider these two simple 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 principal and agent are risk-neutral. First, we alter the timing of our model so that the agent engages in risk-taking before he observes the intermediate output. For example, an entrepreneur might be able to adjust the riskiness of a project only before she learns whether it will bear fruit. Given this timing, the agent's risk-taking concavifies his expected payoff conditional on his *effort*, rather than on intermediate output. Under mild conditions, we show that a linear contract is again optimal in this setting.

Second, we consider optimal contracts if the agent incurs a cost to engage in risk-taking that is increasing in the variance of that risk. We show that this contracting problem can be reformulated as a simple variant of our baseline model, so that our basic intuition extends. The unique optimal contract is convex and converges to a linear contract as the cost of taking on risk converges to zero.

Our third extension reinterprets our baseline model as a dynamic setting in which the principal offers a stationary contract that the agent can game by shifting output over time. Oyer (1998) and Larkin (2014) empirically document how convex incentive schemes and long sales cycles can encourage such gaming. We assume that the agent's effort generates output, but he can costlessly manipulate *when* that output is realized over an interval of time. We show that this model is equivalent to our baseline setting; in particular, a linear contract is optimal, since a convex contract would induce the agent to bunch sales over short time intervals, while a strictly concave contract would provide subpar effort incentives.

Our analysis is inspired in part by Diamond (1998), which uses several examples to argue that linear contracts are approximately optimal if the agent has sufficient control over the distribution of output, and in particular are exactly optimal if the agent can choose *any* distribution over output such that expected output equals effort. Relative to that paper, our model puts different constraints on the agent's risk-taking by assuming that he can add mean-preserving noise to an exogenous distribution, but cannot eliminate all randomness in output. We cleanly characterize how risk-taking constrains incentives in a flexible but tractable framework, which allows us to analyze optimal contracts for more general risk preferences, output distributions, and effort functions. Garicano and Rayo (2016) similarly builds on Diamond (1998) but fixes an exogenous (convex) contract to focus on the social costs of agent risk-taking.

Palomino and Prat (2003) considers a delegated portfolio management problem in which the agent chooses both expected output and the riskiness of the portfolio (in the sense of second-order stochastic dominance) from a parametric family of distributions. The resulting optimal contract consists of a base salary and a fixed bonus that is paid whenever output exceeds a threshold that is determined by features of the parametric family. In contrast, our agent can choose any mean-preserving spread, which means that our optimal contract must deter a more flexible form of gaming. Hébert (2015) finds conditions such that, if an agent can manipulate the output distribution at a cost that depends on the difference between that distribution and some exogenous baseline, then the optimal contract resembles a debt contract. Demarzo, Livdan, and Tchistyi (2014) characterizes the optimal contract if the agent can take socially inefficient risks in a dynamic setting. It argues that backloading can mitigate, but not necessarily eliminate, the agent's incentive to take such risks. Makarov and Plantin (2015) considers a model of career concerns in which the agent can take excessive risk to temporarily manipulate the principal's beliefs about her ability, and characterizes a backloaded incentive scheme that eliminates these incentives. Attar, Mariotti, and Salanie (2011) argues that linear contracts are optimal in a market with adverse selection if the seller can privately sign contracts for additional products.

More broadly, our work is related to a long-standing literature which argues that real-world contracts are simple in order to deter gaming. Holmström and Milgrom (1987) displays a dynamic environment in which linear contracts are optimal but notes that the point that linear contracts are robust to gaming "is not made as effectively as we would like by our model; we suspect that it cannot be made effectively in any traditional Bayesian model." Recent papers, including Chassang (2013), Carroll (2015), and Antic (2016), take up this argument 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 our paper. However, if the principal offers a concave incentive scheme, then Nature might 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 timing is as follows:

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

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 $F(\cdot|a) \in \Delta(\mathcal{Y})$, where (i) F is infinitely differentiable with density f, (ii) F has full support for

[1] 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.

all $a \geq 0$, (iii) f is strictly MLRP-increasing in a with $\frac{f_a(\cdot|a)}{f(\cdot|a)}$ uniformly bounded for all a, and (iv) $\mathbb{E}_{F(\cdot|a)}[x] = a$.

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

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. We also assume $c(\cdot)$ is infinitely differentiable, strictly increasing, and strictly convex.

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

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

We can treat the agent as choosing a and $G \in \mathcal{G}$ simultaneously, since the agent chooses each G_x to maximize his ex-ante expected payoff conditional on x.

After the agent observes x but before y is realized, this is a problem 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.[2]

[2] With a risk-neutral principal and agent, and a non-binding participation constraint, Online Appendix E shows that a linear contract remains optimal even if the principal can demand a report after x but before y is realized.

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 final output y. If we define $\underline{u} \equiv u(-M)$, then an optimal contract solves the following constrained 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 Lemma 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.

Lemma 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. The intuition is simple. 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. Note that 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 Lemma 1, we henceforth consider contracts that satisfy 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 the principal and agent are risk-neutral. If $v(\cdot)$ is constrained to be concave, then the agent will optimally choose G^D. Therefore, an optimal incentive scheme solves the simplified problem:

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

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

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

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

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

For a fixed effort $a \geq 0$, we say that $v(\cdot)$ *implements* a if it satisfies (IC)-(Conc) for a, and does so *at maximum profit* if it maximizes (Obj) subject to (IC)-(Conc). 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 2. *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 7 in Online Appendix D.[3] The no-gaming constraint (Conc) is required for existence to be guaranteed; for example, without this constraint,

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

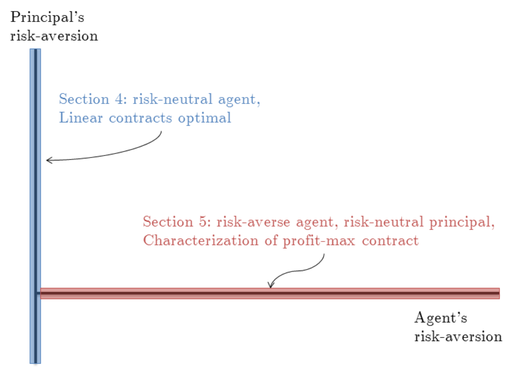

Figure 1: Roadmap

no profit-maximizing contract would exist with a risk-neutral agent.[4] 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.

Roadmap of our Main Results

The next two sections consider how the no-gaming constraint affects optimal incentives. These sections are designed to be modular so that the interested reader can focus on the results that most interest them. We offer here a roadmap that describes how these analyses connect to one another.

Consider Figure 1. Both principal and agent are risk-neutral at the origin. Moving right makes the agent more risk averse, while moving upwards or downwards makes the principal more risk-averse or more risk-seeking, respectively.

Section 4 explores contracting with a risk-neutral agent, which corresponds to the vertical axis on Figure 1. Regardless of the principal's risk preferences, a linear contract is optimal in this setting, and it is uniquely so if the principal is strictly risk-averse (above the origin). Section 5 considers a risk-neutral principal and a risk-averse agent (the horizontal axis in Figure 1). Here, we develop necessary and sufficient conditions that characterize the unique contract that

[4]With a risk-neutral agent, the principal wants to pay the agent only after very high output, since that output is most indicative of high effort. See, e.g., Innes (1990).

implements any $a \geq 0$ at maximum profit. We give conditions under which (Conc) binds for some output, so that the agent's utility is either partially or fully linear in output. For instance, if (IR) is slack (so that (LL) binds), then we show that the agent's payoff under the profit-maximizing contract is linear at any output where the likelihood ratio is negative.

So long as at least one player is strictly risk-averse (the upper quadrant excluding the origin in Figure 1), the Theorem of the Maximum implies that the (unique) contract that implements a at maximum profit is continuous in the topology of almost everywhere pointwise convergence. See Proposition 7 in Online Appendix D for a proof. Therefore, the characterizations in Sections 4 and 5 shed light on profit-maximizing incentives near the two axes as well. For instance, if the principal is risk-averse, the agent is approximately risk-neutral, and effort is fixed at the optimal effort level for a risk-neutral agent, then the profit-maximizing contract is approximately linear. Similarly, if the agent is risk-averse and the principal is approximately risk-neutral, then profit-maximizing incentives approximate the characterization in Section 5. Continuity does not extend to the lower quadrant, so we should be cautious about applying our intuition if the principal is risk-seeking and the agent is risk-averse.

4 Optimal Contracts for a Risk-Neutral Agent

Suppose the agent is risk-neutral, so $u(y) = y$, $v(\cdot) = s(\cdot)$, and $\underline{u} = -M$. 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

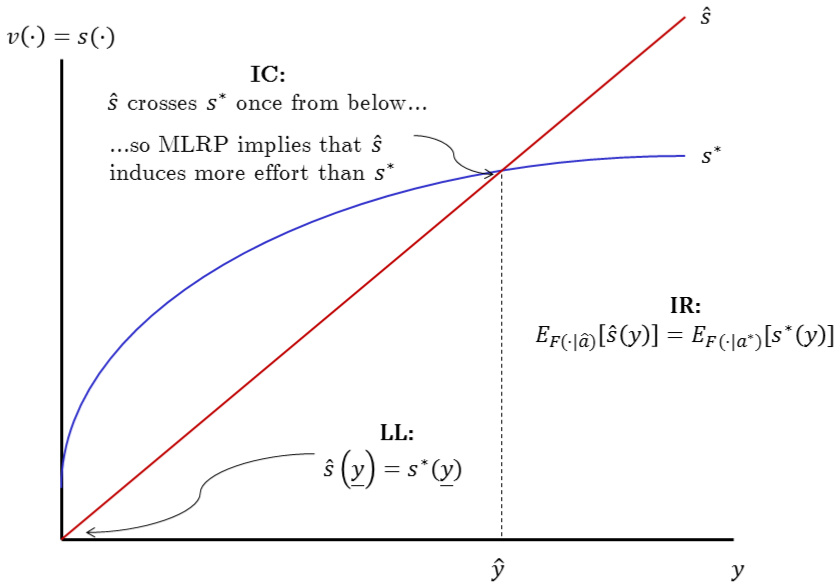

Figure 2: Intuition for the proof of Proposition 1.

$y - c(y)$ 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 1. *Let $u(s) \equiv s$. If a^* is optimal, then $a^* \leq a^{FB}$ and $s^L_{a^*}(\cdot)$ is optimal.*

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

Suppose instead that (IR) is slack for $s^L_{a^{FB}}(\cdot)$, in which case (LL) must bind. Consider any optimal $(a^*, s^*(\cdot))$. Define $\hat{s}(\cdot)$ as the linear contract that agrees with $s^*(\cdot)$ at $\underline{y}$ and gives the agent the same utility if he chooses effort optimally under $\hat{s}(\cdot)$. As shown in Figure 2, $\hat{s}(\cdot)$ must single-cross $s^*(\cdot)$ from below, effectively moving payments from low to high outputs. Since $F(\cdot|a)$ satisfies MLRP, we can show that $\hat{s}(\cdot)$ implements effort $\hat{a} \geq a^*$. Assume that $\hat{a} \geq a^{FB}$. Then $\hat{s}(\cdot) \geq s^L_{a^{FB}}(\cdot)$, and so the principal prefers to offer the agent $s^L_{a^{FB}}(\cdot)$ because it induces first-best effort, perfectly insures the principal, and gives the agent less utility than $s^*(\cdot)$, contradicting that $s^*(\cdot)$ is optimal.

Let $\hat{y}$ be the point at which $\hat{s}(\cdot)$ and $s^*(\cdot)$ cross. Then, since $\hat{a} < a^{FB}$, $\hat{s}'(\cdot) < 1$ and so the principal is richer under both $\hat{s}(\cdot)$ and $s^*(\cdot)$ everywhere to the right of $\hat{y}$ than anywhere to the left. The principal's marginal utility is decreasing in her income, and so holding effort fixed at a^*, the principal is happy that $\hat{s}(\cdot)$ pays less than $s^*(\cdot)$ when she is relatively poor and pays more when she is relatively rich. Furthermore, increasing effort from a^* to $\hat{a} < a^{FB}$ increases expected output and so benefits the principal because $\hat{s}'(\cdot) < 1$. Since $s^L_{\hat{a}}(\cdot)$ is parallel to but weakly below $\hat{s}(\cdot)$, the principal is better off still with $s^L_{\hat{a}}(\cdot)$. So for $s^*(\cdot)$ to be optimal, it must be that $s^*(\cdot) \equiv s^L_{a^*}(\cdot)$ and $\hat{a} = a^*$.

Lemma 2 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.

In some applications, the principal might have risk-seeking preferences over output, for instance because she also faces convex incentives. For example, Rajan (2011) notes that investment funds and banks were motivated to take on significant amounts of risk in the years leading up to the 2008 financial crisis, while Chevalier and Ellison (1997) argue that mutual funds earn disproportionate profits from outperforming their competitors, which encourages excessive risk-taking.

We can model such settings by allowing $\pi(\cdot)$ to be any strictly increasing and continuous function. Lemma 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 1 to show that a linear contract is optimal.

Proposition 2. *Let $u(s) \equiv s$ and let $\pi(\cdot)$ be an arbitrary continuous and strictly increasing function. If a^* is optimal, then $a^* \leq a^{FB}$ and $s^L_{a^*}(\cdot)$ is optimal.*

As in (1), define $\pi^c(\cdot)$ as the concave closure of $\pi(\cdot)$. To see the proof of Proposition 2, note that the principal's expected payoff cannot exceed $\pi^c(\cdot)$ for reasons similar to Lemma 1. Therefore, the contract that maximizes

$E_{F(\cdot|a)}\left[\pi^c(x - s(x))\right]$ subject to (IC)-(Conc) provides an upper bound on the principal's payoff. But Proposition 1 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 $\pi^c(\cdot)$.

Our final result in this section considers how a^* changes with the lower bound $\underline{y}$ on output. A decrease in $\underline{y}$ implies that the agent can take on more severe left-tail risk by gambling over worse outcomes. We prove that a lower $\underline{y}$ makes it costlier for the principal to induce any non-zero effort level. As $\underline{y}$ approaches $-\infty$, inducing any positive effort becomes arbitrarily expensive and so the agent exerts no effort in the optimal contract.

Corollary 1. *Let a^* be the optimal effort level. Then $\lim_{\underline{y}\to-\infty} a^* = 0$, and if $\pi(y) \equiv y$, then a^* is increasing in $\underline{y}$.*

Proposition 1 implies that the principal's expected payment from inducing $a^* \geq 0$ equals $E_{F(\cdot|a^*)}[\pi(y - c'(a^*)(y - \underline{y}) + w)]$. For small enough $\underline{y}$, $s^L_{a^*}(\underline{y}) = -M$. But then implementing $a^* > 0$ becomes arbitrarily costly as $\underline{y} \to -\infty$, in which case the principal is better off not motivating the agent at all. If the principal is risk-neutral, then we can show that the principal's profit under $s^L_{a^*}(\cdot)$ is supermodular in a^* and $\underline{y}$, so that a^* is increasing in $\underline{y}$.

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.

Let $\pi(y) = y$, and define $\underline{w}$ as the infimum of the domain of $u(\cdot)$. We assume that $\lim_{w\downarrow\underline{w}} u'(w) = \infty$ and $\lim_{w\uparrow\infty} u'(w) = 0$. Our first step is to replace (IC) with the weaker condition that local incentives are slack,

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

Given (Conc), replacing (IC) with (IC-FOC) entails no loss if $F(\cdot|\cdot)$ satisfies weak regularity conditions.[5,6] For a fixed effort $a \geq 0$, define the principal's problem (P) as maximizing (Obj) subject to (IC-FOC), (IR), (LL), and (Conc).

If $\underline{u} > -\infty$, then Lemma 2 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. The characterization in this section applies to either setting.

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

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

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 the *net cost of increasing* $v(\cdot)$ *at* y as

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

Intuitively, $n(y)$ represents the marginal cost of increasing $v(y)$ at y, taking into account how that increase affects (IR) and (IC-FOC). In particular, consider increasing $v(y)$. Doing so increases the principal's cost at rate $\rho^{-1}(v(y))f(y|a)$. It relaxes (IR) at rate $f(y|a)$, which has implicit value λ, and similarly 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 $n(y)$.

If the principal could offer non-concave contracts, then the optimal contract

[5] A sufficient condition is that $\int_{\underline{y}}^{z} F_{aa}(y|a)dy \geq 0$ for all $z \in \mathcal{Y}$ and $a \geq 0$. In particular, $E_{F(\cdot|a)}[y] = a$ implies that $\int F_{aa}(y|a)dy = 0$, so this condition holds if $F_{aa}(\cdot|a)$ never changes sign from negative to positive. See Jewitt (1988) and Chade and Swinkels (2016).

[6] For expositional convenience, we use an indefinite integral to denote an integral from $\underline{y}$ to $\overline{y}$.

[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 because $\lim_{w \downarrow \underline{w}} u(w) = -\infty$.

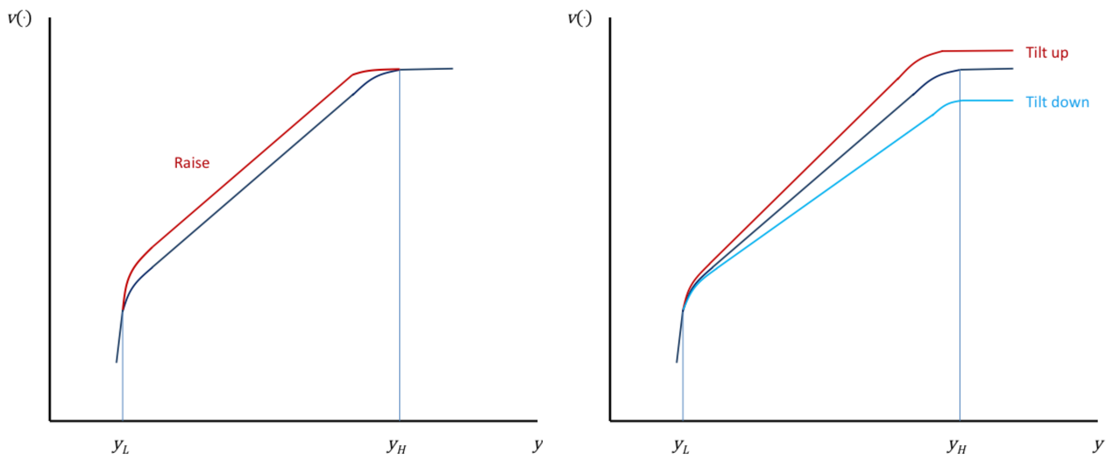

Figure 3: *Raise* and *tilt*. These perturbations require care around y_L and y_H to ensure that concavity is preserved. For this reason, we need both y_L and y_H to be free for *raise*. For *tilt up*, we need y_L to be free, while y_H must be free for *tilt down.*

would set $n(y) = 0$ for all y. Indeed, this incentive scheme (with appropriate λ and μ) is the *Holmström-Mirrlees contract* characterized in Mirrlees (1976) and Holmström (1979). However, this contract may not be concave, so we instead characterize the profit-maximizing contract by identifying a set of perturbations that respect concavity.

Given $v(\cdot)$, 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. Say that y is *free* if it is not on the interior of any linear segment of $v(\cdot)$. Intuitively, if y is free, then $v(\cdot)$ is strictly concave on at least one side of y. The points $\underline{y}$ and $\bar{y}$ are always free.

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 introduces a non-concavity at the lower end of the interval, and *tilting* it a negative amount introduces 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 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).[8]

Consider *raising* $v(\cdot)$ on an interval between two free points $y_L < y_H$. For $v(\cdot)$ to be optimal, the net cost of this perturbation must be positive, or

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

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.

We can make a similar argument using *tilt*. Suppose $y_L < y_H$ with y_L free. Then $v(\cdot)$ is optimal only if it cannot be improved by applying positive *tilt*:

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

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 1. *A contract $v(\cdot)$ is Generalized Holmström-Mirrlees (GHM) if (IC-FOC) holds with equality, (IR)-(Conc) are satisfied, and there exist $\lambda \geq 0$*

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

and $\mu > 0$ *such that*

$$\lambda\left(\int v(y) f\,(y|a)\,dy - c(a) - u_0\right) = 0,$$

and for any $y_L < y_H$,

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

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 incentive scheme that implements any $a > 0$ at maximum profit.

Proposition 3. *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 proofs for all results in this section are in Online Appendix B. 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 perhaps surprising 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.

Proposition 3 highlights how the no-gaming constraint (Conc) affects optimal incentives. This constraint does not always bind; for example, the contract that sets $n(y) = 0$ for all y might satisfy (LL) and (Conc), in which case the Holmström-Mirrlees contract is also profit-maximizing in our setting.[9] More generally, we show that $n(y) = 0$ whenever (Conc) is "slack" near y. For any free and interior $y \in (\underline{y}, \bar{y})$, say y is a *kink point* of $v(\cdot)$ if two linear segments meet at y, and a *point of normal concavity* otherwise.

[9]This case obtains if $\rho(\lambda + \mu l(\cdot|a))$ is concave, for instance because both $\rho(\cdot)$ and $l(\cdot|a)$ are concave, and M is large enough that $\rho(\lambda + \mu l(\underline{y}|a)) \geq \underline{u}$.

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.*

For any point of normal concavity y, (Conc) is "slack" around y in the sense that we can find two free points that are arbitrarily close to y (see Claim 1 in Appendix B). Then Proposition 3 implies that (3) holds with equality between these points. Taking a limit as these points approach y yields $n(y) = 0$. If y is a kink point, then we cannot perturb $v(\cdot)$ around y and preserve concavity. However, there is a sense in which (Conc) binds on the linear segments on either side of y: Lemma 3 in Online Appendix B proves that absent (Conc), 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.

Now, we consider settings in which (Conc) binds for some range of output. Suppose (IR) is slack, which implies that (LL) binds. To relax (IC-FOC), a profit-maximizing contract should pay as little as possible when $l(y|a) < 0$. But paying $-M$ following low output, while giving incentives for higher output, would violate (Conc), so utility is optimally linear in output whenever $l(\cdot|a)$ is negative.

Corollary 3. *Suppose $u(\cdot)$ is strictly concave and $\pi(y) \equiv y$. For any $a > 0$, let $v(\cdot)$ solve (P), and assume that (IR) is slack. Define y_0 by $l(y_0|a) = 0$. Then v is linear on $[\underline{y}, y_0]$.*

If $v(\cdot)$ 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, decreases the principal's expected payment, and is feasible if (IR) is slack.

Next, we identify sufficient conditions under which (Conc) binds everywhere and so the optimal incentive scheme is linear in utility.

Corollary 4. *Suppose $u(\cdot)$ is strictly concave and $\pi(y) \equiv y$. Fix $a \geq 0$, let $v(\cdot)$ solve (P), and let λ and μ be the corresponding shadow values on (IR) and (IC-FOC). If $\rho(\lambda + \mu l(\cdot|a))$ is convex, then $v(\cdot)$ is linear.*

To see this result, note that $n(\hat{y}) \leq 0$ at any interior free $\hat{y}$ by Corollary 2, and so $v(\hat{y}) \leq \rho(\lambda + \mu l(\hat{y}|a))$. But $\rho(\lambda + \mu l(\cdot|a))$ is convex and $v(\cdot)$ is concave, so either $v(y) < \rho(\lambda + \mu l(y|a))$ for all $y > \hat{y}$, or $v(y) > \rho(\lambda + \mu l(y|a))$ for all $y < \hat{y}$. Since $\underline{y}$, $\hat{y}$, and $\bar{y}$ are all free, either violates (3). Hence, $v(\cdot)$ has no interior free points and so must be linear.

If $l(\cdot|a)$ is convex and $\rho(\cdot)$ is convex on the range of $\lambda + \mu l(\cdot|a)$, then $\rho(\lambda + \mu l(\cdot|a))$ is convex.[10] Intuitively, $\rho(\lambda + \mu l(\cdot|a))$ is 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. Such an incentive scheme would be convex in utility, so (Conc) binds and the profit-maximizing $v(\cdot)$ is instead linear. 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*.[11]

Finally, the logic of Corollary 2 implies that in many settings, the profit-maximizing $v(\cdot)$ has at most one linear segment. In particular, (Conc) binds on any linear segment, so $n(\cdot)$ must be negative at the endpoints and positive somewhere in the middle. Therefore, for $v(\cdot)$ to have two linear segments, $\rho(\lambda + \mu l(\cdot|a))$ must have a strictly concave region followed by a weakly convex region, which can be ruled out with reasonably weak conditions.

Given an interval $X \subseteq \mathbb{R}$ and analytic function $h : X \to \mathbb{R}_+$, define the *concavity of h* by $con(h) = \inf_X \left\{1 - (hh^{''})/(h^{'})^2\right\}$. Intuitively, $con(h)$ is the largest value t for which h^t/t is concave.[12]

Corollary 5. *Suppose $con(\rho') + con(l_y) > -1$ and $u(\cdot)$ and $F(\cdot)$ are analytic.*

[10] Note that $\rho(\cdot)$ cannot be convex over its entire domain, because $\rho(0) = -\infty$.

[11] 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.

[12] In particular, $h(\cdot)$ is concave if $con(h) \geq 1$ and is log-concave if $con(h) \geq 0$. See Prekopa (1973) and Borell (1975) for details.

Fix $a > 0$ and let $v(\cdot)$ solve (P). Then $v(\cdot)$ has at most one linear segment, which must begin at $\underline{y}$.

The condition $con(\rho') + con(l_y) > -1$ is satisfied, for instance, if either $u(w) = \log w$ and $\log(l_y)$ is strictly concave or $u(w)$ is *HARA* and $-\frac{1}{l_y^2}$ is strictly concave.[13]

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

This section argues that linear contracts may be optimal even if the agent cannot condition his risk-taking on intermediate output.

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 $\mathbb{E}_G[x|a] = a$.[14]

[13] A utility function $u(w)$ is *HARA* if $-\frac{u'(w)}{u''(w)}$ is linear, as, for example, if $u(w) = \sqrt{w}$.

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

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 density $f(\cdot|x)$ with full support, satisfies strict MLRP in x, and $\mathbb{E}_{F(\cdot|x)}[y] = x$.

The principal and agent earn $y - s(y)$ and $s(y) - c(a)$, respectively, where $c(\cdot)$ is strictly convex. The distribution $G(\cdot)$ can be interpreted as the agent mixing among distributions $F(\cdot|x)$ over output, where a increases the expected value of that mixture. As an example, suppose the agent is an entrepreneur who chooses a product to bring to market. Effort a improves the quality of whichever product she chooses, while $G(\cdot)$ captures the entrepreneur's choice between a product that would be modestly profitable regardless of economic conditions, and one that would have more variable profitability. The distribution $F(\cdot|x)$ represents residual demand uncertainty that depends on the economic conditions. Note that 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).

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

$$V_s(x) \equiv \int s(y)f(y|x)dy. \tag{6}$$

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

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

We prove that a linear contract is optimal in this problem.

Proposition 4. *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 this 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)-(Conc), except that $V_s^c(\cdot)$ is a function of effort rather than of intermediate output. Nevertheless, much as in the proof of Proposition 1, we show that 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 with any given slope and intercept if $s(\cdot)$ is linear with the same slope and intercept 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 must be 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 $s(y) - c(a) - \mathbb{E}_{G_x}[d(y)] + d(x)$, respectively.[15]

For any contract $s(\cdot)$, define the agent's "modified payment" and "modified cost" as

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

respectively. Then the agent's payoff equals $\tilde{v}(y) - \tilde{c}(a)$. 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 optimally 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 1 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 $d(\cdot)$.

[15]We are very grateful to Doron Ravid for suggesting this formulation of the cost function.

Proposition 5. *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.*

As in Proposition 1, a contract that makes $\tilde{v}(\cdot)$ linear implements higher effort than any strictly concave $\tilde{v}(\cdot)$. Therefore, $\tilde{v}(\cdot)$ is optimally linear, which implies that $s^*(\cdot)$ is optimal. Intuitively, the profit-maximizing contract is the most convex contract that does not induce the agent to gamble.

Importantly, $s^*(\cdot)$ does not equal the optimal linear contract from Proposition 1. In fact, one can show that the principal pays strictly less in expectation than in that contract. This is intuitive: if the agent finds risk-taking costly, then somewhat convex contracts do not induce gaming.

6.3 Manipulating the Timing of Output[16]

This section proposes 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 setting is isomorphic to the model in Section 4.

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 $\int_0^1 y_x(t)dt = x$.

[16] We are very grateful to Lars Stole for suggesting this interpretation of the model.

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, this model constrains the principal to offer a stationary contract $s(\cdot)$. Without this assumption, the principal could eliminate gaming incentives entirely by paying only for output realized at time $t = 1$. 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 $y_x(t) \leq y$.[17] Then $G_x(\cdot)$ is a distribution: it is increasing, with $G_x(\underline{y}) = 0$ and $G_x(\bar{y}) = 1$. The amount of time the agent spends producing exactly y equals $dG_x(y)$, so the agent's and principal's payoffs are $\int s(y)dG_x(y) - c(a) = \mathbb{E}_{G_x}[s(y)] - c(a)$ and $\mathbb{E}_{G_x}[y - s(y)]$, respectively, where $G_x(\cdot)$ must satisfy $\mathbb{E}_{G_x}[y] = x$. Therefore, for each x, both players' expected payoffs are as in (Obj$_F$)-(LL$_F$), and consequently, the results from Sections 3 and 4 apply.

Proposition 6. *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 his output over time if $s(\cdot)$ is concave, and 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.

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

The assumption that the agent can continuously adjust his realized output implies that he has substantial freedom to game the contract. For instance, suppose that the agent could adjust her output only once, at $t = \frac{1}{2}$. Then for each x, the agent could choose at most two output levels $y_L, y_H \in \mathcal{Y}$, each of which would be produced one-half of the time. Consequently, $G_x(y) = 0$ for $y < y_L$, $G_x(y) = \frac{1}{2}$ for $y \in [y_L, y_H)$, and $G_x(y) = 1$ for $y \geq y_H$. In contrast, if the agent can adjust output continuously, then he can choose any $G_x(\cdot)$ that satisfies $\mathbb{E}_{G_x}[y] = x$.

7 Concluding Remarks

While we have focused on the relationship between a single principal and agent, incentive contracts are rarely offered in a competitive vacuum. For instance, Chevalier and Ellison (1997) describe how tournament-like incentives drive financial advisors to make risky investments. More generally, an individual's (or a firm's) competitive context shapes the incentives they face, which in turn determine the kinds of risks they optimally pursue. A more complete analysis of how competition interacts with risk-taking could shed light on behavior in both financial and product markets. See Fang and Noe (2015) for a step in this direction.

We explore risk-taking in the context of a formal incentive scheme. As emphasized in Makarov and Plantin (2015), an agent might also engage in risk-taking to try to manipulate implicit incentives, for instance by gaming his superiors' beliefs about his ability. Further research is required to understand how such implicit incentives might encourage or deter risk-taking.

References

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

Attar, A., Mariotti, T., and Salanié, F., 2011. Non-exclusive competition in the market for lemons. *Econometrica*, 79(6), pp. 1869-1918.

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.

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

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.

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.

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.

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

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.

Repenning, N.P. and Henderson, R.M., 2010. Making the Numbers? "Short Termism" & the Puzzle of Only Occasional Disaster. National Bureau of Economic Research.

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

A.1 Proof of Lemma 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.

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

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

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

A.2 Proof of Proposition 1

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

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,[18] (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 $\hat{s}_{\hat{a}}$,

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

Here, the second equality uses that $\mathbb{E}_{F(\cdot|a)}\left[\hat{s}_{\hat{a}}(y)\right]$ 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 $t > s^L_{\hat{a}}(y)$,

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

[18] 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.

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 $\mathbb{E}_{F(\cdot|a^*)}\left[s^L_{\hat{a}}(y)\right] \leq \mathbb{E}_{F(\cdot|a^*)}\left[s^*(y)\right]$ 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.3 Proof of Proposition 2

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)-(Conc) 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 $s^L_a(\underline{y}) = -M$, so Proposition 1 implies that $s^L_a(\cdot)$ implements a at maximum profit in this modified problem. So the principal's expected payoff equals

$\mathbb{E}_{F(\cdot|a)}\left[\pi^c(x - s^L_a(x))\right]$ in this modified problem.

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

A.4 Proof of Corollary 1

Fix $\hat{a} > 0$. Define

$$y_1 \equiv \min_{a \in [\hat{a}, a^{FB}]} \left\{ a - \frac{c(a) + u_0 + M}{c'(a)} \right\},$$

and

$$y_2 \equiv \min_{a \in [\hat{a}, a^{FB}]} \left\{ \frac{u^{-1}(u_0) - (1 - c'(a)a) - M}{c'(\hat{a})} \right\},$$

and note that since $c'(a) \geq c'(\hat{a}) > 0$ for all $a \geq \hat{a}$, $y_{min} \equiv \min\{0, y_1, y_2\} > -\infty$.

Let $\underline{y} < y_{min}$, and suppose towards a contradiction that there exists a distribution $F(\cdot|a)$ on $[\underline{y}, \bar{y}]$ such that effort $a^* \geq \hat{a}$ is optimal under $F(\cdot|a)$. Note first that Proposition 1 implies that the principal's expected payoff equals

$$\mathbb{E}_{F(\cdot|a^*)}\left[\pi(y - s^L_{a^*}(y))\right] = \mathbb{E}_{F(\cdot|a^*)}\left[\pi\left(y - c'(a^*)(y - \underline{y}) + \min\left\{M, c'(a^*)(a^* - \underline{y}) - c(a^*) - u_0\right\}\right)\right].$$

Since $\underline{y} < y_1$, $c'(a^*)(a^* - \underline{y}) - c(a^*) - u_0 > M$. Furthermore, the principal's payoff is bounded above by

$$\pi\left((1 - c'(a^*))a^* + c'(a^*)\underline{y} + M\right)$$

by Jensen's inequality. Since $\underline{y} < \min\{0, y_2\}$, $(1-c'(a))a + c'(a)\underline{y} + M < u^{-1}(u_0)$

for any $a \in [\hat{a}, a^{FB}]$. But then $a^* \geq \hat{a}$ cannot be optimal because it is strictly dominated by $a^* = 0$ and $s(\cdot) \equiv u^{-1}(u_0)$, a contradiction. Hence, for $\underline{y} < y_{min}$, any distribution $F(\cdot|a)$, and any optimal a^*, it must be that $a^* < \hat{a}$. Since $\hat{a} > 0$ is arbitrary, $\lim_{\underline{y} \to -\infty} a^* = 0$.

Suppose $\pi(y) \equiv y$. To prove that a^* is increasing in $\underline{y}$, it suffices to show that the principal's payoff from implementing a in an optimal contract, $\Pi(a, \underline{y}) = a - c'(a)(a - \underline{y}) + w$, is supermodular in a and $\underline{y}$.

Recall that $w = \min\left\{M,\ c'(a)(a - \underline{y}) - c(a) - u_0\right\}$ is a function of $(a, \underline{y})$. Therefore,

$$\frac{\partial \Pi}{\partial a} = 1 - c''(a)(a - \underline{y}) - c'(a) + \frac{\partial w}{\partial a}$$

and so

$$\frac{\partial^2 \Pi}{\partial \underline{y} \partial a} = c''(a) + \frac{\partial^2 w}{\partial \underline{y} \partial a}.$$

But $\frac{\partial^2 w}{\partial \underline{y} \partial a} = 0$ if $M < c'(a)(a - \underline{y}) - c(a) - u_0$ and $\frac{\partial^2 w}{\partial \underline{y} \partial a} = -c''(a)$ otherwise. In either case, $\frac{\partial^2 \Pi}{\partial \underline{y} \partial a} \geq 0$ and so optimal effort a^* is increasing in $\underline{y}$, as desired. ■

B For Online Publication: Proofs for Sections 5

First, we prove some preliminary properties of optimal incentives schemes. If $\underline{u} > -\infty$, Lemma 2 shows that any profit-maximizing incentive scheme $v(\cdot)$ must be unique, and Online Appendix D shows the same for $\underline{u} = -\infty$. We prove that $v(\cdot)$ must be monotonically increasing and satisfy (IC-FOC) with equality.

Suppose $v(\cdot)$ is concave and not everywhere increasing. Then, we can find $\tilde{y} \in \mathcal{Y}$ such that if we replace $v(y)$ by a constant $v(\tilde{y})$ to the right of $\tilde{y}$, the resultant contract is concave, gives the same utility to the agent, is cheaper, and, using MLRP and Beesack's inequality makes (IC-FOC) slack. So any optimal $v(\cdot)$ must be increasing.

Suppose $v(\cdot)$ does not satisfy (IC-FOC) with equality. Then, a convex combination of v and the contract which gives utility constant and equal to $\max\{\underline{u}, u_0 + c(a)\} \geq 0$ implements a, is strictly cheaper than v, and satisfies (IC-FOC) with equality. So any optimal $v(\cdot)$ must satisfy (IC-FOC) with equality.

Consider an interval $[y_L, y_H]$. The initial impact of *raising* the agent's utility on this interval is given by

$$r_{y_L,y_H}(y) = \begin{cases} 1 & y \in [y_L, y_H] \\ 0 & else \end{cases}.$$

Similarly, *tilting* this interval has an initial impact on the agent's utility given by

$$t_{y_L,y_H}(y) = \begin{cases} 0 & y \leq y_L \\ y - y_L & y \in (y_L, y_H) \\ y_H - y_L & y \geq y_H \end{cases}.$$

We will carefully define the perturbations *raise* and *tilt* and show that they respect concavity in Section B.1.2.

Our first result proves two useful properties of any contract that is GHM.

Lemma 3. *Let v be GHM, and let $[y_L, y_H]$ be a linear segment of v. Then, for* *each* $\hat{y} \in (y_L, y_H)$*, there is* $\tilde{y} \in (\hat{y}, y_H)$ *such that*

$$n(\tilde{y}) \leq 0.$$

If $v(y_L) > \underline{u}$*, then such a* $\tilde{y}$ *exists in* $(y_L, \hat{y})$ *as well. But, somewhere on* (y_L, y_H)*,* $n(y) \geq 0$*.*

Proof. Note that for $y > y_H$, $t_{\hat{y},y_H}(y) = y_H - \hat{y} = (y_H - \hat{y})\, r_{y_H,\bar{y}}(y)$. Since v satisfies IC, since $a > 0$, and since v is concave and weakly increasing, v must be strictly increasing near $\underline{y}$. Hence, since $y_H > \underline{y}$, $v(y_H) > \underline{u}$. We thus have $\int n(y)\, r_{y_H,\bar{y}}(y)\, f(y|a)\, dy = 0$ by Definition 1.1. Hence, by Definition 1.3, we have

$$\begin{aligned} 0 &\geq \int n(y)\, t_{\hat{y},y_H}(y)\, f(y|a)\, dy \\ &= \int n(y)\, t_{\hat{y},y_H}(y)\, f(y|a)\, dy - (y_H - \hat{y}) \int n(y)\, r_{y_H,\bar{y}}(y)\, f(y|a)\, dy \\ &= \int_{\hat{y}}^{y_H} n(y)\, t_{\hat{y},y_H}(y)\, f(y|a)\, dy, \end{aligned}$$

and so at some point $\tilde{y} \in (\hat{y}, y_H)$, the integrand is weakly negative. Since $t_{\hat{y},y_H}(\tilde{y}) > 0$, it follows that $n(\tilde{y}) \leq 0$.

Similarly, note that if $v(y_L) > \underline{u}$, then $\int n(y)\, r_{y_L,\bar{y}}(y)\, f(y|a)\, dy = 0$ by Definition 1.1, and so by Definition 1.2,

$$\begin{aligned} 0 &\leq \int n(y)\, t_{y_L,\hat{y}}(y)\, f(y|a)\, dy \\ &= \int n(y)\, t_{y_L,\hat{y}}(y)\, f(y|a)\, dy - (\hat{y} - y_L) \int n(y)\, r_{y_L,\bar{y}}(y)\, f(y|a)\, dy \\ &= \int_{y_L}^{\hat{y}} n(y) \left[t_{y_L,\hat{y}}(y) - (\hat{y} - y_L) \right] f(y|a)\, dy, \end{aligned}$$

where, since the bracketed term is strictly negative on $(y_L, \hat{y})$, it follows that $n(y)$ is somewhere weakly negative on $(y_L, \hat{y})$.

Finally, since $\int n(y)\, r_{y_L,y_H}(y)\, f(y|a)\, dy \geq 0$, and since we have established that $n(y)$ is weakly negative somewhere on (y_L, y_H), we must also have $n(y)$

weakly positive somewhere on the same interval.

□

B.1 Proof of Proposition 3

The discussion prior to the statement of Proposition 3 proves necessity, given well-defined perturbations that satisfy concavity, and well-defined shadow values. This section begins by formally defining the relevant perturbations, showing that they preserve concavity, and then showing how they can be used to establish shadow values for (IR) and (IC-FOC). We then turn to sufficiency.

B.1.1 Preliminaries

Definition 1 and Proposition 3 are phrased in terms of free points. But, not every free point is a convenient place to define a perturbation. Instead, for any given v, let C_v be the set of points y at which there exists a supporting plane L such that $L(y') > v(y')$ for all $y' \neq y$.

Clearly any kink point (see the discussion immediately before Corollary 2) is an element of C_v. The next claim shows that for every other free point, there is an arbitrarily close-by element of C_v.

Claim 1. *Let $\hat{y}$ be any point of normal concavity. Then, for each δ, there is a point in $\{(\hat{y}-\delta, \hat{y}+\delta) \setminus \hat{y}\} \cap C_v$. From this, it follows that for each $\varepsilon > 0$, there exists $y_L < y_H$ such that $y_L, y_H \in C_v$, and such that $y_L, y_H \in [\hat{y}-\varepsilon, \hat{y}+\varepsilon]$.*

Proof of Claim. We will show first that for each δ, there is a point in $\{(\hat{y}-\delta, \hat{y}+\delta) \setminus \hat{y}\} \cap C_v$. To see that this suffices to show the second part, apply the result first to find a point y_1 in $\{(\hat{y}-\varepsilon, \hat{y}+\varepsilon) \setminus \hat{y}\} \cap C_v$. Apply the result again to find y_2 in $\{(\hat{y}-\delta, \hat{y}+\delta) \setminus \hat{y}\} \cap C_v$ where $\delta = (1/2)|y_1 - \hat{y}|$, and finally take y_L and y_H as the smaller and larger of y_1 and y_2.

So, fix $\delta > 0$. Since $\hat{y}$ is not on the interior of a linear segment and not a kink point, there is at least one side of $\hat{y}$, without loss of generality the right side, such that $v(\cdot)$ is not linear on $(\hat{y}, \hat{y}+\delta)$. Let $S(\cdot)$ be the correspondence which for each y assigns the set of slopes of supporting planes at y, and let

$s(\cdot)$ be any selection from $S(\cdot)$. Note that since v is concave, for any $y'' > y'$, $\max\{S(y'')\} \leq \min\{S(y')\}$, and hence s is decreasing. Assume first that there is a point $\tilde{y} \in (\hat{y}, \hat{y} + \delta)$ where $s(\cdot)$ jumps downward, say from s'' to $s' < s''$. Then, the supporting plane at $\tilde{y}$ with slope $(s' + s'')/2$ qualifies. Assume instead that $s(\cdot)$ is continuous on $(\hat{y}, \hat{y} + \delta)$. It cannot be everywhere constant, since $v(\cdot)$ is not linear on $(\hat{y}, \hat{y} + \delta)$. Hence, since $s(\cdot)$ is continuous, there is a point $\tilde{y}$ at which it is strictly decreasing, so that in specific, $s(\tilde{y}) < s(y)$ for all $y < \tilde{y}$, and $s(\tilde{y}) > s(y)$ for all $y > \tilde{y}$. The supporting plane at $\tilde{y}$ with slope $s(\tilde{y})$ then qualifies.

□

To see that why Claim 1 is helpful, assume that some part of Definition 1 is violated. For example, assume some optimal contract has a pair of free points y_L and y_H such that $\int n(y) r_{y_L,y_H} f(y) dy < 0$. If either y_L or y_H is a kink point, then it is also an element of C_v. If not, then we can apply Claim 1 to replace each relevant point by a sufficiently close-by element of C_v that the strict inequality is maintained. Hence, it is enough to prove Proposition 3 when each restriction to a free point is tightened to a restriction to C_v.

B.1.2 Formal Definition and Properties of the Perturbations

This section defines *raise* and *tilt*, being careful in particular to maintain concavity at the endpoints of the perturbed interval. We will need to consider as many as three perturbations at once, where, given the previous discussion, we will require the relevant points to be in C_v. First, we will have some small amount ε_p of a perturbation p where p could be r_{y_L,y_H} or t_{y_L,y_H} in each case with ε_p positive or negative. Second, for some $\hat{y} \in C_v$, we will need to consider some amount ε_t of $t_{\hat{y},\bar{y}}$ and ε_r of $r_{\hat{y},\bar{y}}$. Intuitively, we will use $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$ to establish shadow values for (IC-FOC) and (IR), and then, for any particular perturbation p, consider the three deviations together where one uses $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$ to undo the effect of p on (IC-FOC) and (IR).

Fix y_L, y_H, and $\hat{y}$. *A priori*, $\hat{y}$ may have arbitrary position relative to y_L and y_H, and moreover, in the case where p is t_{y_L,y_H}, one of y_L or y_H may not be in C_v, depending on whether ε_p is negative or positive. Define $y_0 < y_1 < \cdots < y_K$, $K \leq 4$, as elements of the set $\left\{\underline{y}, y_L, y_H, \hat{y}, \bar{y}\right\} \cap C_v$. For any given $\boldsymbol{\varepsilon} = (\varepsilon_p, \varepsilon_t, \varepsilon_r)$, let $d(\cdot; \boldsymbol{\varepsilon}) : [\underline{y}, \bar{y}] \to \mathbb{R}$ be given by

$$d(\cdot; \boldsymbol{\varepsilon}) = \varepsilon_p p(\cdot) + \varepsilon_t t_{\hat{y},\bar{y}}(\cdot) + \varepsilon_r r_{\hat{y},\bar{y}}(\cdot).$$

If y_L and y_H are both elements of $\{y_0, \cdots, y_K\}$, as must be true if p is r_{y_L,y_H}, then it follows that d is linear on each interval of the form (y_{k-1}, y_k). Assume that $y_H \notin \{y_0, \cdots, y_K\}$. Then, it must be that p is t_{y_L,y_H} with $\varepsilon_p \geq 0$. In this case, if $y_H \notin (y_{k-1}, y_k)$, then $d(\cdot; \boldsymbol{\varepsilon})$ is linear on (y_{k-1}, y_k), while if $y_H \in (y_{k-1}, y_k)$, then, since $\varepsilon_p \geq 0$, $d(\cdot; \boldsymbol{\varepsilon})$ is concave with two linear segments on (y_{k-1}, y_k). Finally, assume $y_L \notin \{y_0, \cdots, y_K\}$. Then, p is t_{y_L,y_H} with $\varepsilon_p \leq 0$, and once again, if $y_L \notin (y_{k-1}, y_k)$, then $d(\cdot; \boldsymbol{\varepsilon})$ is linear on (y_{k-1}, y_k), while if $y_L \in (y_{k-1}, y_k)$, then since $\varepsilon_p \leq 0$, $d(\cdot; \boldsymbol{\varepsilon})$ is once again concave with two linear segments on (y_{k-1}, y_k).

For each k, let $L_k^-(\cdot; \boldsymbol{\varepsilon})$ be the line that coincides with the linear segment of $d(\cdot; \boldsymbol{\varepsilon})$ immediately to the right of y_{k-1} and let $L_k^+(\cdot; \boldsymbol{\varepsilon})$ be the line that coincides with the linear segment immediately to the left of y_k (these are the same line if d is linear on (y_{k-1}, y_k)), and let

$$d_k(y; \boldsymbol{\varepsilon}) = \begin{cases} L_k^-(y; \boldsymbol{\varepsilon}) & y \leq y_{k-1} \\ d(y; \boldsymbol{\varepsilon}) & y \in (y_{k-1}, y_k) \\ L_k^+(y; \boldsymbol{\varepsilon}) & y \geq y_k \end{cases}.$$

Note that d_k is concave, and that as $|\boldsymbol{\varepsilon}| \equiv |\varepsilon_p| + |\varepsilon_t| + |\varepsilon_r| \to 0$, d_k converges uniformly to the function that is constant at 0.

For each k, let L_k be a supporting line to v at y_k, where since $y_k \in C_v$, we can choose L_k such that $L_k(y) > v(y)$ for all $y \neq y_k$, and let

$$v_k(y) = \begin{cases} L_{k-1}(y) & y \leq y_{k-1} \\ v(y) & y \in (y_{k-1}, y_k) \\ L_k(y) & y \geq y_k \end{cases},$$

so that $v_k(\cdot)$ is concave. Define $\hat{v}(\cdot;\boldsymbol{\varepsilon})$ by

$$\hat{v}(y;\boldsymbol{\varepsilon}) = \min_{k\in\{1,\cdots,K\}} (v_k(y) + d_k(y;\boldsymbol{\varepsilon})).$$

As the minimum over concave functions, $\hat{v}(\cdot;\boldsymbol{\varepsilon})$ is concave.

Fix k and consider any $y \in (y_{k-1}, y_k)$. Since $d_k(y, \mathbf{0}) = 0$, and by the fact that for each k', $L_{k'}(y) > v(y)$ for all $y \neq y_{k'}$, k is the unique minimizer of $v_k(y) + d_k(y;\mathbf{0})$. From this, it follows first that $\hat{v}(y;\mathbf{0}) = v_k(y) = v(y)$, and second, that for all $\boldsymbol{\varepsilon}$ in some neighborhood of $\mathbf{0}$ (where ε_p is restricted in sign if $p = t_{y_L,y_H}$ and if one of y_L or y_H is not in C_v),

$$\begin{aligned}
\hat{v}_{\varepsilon_p}(y;\boldsymbol{\varepsilon}) &= d_{\varepsilon_p}(y;\boldsymbol{\varepsilon}) = p(y), \\
\hat{v}_{\varepsilon_t}(y;\boldsymbol{\varepsilon}) &= d_{\varepsilon_t}(y;\boldsymbol{\varepsilon}) = t_{\hat{y},\bar{y}}(y), \text{ and} \\
\hat{v}_{\varepsilon_r}(y;\boldsymbol{\varepsilon}) &= d_{\varepsilon_r}(y;\boldsymbol{\varepsilon}) = r_{\hat{y},\bar{y}}(y).
\end{aligned}$$

But then, except on the zero-measure set of points $\{y_0, \cdots, y_K\}$,

$$\begin{aligned}
\hat{v}_{\varepsilon_p}(\cdot;\mathbf{0}) &= p(\cdot), \\
\hat{v}_{\varepsilon_t}(\cdot;\mathbf{0}) &= t_{\hat{y},\bar{y}}(\cdot), \text{ and} \\
\hat{v}_{\varepsilon_r}(\cdot;\mathbf{0}) &= r_{\hat{y},\bar{y}}(\cdot).
\end{aligned} \tag{9}$$

B.1.3 Shadow Values

We need to establish that starting from $\varepsilon = \mathbf{0}$ the effects of perturbation p can be undone via $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$. To do so, let

$$Q(\boldsymbol{\varepsilon}) = \begin{bmatrix} \int \hat{v}_{\varepsilon_t}(y,\boldsymbol{\varepsilon}) f_a(y|a)\,dy & \int \hat{v}_{\varepsilon_r}(y,\boldsymbol{\varepsilon}) f_a(y|a)\,dy \\ \int \hat{v}_{\varepsilon_t}(y,\boldsymbol{\varepsilon}) f(y|a)\,dy & \int \hat{v}_{\varepsilon_r}(y,\boldsymbol{\varepsilon}) f(y|a)\,dy \end{bmatrix}.$$

The top row of Q tracks the rate at which ε_t and ε_r respectively affect (IC-FOC), while the bottom row tracks the rate at which ε_t and ε_r respectively affect (IR). Then, from (9),

$$
\begin{aligned}
Q(\mathbf{0}) &= \begin{bmatrix} \int t_{\hat{y},\bar{y}} f_a(y|a)\,dy & \int r_{\hat{y},\bar{y}} f_a(y|a)\,dy \\ \int t_{\hat{y},\bar{y}} f(y|a)\,dy & \int r_{\hat{y},\bar{y}} f(y|a)\,dy \end{bmatrix} \\
&= \begin{bmatrix} \int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f_a(y|a)\,dy & \int_{\hat{y}}^{\bar{y}} f_a(y|a)\,dy \\ \int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f(y|a)\,dy & \int_{\hat{y}}^{\bar{y}} f(y|a)\,dy \end{bmatrix},
\end{aligned}
$$

and so

$$
\begin{aligned}
|Q(\mathbf{0})| &= \int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f_a(y|a)\,dy \int_{\hat{y}}^{\bar{y}} f(y|a)\,dy - \int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f(y|a)\,dy \int_{\hat{y}}^{\bar{y}} f_a(y|a)\,dy \\
&\underset{s}{=} \frac{\int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f_a(y|a)\,dy}{\int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f(y|a)\,dy} - \frac{\int_{\hat{y}}^{\bar{y}} f_a(y|a)\,dy}{\int_{\hat{y}}^{\bar{y}} f(y|a)\,dy} \\
&= \int_{\hat{y}}^{\bar{y}} l(y|a) \frac{(y-\hat{y}) f(y|a)}{\int_{\hat{y}}^{\bar{y}} (y-\hat{y}) f(y|a)\,dy} dy - \int_{\hat{y}}^{\bar{y}} l(y|a) \frac{f(y|a)}{\int_{\hat{y}}^{\bar{y}} f(y|a)\,dy} dy,
\end{aligned}
$$

where the symbol $\underset{s}{=}$ means "has (strictly) the same sign as."

Thus, $|Q(\mathbf{0})|$ has the same sign as the difference between two expectations of $l(\cdot|a)$. Using that $(y-\hat{y})$ is strictly increasing, the density in the first integral strictly likelihood-ratio dominates the density in the second integral. Since $l(\cdot|a)$ is strictly increasing, it follows that $|Q(\mathbf{0})|$ is strictly positive (and so remains so for all $\boldsymbol{\varepsilon}$ in some ball around $\mathbf{0}$.) But then by the implicit function theorem, for each $p \in \{t_{y_L,y_H}, r_{y_L,y_H}\}$, we can on the appropriate neighborhood implicitly define $\varepsilon_r(\cdot)$ and $\varepsilon_t(\cdot)$ by

$$
\begin{aligned}
\int \hat{v}(y;\varepsilon_p,\varepsilon_t(\varepsilon_p),\varepsilon_r(\varepsilon_p)) f(y|a)\,dy &= c(a) + u_0, \text{ and} \\
\int \hat{v}(y;\varepsilon_p,\varepsilon_t(\varepsilon_p),\varepsilon_r(\varepsilon_p)) f_a(y|a)\,dy &= c'(a),
\end{aligned}
$$

so that starting from $\boldsymbol{\varepsilon} = \mathbf{0}$, if we make the small perturbation ε_p to v, we can restore (IC-FOC) and (IR) by a suitable combination of small applications ε_t and ε_r of $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$.

Let λ be the rate of change of costs as one relaxes (IR) using $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$.

That is, if we let

$$\begin{pmatrix} q_t^{IR} \\ q_r^{IR} \end{pmatrix} = [Q(\mathbf{0})]^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

then

$$\lambda = \int \rho^{-1}(v(y)) \left(q_t^{IR} t_{\hat{y},\bar{y}}(y) + q_r^{IR} r_{\hat{y},\bar{y}}(y) \right) f(y|a)\, dy.$$

Similarly, if

$$\begin{pmatrix} q_t^{IC} \\ q_r^{IC} \end{pmatrix} = [Q(\mathbf{0})]^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

then the rate of change of costs as one relaxes (IC-FOC) using $t_{\hat{y},\bar{y}}$ and $r_{\hat{y},\bar{y}}$ is

$$\mu = \int \rho^{-1}(v(y)) \left(q_t^{IC} t_{\hat{y},\bar{y}}(y) + q_r^{IC} r_{\hat{y},\bar{y}}(y) \right) f(y|a)\, dy.$$

Given the shadow values λ and μ, the argument in Section 5 (prior to Definition 1) completes the proof of necessity in Proposition 3. ■

B.1.4 Proof of Sufficiency

We begin by proving the following useful result.

Lemma 4. *Let $v(\cdot)$ be GHM 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 (as defined immediately before Corollary 2).*

Proof. If y is a kink point, then Lemma 3 applied to the left of y implies that $n(y) \leq 0$. If y is a point of normal concavity, then by Lemma 1 there exist sequences of points $\{y_k^L\}, \{y_k^H\} \in C_v$ such that $y_k^L < y < y_k^H$ for all $k \in \mathbb{N}$ and $\lim_k y_k^L = \lim_k y_k^H = y$. These points are free, so 3 holds with equality on each interval $[y_k^L, y_k^H]$. Hence, in the limit, $n(y) = 0$.

□

Now, let v, with associated λ and μ, be GHM. Let us show that v is optimal. We will argue by contradiction. Assume v is not optimal, and let v^* be a lower cost contract satisfying (IC-FOC) and (IR)-(Conc). As in the argument at the beginning of Appendix B, v^* can be taken to be increasing, satisfy (IC-FOC) exactly, and as in the proof of Lemma 6 in Appendix D, $v^*(\bar{y})$ and $v^*(\underline{y})$ can be taken to be finite.

Enumerate the closed linear segments $S_1, S_2, \cdots$, of v, and let $S = \cup S_i$. Let $\delta(y) = v^*(y) - v(y)$, and let $\hat{v}(y;\varepsilon) = v(y) + \varepsilon\delta(y)$, so that $\hat{v}(\cdot, 0) = v(\cdot)$ and $\hat{v}(\cdot, 1) = v^*(\cdot)$. Then, for each ε, $\hat{v}(\cdot;\varepsilon)$ is a convex combination of the concave contracts v and v^*. Hence, $\hat{v}(\cdot;\varepsilon)$ satisfies (IC-FOC) and (IR)-(Conc). Since $u^{-1}(\cdot)$ is convex, and since for each y, $\hat{v}(y;\varepsilon)$ is linear in ε, it follows that $\int u^{-1}(\hat{v}(y;\varepsilon)) f(y|a)\,dy$ is convex in ε. Thus, since

$$\begin{aligned}
\int u^{-1}(\hat{v}(y;0)) f(y|a)\,dx &= \int u^{-1}(v(y)) f(y|a)\,dy \\
&> \int u^{-1}(v^*(y)) f(y|a)\,dy \\
&= \int u^{-1}(\hat{v}(y;1)) f(y|a)\,dy,
\end{aligned}$$

it follows that

$$\begin{aligned}
0 &> \frac{d}{d\varepsilon}\int u^{-1}(\hat{v}(y;0)) f(y|a)\,dy \\
&= \int \frac{1}{u'(u^{-1}(\hat{v}(y;0)))}\delta(y) f(y|a)\,dy \\
&= \int \rho^{-1}(v(y))\,\delta(y) f(y|a)\,dy \\
&= \int_S \rho^{-1}(v(y))\,\delta(y) f(y|a)\,dy + \int_{\mathcal{Y}\backslash S} \rho^{-1}(v(y))\,\delta(y) f(y|a)\,dy,
\end{aligned}$$

and so, since every point in $\mathcal{Y}\backslash S$ is a point of normal concavity (noting that we took the sets S_i to be closed, and so any kink point is in S), we have

$$\begin{aligned}
\int_S \rho^{-1}(v(y))\,\delta(y) f(y|a)\,dy &< -\int_{\mathcal{Y}\backslash S} \rho^{-1}(v(y))\,\delta(y) f(y|a)\,dy \\
&= -\int_{\mathcal{Y}\backslash S} (\lambda + \mu l(y|a))\,\delta(y) f(y|a)\,dy \\
&= -\lambda\int_{\mathcal{Y}\backslash S} \delta(y) f(y|a)\,dy - \mu\int_{\mathcal{Y}\backslash S} \delta(y) f_a(y|a)\,dy.
\end{aligned}$$

where the first equality follows by Lemma 4.

Both v and v^* satisfy (IC-FOC) with equality, and hence $\int \delta(y) f_a(y|a)\, dy = 0$, from which

$$-\mu \int_{\mathcal{Y}\setminus S} \delta(y) f_a(y|a)\, dy = \mu \int_S \delta(y) f_a(y|a)\, dy.$$

Similarly, either (IR) is binding at v, in which case $\int \delta(y) f(y|a)\, dy \geq 0$, or (IR) does not bind at v, in which case $\lambda = 0$, and hence in either case

$$-\lambda \int_{\mathcal{Y}\setminus S} \delta(y) f(y|a)\, dy \leq \lambda \int_S \delta(y) f(y|a)\, dy.$$

Making these two substitutions thus yields

$$\int_S \rho^{-1}(v(y))\, \delta(y) f(y|a)\, dy < \lambda \int_S \delta(y) f(y|a)\, dy + \mu \int_S \delta(y) f_a(y|a)\, dy.$$

Hence, since $S = \cup S_i$, where the S_i's are disjoint except possibly at their zero-measure boundaries, there must be some i such that

$$\int_{S_i} \rho^{-1}(v(y))\, \delta(y) f(y|a)\, dy < \lambda \int_{S_i} \delta(y) f(y|a)\, dy + \mu \int_{S_i} \delta(y) f_a(y|a)\, dy,$$

or equivalently,

$$\int_{S_i} n(y)\, \delta(y) f(y|a)\, dy < 0.$$

Fix such an i, and consider δ_1, the restriction of δ to $S_i = [y_L, y_H]$. Since v is linear on S_i, and v^* is concave, δ_1 is concave. For any given K, let $\Delta = (y_H - y_L)/2^K$, and consider the function δ_K on $[y_L, y_H]$ that agrees with δ_1 on the set of points $\{y_L, y_L + \Delta, \cdots, y_H\}$, and is linear between these points. Note that δ_K is concave and continuous on $[y_L, y_H]$, and that for each y, $\delta_K(y)$ is monotonically increasing in K with limit $\delta(y)$. Hence, we can choose $\hat{K}$ large enough that

$$\int_{S_i} n(y)\, \delta_{\hat{K}}(y) f(y|a)\, dy < 0.$$

Finally, define $\tilde{\delta}$ on $[\underline{y}, \bar{y}]$ by

$$\tilde{\delta}(y) = \begin{cases} 0 & y \leq y_L \\ \delta_{\hat{K}}(y) & y \in [y_L, y_H] \\ \delta_{\hat{K}}(y_H) & y > y_H \end{cases}.$$

Note that y_H and $\bar{y}$ are free. Note also that as in the proof of Lemma 3, $v(y_H) > \underline{u}$. It follows from Definition 1.1 that since $\tilde{\delta}$ is constant on $[y_H, \bar{y}]$,

$$\int_{y_H}^{\bar{y}} n(y)\,\tilde{\delta}(y)\,f(y|a)\,dy = 0,$$

and hence,

$$\int n(y)\,\tilde{\delta}(y)\,f(y|a)\,dy < 0.$$

Let us next argue that $\tilde{\delta}$ can be expressed as a sum of raises and tilts. For $k \in \left\{0, \ldots, 2^{\hat{K}}\right\}$, let $y_k = y_L + k\Delta$, and let s_k be the slope of $\tilde{\delta}$ on (y_{k-1}, y_k). Then, we claim that for all y in $[y_L, y_H]$,

$$\tilde{\delta}(y) = \delta(y_0)\,r_{y_0,\bar{y}}(y) + \sum_{k=1}^{2^{\hat{K}}-1} (s_k - s_{k+1})\,t_{y_0,y_k}(y) + s_{2^{\hat{K}}} t_{y_0,y_{2^{\hat{K}}}}(y). \qquad (10)$$

To see (10) note first that for $y < y_0 = y_L$, both sides of the equation are 0. At y_0, each side is $\delta(y_0)$, since $r_{y_0,\bar{y}}(y_0) = 1$, and since $t_{y_0,y_k}(y_0) = 0$ for all k. Thus, since both sides are continuous and piecewise linear on $[y_0, \bar{y}]$, it is enough that the two sides have that same derivative where defined. So, fix $\hat{k} \in \left\{1, \ldots, 2^{\hat{K}}\right\}$, and let $y \in \left(y_{\hat{k}-1}, y_{\hat{k}}\right)$. Note that for $k < \hat{k}$, $t'_{y_0,y_k}(y) = 0$, and for $k \geq \hat{k}$, $t'_{y_0,y_k}(y) = 1$. Hence, the derivative of the right-hand side is

$$\sum_{k=\hat{k}}^{2^{\hat{K}}-1} (s_k - s_{k+1}) + s_{2^{\hat{K}}} = s_{\hat{k}},$$

as desired, and so, noting that $\tilde{\delta}'(y) = 0$ for $y > y_K = y_H$, we have established

(10).

Since $\int n(y)\,\tilde{\delta}(y)\,f(y|a)\,dy < 0$, we must thus have at least one of

1. $\delta(y_0)\int n(y)\,r_{y_0,\bar{y}}(y)\,f(y|a)\,dy < 0$,

2. for some $k < 2^{\hat{K}}$, $(s_k - s_{k+1})\int n(y)\,t_{y_0,y_k}(y)\,f(y|a)\,dy < 0$, or

3. $s_{2^{\hat{K}}}\int n(y)\,t_{y_0,y_{2^{\hat{K}}}}(y)\,f(y|a)\,dy < 0$.

By Definition 1.1, and since y_0 is free, $\int n(y)\,r_{y_0,\bar{y}}(y)\,f(y|a)\,dy = \int_{y_0}^{\bar{y}} n(y)\,f(y|a)\,dy \geq 0$, and so 1 cannot hold. Since $\tilde{\delta}$ is concave on $[y_L, y_H]$, it follows that $s_k - s_{k+1} \geq 0$, and so, since y_0 is free, it follows by Definition 1.2 that 2 cannot hold either. Finally, since y_0 and $y_{2^{\hat{K}}}$ are both free, the integral in 3 is in fact 0 by Definition 1.2 and Definition 1.3. We thus have the required contradiction, and v is in fact optimal. ■

B.2 Proof of Corollary 2

This result follows immediately from Proposition 3 and Lemma 4. ■

B.3 Proof of Corollary 3

Let $v(\cdot)$ be an optimal incentive scheme, and suppose (IR) does not bind. Towards a contradiction, suppose that $v(\cdot)$ is strictly concave at some $y < y_0$. Consider the alternative contract

$$\tilde{v}(y) = \begin{cases} \alpha v(y) + (1-\alpha)\left[v(\underline{y}) + (y-\underline{y})\frac{v(y_0)-v(\underline{y})}{y_0-\underline{y}}\right] & y \leq y_0 \\ v(y) & y > y_0 \end{cases}.$$

Note that $\tilde{v}(\cdot)$ is concave, $\tilde{v}(y) \leq v(y)$ for all $y \in \mathcal{Y}$, $\tilde{v}(\underline{y}) \geq \underline{u}$, and there exists an interval in $[\underline{y}, y_0]$ such that $\tilde{v}(y) < v(y)$ on that interval. Therefore, $\tilde{v}(\cdot)$ is strictly less expensive than $v(\cdot)$ to the principal. Since (IR) does not bind, there exists some $\alpha \in [0,1)$ such that $\tilde{v}(\cdot)$ satisfies (IR). Furthermore,

$$\int \tilde{v}(y) f_a(y|a)dy = \int_{\underline{y}}^{y_0} \tilde{v}(y) f_a(y|a)dy + \int_{y_0}^{\bar{y}} v(y) f_a(y|a)dy >$$
$$\int_{\underline{y}}^{y_0} v(y) f_a(y|a)dy + \int_{y_0}^{\bar{y}} v(y) f_a(y|a)dy = \int v(y) f_a(y|a)dy,$$

where the strict inequality follows because $f_a(y|a)$ is negative on $y \in [\underline{y}, y_0]$. Hence, $\tilde{v}(\cdot)$ satisfies (IC-FOC). So $\tilde{v}(\cdot)$ implements a, contradicting that $v(\cdot)$ is optimal. ■

B.4 Proof of Corollary 4

Suppose $\rho(\lambda + \mu l(\cdot|a))$ is convex. Let $v(\cdot)$ be optimal, and towards a contradiction assume that there exists $\tilde{y} \in (\underline{y}, \bar{y})$ that is free. Then $n(\tilde{y}) \leq 0$ by Corollary 2. But $v(\cdot)$ is concave and $\rho(\lambda + \mu l(\cdot|a))$ is convex, and so either $n(y) < 0$ for all $y < \tilde{y}$, or $n(y) < 0$ for all $y > \tilde{y}$. Either violates (3), since in particular $\underline{y}$, $\tilde{y}$, and $\bar{y}$ are all free. ■

B.5 Proof of Corollary 5

Given the discussion immediately preceding the proposition, it is enough to show that $\rho(\lambda + \mu l(\cdot|a))$ is never first strictly concave and then weakly convex.

For any analytic function q with domain a subset of the reals, let $q^{(k)}$ be the k^{th} derivative of q.

Lemma 5. *Assume $q > 0$ is not everywhere a constant, is analytic, and has $con(q) = \omega > -\infty$. Assume also that for some $\hat{y}$ on the interior of its domain, $q'(\hat{y}) = 0$. Let $\hat{k} = \min\left\{k | q^{(k)}(\hat{y}) \neq 0\right\}$. Then, $q^{(\hat{k})}(\hat{y}) < 0$.*

Proof. Note that $\hat{k} \geq 2$. Recall that q has concavity ω if q^{ω}/ω is concave, or, equivalently (cancelling the strictly positive term $q^{\omega-2}$), if for all y in the domain of q,

$$\xi(y) \equiv (\omega - 1)(q'(y))^2 + q(y)\,q''(y) \leq 0.$$

So, in particular, if $\hat{k} = 2$, then we must have $q''(\hat{y}) < 0$, since $\xi(\hat{y}) \leq 0$. Note that for $k \in \{0, 1, 2, \cdots\}$

$$\xi^{(k)}(\hat{y}) = d(\hat{y}) + q(\hat{y})\,q^{(k+2)}(\hat{y}),$$

where d is an expression involving derivatives of q of order less than $k+2$. So, the first non-zero term of the Taylor expansion of ξ is $\frac{\xi^{(\hat{k}-2)}(\hat{y})}{(\hat{k}-2)!}(y - \hat{y})^{\hat{k}-2}$, where $\xi^{(\hat{k}-2)}(\hat{y}) = q(\hat{y})\, q^{(\hat{k})}(\hat{y})$. Hence, since $(y-\hat{y})^{\hat{k}-2} > 0$ for $y > \hat{y}$, while $\xi(y) \leq 0$, $q^{(\hat{k})}(\hat{y})$, which is non-zero by assumption, must be strictly negative. □

Using this lemma, we can prove the following claim, from which Corollary 5 is immediate.

Claim 2. *Let g and h be strictly positive analytic functions with $con(g') + con(h') > -1$, and g' and h' everywhere strictly positive. Then, $(g(h(\cdot)))$ is never first strictly concave and then weakly convex.*

Proof. Let

$$\theta(\cdot) = (g(h(\cdot)))'' = g''(h')^2 + g'h''. \tag{11}$$

If both g and h are linear, then $\theta \equiv 0$, and we are done. Assume g and h are not both linear, and consider any point $\hat{y}$ at which $\theta = 0$. We will show that immediately to the right of $\hat{y}$, $\theta < 0$. This rules out that θ is ever first strictly negative and then weakly positive over any interval of non-zero length.

To see this, note that

$$\theta' = g'''(h')^3 + 3g''h'h'' + g'h'''. \tag{12}$$

Consider any point $\hat{y}$ at which $\theta = 0$. Consider first the case that $g''(\hat{y})\, h''(\hat{y}) \neq 0$. Then, since $g' > 0$, it follows by (11) that $g''(\hat{y})$ and $h''(\hat{y})$ have opposite sign. Hence, $g''(\hat{y})\, h''(\hat{y})\, h'(\hat{y}) < 0$, and so, evaluated at $\hat{y}$,

$$\begin{aligned} \theta' &\underset{s}{=} -\frac{g'''(h')^2}{g''h''} - 3 - \frac{g'h'''}{g''h''h'} \\ &= \frac{g'g'''}{(g'')^2} - 3 + \frac{h'h'''}{(h'')^2} \\ &\leq -con(g') - con(h') - 1 \\ &< 0 \end{aligned}$$

where in the second line we substitute for $(h')^2$ in the first term using (11) and that $\theta(\hat{y}) = 0$, and similarly for g' in the third term. Hence, θ is negative on an interval to the right of $\hat{y}$.

Assume instead that $g''(\hat{y})\, h''(\hat{y}) = 0$, where, since $\theta(\hat{y}) = 0$, it follows that $g''(\hat{y}) = h''(\hat{y}) = 0$. Thus, since $con(g') > -\infty$, it follows from Lemma 5 applied to $q = g'$ that the first non-zero derivative of g' is strictly negative, and similarly for h'. But then, the first non-zero derivative of θ will be of the form $g^{(k)}(h')^k + g'h^{(k)}$ with $k \geq 3$, and at least one term strictly negative, and so, taking a Taylor expansion, θ is strictly negative on an interval to the right of $\hat{y}$, and we are done.

□