Simple Optimal Contracts with a Risk-Taking Agent*

Daniel Barron, George Georgiadis, and Jeroen Swinkels†

March 16, 2017

Abstract

Consider an agent who can costlessly add mean-preserving noise to his output. Then, the principal can do no better than offer weakly concave incentives to deter risk-taking. 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 optimal contract and provide conditions under which it takes an intuitive form. We extend our model to analyze costly risk-taking, and we show that the model can be reinterpreted as a dynamic setting in which the agent can manipulate the timing of output.

*We are grateful to Ralph Boleslavsky, Gabriel Carroll, Hugo Hopenhayn, Eddie Lazear, Niko Matouschek, Meg Meyer, Konstantin Milbradt, Dimitri Papanikolaou, Lars Stole, Doron Ravid, Balazs Szentes, Dimitri Vayanos, Glen Weyl, and participants at Cambridge University, Northwestern University, University College London, the University of Miami, the University of Exeter, the University of Warwick, CRETE 2016, the 2016 ASU Theory Conference, and the 2016 Transatlantic Theory Conference 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

Organizations offer incentive schemes to motivate their employees, suppliers, and partners to exert effort. However, in many settings, an agent can game these contracts by taking on risks that increase his expected compensation but do not benefit the principal. For example, portfolio managers can choose riskier investments as well as influence their average returns; executives and entrepreneurs control both the expected profitability of their projects and the distribution over possible outcomes; and salespeople can both invest to increase demand and adjust the timing of the resulting sales.[1] If the principal fails to take these opportunities to game the system into account, then she might offer a contract that encourages risk-taking rather than productive effort.

This paper explores how a principal can effectively motivate an agent who can engage in risk-taking to game his performance contract. In our setting, a principal offers a contract to a potentially liquidity-constrained agent. If the agent accepts this contract, then he exerts costly effort that produces a non-contractible intermediate output. The agent privately observes this output and then can manipulate it by costlessly adding mean-preserving noise to it, which in turn determines the final, contractible output.

Echoing Jensen and Meckling (1976) and others, we argue that the agent's ability to engage in risk-taking fundamentally constrains the principal's ability to offer effective incentives. In Section 3, we show that the agent optimally takes on additional risk whenever the intermediate output lies in a region where his utility is convex in output, effectively concavifying his utility. If the principal and agent are both weakly risk-averse, then the principal finds it optimal to deter such gaming entirely by offering an incentive scheme that directly makes the agent's utility concave in output. Motivated by this logic, we consider how the principal can optimally motivate the agent if she is constrained to offer concave incentives.

[1] See Brown, Harlow and Starks (1996), Chevalier and Ellison (1997), and de Figueiredo, Rawley and Shelef (2014) on portfolio managers; Matta and Beamish (2008) and Repenning and Henderson (2010) on executives; Vereshchagina and Hopenhayn (2009) on entrepreneurs; and Oyer (1998) and Larkin (2014) on sales.

Suppose that the agent is risk-neutral and that higher intermediate output indicates higher effort in the sense of an increasing likelihood ratio. Section 4 proves that the optimal contract is linear. Relative to a strictly concave contract, a linear contract motivates the agent at lower cost because it concentrates pay on higher outcomes, which are more indicative of effort. We show that a linear contract remains optimal regardless of the principal's attitude towards risk, and is uniquely optimal if the principal is risk averse.

This linear optimal contract is tractable and highlights new and interesting economic forces. In particular, we show that as the worst output over which the agent can gamble becomes arbitrarily bad, the optimal contract for any given positive effort gives the agent an arbitrarily large rent and so becomes arbitrarily costly for the principal. Consequently, optimal effort tends to zero.

Section 5 considers optimal incentives if the agent is risk-averse and the principal is risk-neutral. In this setting, we prove that a unique optimal contract exists, with a form determined by a variation of the classic Holmström-Mirlees tradeoff between insurance and incentives. The agent's utility under this incentive scheme is typically a combination of linear and strictly concave segments. To characterize it, we develop a set of perturbations that allow us to adjust the agent's expected utility and expected incentives while respecting concavity. This technique yields an intuitive set of inequalities that characterize the optimal contract. We use this characterization to determine conditions under which the agent's utility is linear in output, and to identify salient features of the optimal contract otherwise. For instance, if the agent's limited liability constraint binds but his participation constraint does not, then the optimal contract makes his utility linear in output below a threshold.

Finally, Section 6 considers three extensions, all of which assume that both principal and agent are risk-neutral. First, we analyze a variant of our model in which the agent must choose his risk-taking distribution 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 or not it will bear fruit. We extend our tools to this setting and give mild conditions under which linear contracts remain optimal. 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 extend our basic intuition to this case and show that 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 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 a stochastic output, but that 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 game the timing of his sales while a strictly concave contract would provide subpar effort incentives.

By identifying how optimal contracts deter the agent from gaming, our analysis sheds light on the kinds of risk-taking that an improperly designed incentive system can encourage. For example, in discussing the 2007-2008 financial crisis, Federal Reserve Chairman at the time Ben Bernanke argued 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)). See also Rajan (2011) for an overview, and Garicano and Rayo (2016) for a case study of the American International Group (AIG). In our model, if organizations do not design their compensation packages to deter these kinds of gaming, they face the prospect of managers and employees exerting too little effort and taking on too much risk.

Our analysis is inspired in part by Diamond (1998), which uses several simple 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. In a contemporaneous paper, Garicano, Matouschek, and Rayo (2016) focuses on the social (rather than incentive) costs of risk-taking in the context of a regulator who attempts to deter a firm from taking socially destructive risks.

Palomino and Prat (2003) considers a delegated portfolio management problem in which, in addition to choosing expected output, the agent chooses 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. In that paper, the optimal threshold for the bonus is determined by the details 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 under which, 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) characterize the optimal contract if the agent can take on socially inefficient risk in a dynamic setting. They argue 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 that result relies on the specific formulation of both agent preferences and the distribution over output. Indeed, Holmström and Milgrom note 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 justifies simple contracts in a setting that lies firmly within the Bayesian tradition.

While the solution concept is 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, the agent might take on additional risk in order 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 the agent might choose a distribution with both less risk and a lower expected output. This second type of gaming is fundamentally different from our setting, which assumes that the agent can only add risk. 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 $s(y) : \mathcal{Y} \to [-M, \infty)$.[2]

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

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

4. Intermediate output x is realized according to $F(\cdot|a) \in \Delta(\mathcal{Y})$, where (i) F is analytic with density f, (ii) $\bar{y}$ is in the support of F for 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 analytic, strictly increasing, and strictly convex.

Let $\mathcal{G} = \{G : \mathcal{Y} \to \Delta(\mathcal{Y}) \mid \mathbb{E}_{G_x}[y] = x \text{ for all } x \in \mathcal{Y}\}$. 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. Define *first-best effort* $a^{FB} \in \mathbb{R}_+$ as the unique effort that maximizes $y - c(y)$ and so satisfies $c'\left(a^{FB}\right) = 1$.

Our model is static, but the fact that the agent can observe x and then gamble prior y being revealed implies a temporal component. We assume that the underlying setting is one in which the principal cannot demand that the agent report x before he has the chance to gamble. Otherwise, the principal could use either a final output that differs from the reported x or a delayed report as evidence of gaming, which would change (but not necessarily eliminate) the agent's incentive to engage in risk-taking. For example, such messages would

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

not be useful if (i) the agent observes x at a random time, and (ii) gambling is instantaneous.

3 Risk-taking and optimal incentives

This section explores how the agent's ability to engage in risk-taking constrains the contracts 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 extraneous risk-taking entirely, but doing so constrains her to incentive schemes that are weakly concave.

Lemma 1. *Suppose $v(\cdot)$ satisfies (IC$_F$)-(LL$_F$) for some $a \geq 0$ and $G \in \mathcal{G}$. Then there exists a weakly concave $\hat{v}(\cdot)$ that satisfies (IC$_F$)-(LL$_F$) for that a and degenerate G 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 efficient than inducing risk-taking.

Given Lemma 1, we henceforth restrict attention to contracts that make the agent's utility 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. Given this constraint, 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 also maximizes (Obj). 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.[3]

Lemma 2. *Fix $a \geq 0$. 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 6 in Appendix D.[4] The profit-maximizing contract is unique if at least one player is strictly risk-averse because Jensen's Inequality

[3]Given domain $D \subseteq \mathbb{R}$, a function $u : D \to \mathbb{R}$ is strictly concave if for *every* $x \in D$, there exists an affine function $l : \mathbb{R} \to \mathbb{R}$ such that $u(x) = l(x)$ and $u(x') < l(x')$ for all $x' \neq x$.

[4]Appendix D may be found online at https://sites.google.com/site/danielbarronecon/

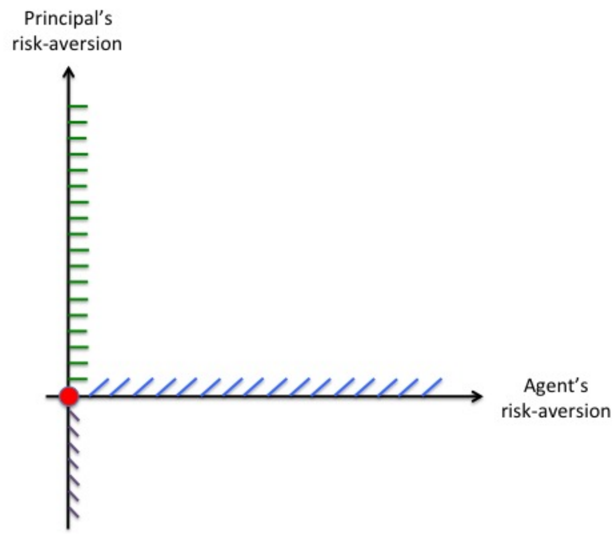

Figure 1: Roadmap

implies that a convex combination of two different contracts that implement a would also implement a and would give the principal a strictly higher payoff.

Roadmap of our Main Results

The next two sections build on Lemma 1 to characterize optimal incentive schemes. 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 give weak conditions under which a unique profit-maximizing contract exists, and we develop techniques to characterize it. In some situations, the profit-maximizing incentive scheme makes the agent's utility 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 y such that $F(y|a)$ has a negative likelihood ratio.

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 6 in Appendix D for a proof. Therefore, the characterizations in Sections 4 and 5 shed light on profit-maximizing incentives in the rest of this quadrant. In particular, 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. The following proposition proves that a linear contract is optimal if the agent is risk-neutral.

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.

For the moment, suppose the principal is risk neutral, so that $\pi(y) \equiv y$, and assume $u_0 + c(a) < -M$ so (IR) never binds. The distribution over intermediate output $F(\cdot|a)$ satisfies MLRP, so if we ignored the requirement that $s(\cdot)$ be concave, then an approximately optimal incentive scheme would pay large rewards for outcomes near $\bar{y}$ and $-M$ otherwise.[5] However, such a contract is convex and hence susceptible to gaming.

Suppose a^* is the optimal effort level, and let $\tilde{s}(\cdot)$ be a strictly concave contract that implements a^*. Consider the unique linear contract $s^L(\cdot)$ that starts at $\tilde{s}(\underline{y})$ and gives the agent the same expected payoff as $\tilde{s}(\cdot)$. Then $s^L(\cdot)$ must cross $\tilde{s}(\cdot)$ exactly once from below. If the slope of $s^L(\cdot)$ is larger than 1, then the proof of Proposition 1 shows that the principal strictly prefers to induce first-best effort with the linear contract $s^L_{a^{FB}}(\cdot)$ rather than offer $\tilde{s}(\cdot)$. If $s^L(\cdot)$ has a slope less than 1, then it assigns larger rewards to high output and smaller rewards to low output than $\tilde{s}(\cdot)$, and hence motivates the agent to exert strictly more effort than a^*. Consequently, we prove that $s^L(\cdot)$ leads to a higher payoff for the principal. So a linear contract is at least as profitable as $\tilde{s}(\cdot)$ and hence $s^L_{a^*}(\cdot)$ is optimal.

If $\pi(\cdot)$ is strictly concave, then the principal's marginal utility of money is strictly decreasing in $y - s(y)$. Therefore, so long as the slope of $s^L(\cdot)$ is no larger than 1, the principal strictly prefers it to $\tilde{s}(\cdot)$ holding effort fixed, since $s^L(y) - \tilde{s}(y)$ is positive exactly where the principal's marginal utility of money is low. So the argument above holds *a fortiori* and a linear contract continues to be optimal. Note that Lemma 2 implies that $s^L_{a^*}(\cdot)$ is uniquely optimal if $\pi(\cdot)$ is strictly concave. If the principal 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 result in identical expected payoffs.

In some applied settings, the principal might have risk-seeking preferences

[5]Without (Conc), no optimal contract would exist in this setting, since the reward for high output can be made arbitrarily large.

over output. Indeed, there is evidence that organizations in certain settings might encourage risk-taking from their agents because the market gives the organization itself a convex payoff. For instance, Rajan (2011) notes that investment funds and banks were motivated to take on significant amounts of risk in the years leading up to the 2007 financial crisis, while Chevalier and Ellison (1997) argue that mutual funds earn disproportionate profits from outperforming their competitors, which encourages lagging funds to take on excess risk.

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, if the agent is risk-neutral, then we can modify the argument from Proposition 1 to show that a linear contract is optimal.

Remark 1. *Let $u(s) \equiv s$ and $\pi(\cdot)$ 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 Remark 1, 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 relaxed problem because $\pi^c(\cdot)$ is concave. Given $s^L_{a^*}(\cdot)$, the agent is indifferent among distributions $G \in \mathcal{G}$, and 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 $\lim_{\underline{y}\to-\infty} a^* = 0$. If $c'''(a) \geq 0$ for all a and $\pi(y) \equiv y$, then a^* is increasing in $\underline{y}$.*

To see the argument, observe that Proposition 1 implies that the principal's expected payment from inducing $a^* \geq 0$ equals $E_{F(\cdot|a^*)}[\pi(y - c'(a^*)(y - \underline{y}) + w)]$. As $\underline{y} \to -\infty$, the limited liability constraint eventually binds at $s^L_{a^*}(\cdot)$ and so $w = M$. But then implementing an $a^* > 0$ 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 her payoff from inducing effort a^* equals $a^* - E_{F(\cdot|a^*)}\left[s^L_{a^*}(y)\right]$; if $c'''(a) \geq 0$, then we can manipulate this expression to directly perform comparative statics on a^*.

5 Optimal contracts if the agent is risk averse

This section explores the implications of Lemma 1 in a setting with a risk-averse agent and a risk-neutral principal and characterizes the unique contract that implements a given $a \geq 0$ at maximum profit.

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

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

Given that $v(\cdot)$ is constrained to be concave, replacing (IC) with (IC-FOC) entails no loss if $F(\cdot|\cdot)$ satisfies weak regularity conditions.[6,7] For a fixed effort $a \geq 0$, define the principal's problem (P) as maximizing (Obj) subject to (IC-FOC), (IR), (LL), and (Conc).

[6] 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, $E_{F(\cdot|a)}[x] = a$ implies that $\int F_{aa}(x|a)dx = 0$, so this condition holds if $F_{aa}(\cdot|a)$ never changes sign from negative to positive.

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

Define $\rho(\cdot)$ as the function that maps $\frac{1}{u'(\cdot)}$ into $u(\cdot)$; that is, for every $w \in (\underline{w}, \infty)$, $\rho\left(\frac{1}{u'(w)}\right) = u(w)$. This function is well-defined because $u'(\cdot)$ and $u(\cdot)$ are strictly monotonic. Note that $\rho^{-1}(v(x))$ equals the marginal cost to the principal of giving the agent extra utility at x. 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* x as

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

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

If the principal could offer non-concave contracts, then the optimal contract would set $n(x) = 0$ for all x. Indeed, this is the *Holmström-Mirrlees contract* characterized in Holmström (1979). However, this contract is not necessarily concave. Instead, we characterize the profit-maximizing contract by identifying a set of perturbations that respect concavity.

Given $v(\cdot)$, say that an interval $[x_L, x_H]$ is a *linear segment* if $v(\cdot)$ is linear on $[x_L, x_H]$, but not on any strictly larger interval. Say that x is *free* if it is not on the interior of any linear segment of $v(\cdot)$. Intuitively, if x is free, then $v(\cdot)$ is strictly concave on at least one side of x. We can make small perturbations to this strictly concave side without violating concavity. The points $\underline{y}$ and $\bar{y}$ are always free.

Consider the following two perturbations, formally defined in Appendix B.

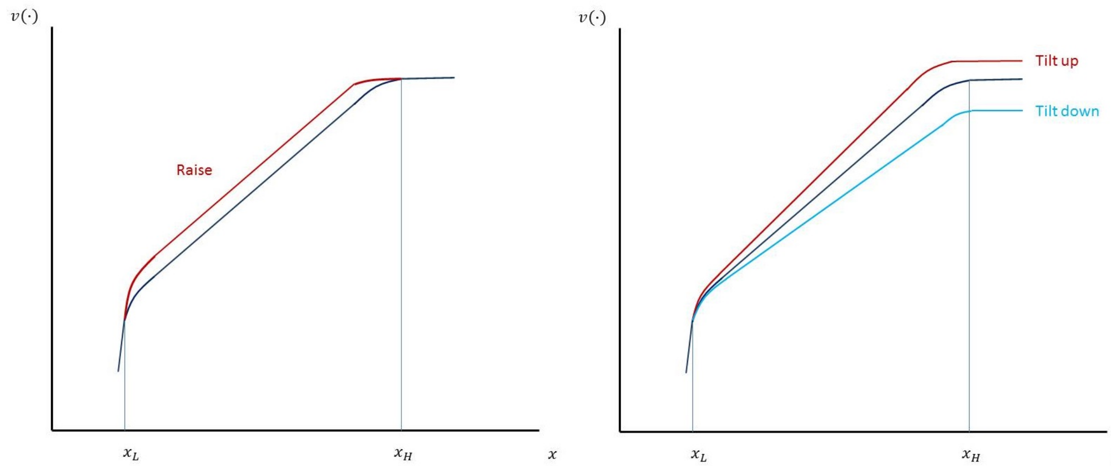

Figure 2: *Raise* and *tilt*. These perturbations require care around x_L and x_H to ensure that concavity is preserved. For this reason, we need both x_L and x_H to be free for *raise*. For *tilt up*, we need x_L to be free, while x_H must be free for *tilt down*.

Raise increases $v(\cdot)$ 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 that 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. Appendix B shows that these non-concavities can be repaired so long as the relevant endpoints are free. Figure 2 presents simplified representations of these perturbations.

Each of *raise* and *tilt* simultaneously affects both (IR) and (IC-FOC). Appendix B shows that these effects are linearly independent, so we can construct a combination of the two perturbations to separately affect the agent's expected utility and effort incentives in any desired manner. Therefore, so long as there exists at least one free point $\hat{x} < \bar{y}$ such that $v(\hat{x}) > \underline{u}$, we can use these perturbations on $[\hat{x}, \bar{y}]$ to establish candidate shadow values λ and μ for (IR) and (IC-FOC).[8]

Consider *raising* $v(\cdot)$ on an interval between two free points $x_L < x_H$. As we argued above, we can undo the effects of this perturbation using a

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

combination of *raise* and *tilt* on $[\hat{x}, \bar{y}]$. For $v(\cdot)$ to be optimal, the net cost of this perturbation must be positive, or

$$\int_{x_L}^{x_H} n(x) f(x|a) dx \geq 0. \tag{3}$$

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

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

$$\int_{x_L}^{x_H} n\left(x\right)\left(x - x_L\right) f\left(x|a\right) dx + \left(x_H - x_L\right) \int_{x_H}^{\bar{y}} n(x) f(x|a) \geq 0, \tag{4}$$

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

$$\int_{x_L}^{x_H} n\left(x\right)\left(x - x_L\right) f\left(x|a\right) dx + \left(x_H - x_L\right) \int_{x_H}^{\bar{y}} n(x) f(x|a) \leq 0. \tag{5}$$

Our characterization combines these perturbations with the usual complementary slackness condition that $\lambda = 0$ if (IR) is slack (in which case $v(\underline{y}) = \underline{u}$).

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

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

and for any $x_L < x_H$,

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

2. *if x_L is free, then (4) holds;*

3. if x_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 2. *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 Appendix B. Intuitively, the necessity of GHM follows from the arguments above. To establish sufficiency, we show that 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 2 implies that optimal incentive schemes must satisfy several intuitive properties. For any free $x \in (\underline{y}, \bar{y})$, say x is a *kink point* of $v(\cdot)$ if two linear segments meet at x, and a *point of normal concavity* otherwise.

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

The optimal contract must equal the Holmström-Mirrlees contract in expectation between any two free points $x_L < x_H$ with $v(x_L) > \underline{u}$, since (3) holds with equality between these points. If x is a point of normal concavity, then there exist two free points that are arbitrarily close to x. Since (3) holds with equality between these points, taking a limit as these points approach x proves that $n(x) = 0$. Moreover, Lemma 3 in Appendix B proves that absent (Conc), the principal would want to increase payments near the ends of a linear interval and decrease them somewhere in the middle of that interval. Therefore, $n(x) \leq 0$ at the right endpoint of any linear segment, which includes any kink point.

Suppose (IR) is slack, which implies that (LL) binds. Then Corollary 2 implies that the optimal contract is linear everywhere that $l(\cdot|a)$ 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, so that (LL) binds. Define y_0 by $l(y_0|a) = 0$. Then v is linear on $[\underline{y}, y_0]$.*

Intuitively, the agent's incentive to exert effort is decreasing in his payment following $y < y_0$, because $f_a(y|a) < 0$. 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 $v(y_0)$ both improves the agent's incentives and decreases the principal's expected payment. This perturbation is feasible so long as (IR) does not bind.

Finally, we identify sufficient conditions under which the optimal incentive scheme is linear, as well as conditions under which it coincides with the Holmström-Mirrlees contract.

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).*

1. *If $\rho(\lambda + \mu l(\cdot|a))$ is convex, then $v(\cdot)$ is linear.*

2. *If $\rho(\lambda + \mu l(\cdot|a))$ is concave and (LL) does not bind, then $n(\cdot) \equiv 0$. If (LL) binds, then $v(\cdot)$ has at most one linear segment on which $n(\cdot)$ is not identically 0, and this segment begins at $\underline{y}$.*

Sufficient (but far from necessary) conditions for $\rho(\lambda + \mu l(\cdot|a))$ to be convex (concave) are that $\rho(\cdot)$ and $l(\cdot|a)$ are convex (concave).

To see this result, note that if $\rho(\lambda + \mu l(\cdot|a))$ is concave and the agent's liability constraint does not bind, then the Holmström-Mirrlees contract is concave and hence profit-maximizing in our setting as well. If $\rho(\lambda+\mu l(\cdot|a))$ is convex, then $n(\hat{x}) \leq 0$ at any free $\hat{x}$ by Corollary 2, so $v(\hat{x}) \leq \rho(\lambda + \mu l(\hat{x}|a))$. But $\rho(\lambda+\mu l(\cdot|a))$ is convex and $v(\cdot)$ is concave, so either $v(x) < \rho(\lambda+\mu l(x|a))$ for all $x > \hat{x}$, or $v(x) > \rho(\lambda + \mu l(x|a))$ for all $x < \hat{x}$. Since $\underline{y}$, $\hat{x}$, and $\bar{y}$ are all free, either violates (3). Hence, $v(\cdot)$ has no interior free points and so must be linear.

These corollaries show that the profit-maximizing contract is particularly simple if (IR) does not bind, or if $\rho(\lambda + \mu l(\cdot|a))$ is either always convex or always concave. However, Proposition 2 imposes substantial structure on the optimal contract even if these conditions do not hold. To illustrate this point, we establish mild conditions under which the optimal $v(\cdot)$ has no more than one linear segment, which must begin at $\underline{y}$. Essentially, for the optimal contract to have two linear segments, Corollary 2 requires that $\rho(\lambda + \mu l(\cdot|a))$ must be at least as steep as the first linear segment but no steeper than the second linear segment. Hence, $\rho(\lambda + \mu l(\cdot|a))$ must be first strictly concave and then weakly convex on some region.

This case can be ruled out with reasonably weak conditions. For any interval $X \subseteq \mathbb{R}$ and any twice continuously differentiable function $h : X \to \mathbb{R}_+$, define the *concavity of* $h(\cdot)$ as the largest value t for which h^t is concave,

$$con\,(h) \equiv \inf_X \left(1 - \frac{hh''}{(h')^2}\right).$$

Note that $h(\cdot)$ is concave if $con(h) \geq 1$ and is log-concave if $con(h) \geq 0$.[9]

Proposition 3. *Assume that* $con\,(\rho') + con\,(l_x) > -1$ *and that* $u(\cdot)$ *is analytic. Then for any* $a \geq 0$*, the contract* $v(\cdot)$ *that solves (P) has at most one linear segment, which must begin at* $\underline{y}$*.*

The condition $con(\rho') + con(l_x) > -1$ is satisfied by many natural cases, including if either $u(w) = \log w$ and $\log(l_x)$ is strictly concave or $u(w)$ is *HARA* and $-\frac{1}{l_x^2}$ is strictly concave.[10]

If $\underline{u} = -\infty$, so that the agent is not liquidity constrained, then Lemma 2 does not apply. Nevertheless, we show in Appendix D that a unique solution exists so long as $u'(\cdot)$ is not excessively convex. Furthermore, our characterization extends to that setting, so an incentive scheme implements $a \geq 0$ at maximum profit in that setting if and only if it is GHM.

[9] See Prekopa (1973) and Borell (1975) for details.

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

6 Extensions and Reinterpretations

This section considers three extensions of the baseline model, 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 that is increasing in the amount of gambling that he chooses to do. 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 these extensions may be found in Appendix C.[11]

6.1 Risk-Taking Before Intermediate Output is Realized

This section proves that a linear contract is optimal even if the agent cannot condition his risk-taking distribution on the 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 and he receives u_0 while the principal receives 0.

3. The agent chooses an effort $a \geq 0$ and a distribution $G(\cdot|a) \in \Delta(\mathcal{Y})$ subject to the constraint $\mathbb{E}_G [x|a] = \phi(a)$. We assume that $\phi(\cdot)$ is a smooth, increasing function such that $c(\phi^{-1}(\cdot))$ is strictly convex.

4. Final output $y \sim H(\cdot|a)$ is realized, where $H(y|a) = \int F(y|x)\, dG(x|a)$, and the agent is paid $s(y)$.[12] We assume that $F(\cdot|x)$ has density $f(\cdot|x)$, satisfies strict MLRP in x, and $\mathbb{E}_{F(\cdot|x)}[y] = x$.

[11] Available online at https://sites.google.com/site/danielbarronecon/

[12] In effect, $G(\cdot|a)$ determines a stochastic "intermediate effort" x, which then determines output according to $F(\cdot|x)$. Alternatively, we could have modelled the agent as choosing a random, additively separable noise term that affects output. We do not believe that optimal contracts would be linear in that alternative framework.

The principal and agent earn $y - s(y)$ and $s(y) - c(a)$, respectively. The distribution $G(\cdot|a)$ 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 $G(\cdot|a)$ 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 $G(\cdot|a)$ leads to a riskier distribution over output: $H(\cdot|a)$ is increasing in $G(\cdot|a)$ in the sense of second-order stochastic dominance.

Given a contract $s(\cdot)$ and a realization x from $G(\cdot|a)$, 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). It follows by an argument similar to Lemma 1 that the agent will optimally choose G such that $E_{G(\cdot|a)}[V_s(x)] = V_s^c(\phi(a))$. Since $\mathbb{E}_{G(\cdot|a)}\left[\mathbb{E}_{F(\cdot|x)}[y]\right] = \phi(a)$ for any $G(\cdot|a)$, the principal's problem is

$$\begin{aligned}
\max_{a, V_s(\cdot)} \quad & \phi(a) - V_s(\phi(a)) \qquad (7)\\
\text{s.t.} \quad & a \in \arg\max_{\tilde{a}} \{V_s(\phi(\tilde{a})) - c(\tilde{a})\} \\
& V_s(\phi(a)) - c(a) \geq u_0 \\
& V_s(\cdot) \ \text{weakly concave},
\end{aligned}$$

with the additional restriction that $V_s(\cdot)$ must be the concave closure of (6) for some $s(\cdot) \geq -M$.

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

Proposition 4. *For optimal effort $a^* \geq 0$, define $s^*(y) = \frac{c'(a^*)}{\phi'(a^*)}\left(y - \underline{y}\right) - w$, where $w = \min\left\{M, \frac{c'(a^*)}{\phi'(a^*)}\left(\phi(a^*) - \underline{y}\right) - c(a^*) - u_0\right\}$. Then $a^* \leq a^{FB}$ and*

$s^*(\cdot)$ *is optimal.*

To see the argument, relax this problem by ignoring the constraint that $V_s(\cdot)$ must be the concave closure of (6). Redefine $\beta = \phi(a)$. Then this relaxed problem is a special case of (Obj)-(Conc) in which the distribution over intermediate output is degenerate at a. Consequently, we can modify the proof of Theorem 1 to show that a linear $V_s(\cdot)$ is optimal. But a linear $V_s(\cdot)$ can be implemented by a linear $s(\cdot)$ because $\mathbb{E}_{F(\cdot|x)}[y] = x$, so a linear contract is optimal. Note that it is not uniquely so; since both parties are risk-neutral, any contract such that $V_s^c(y) = \frac{c'(a^*)}{\phi'(a^*)}(y - \underline{y}) - w$ is also 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 $d(y) = y^2 - \underline{y}^2$. 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.[13]

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 $\tilde{v}^c(x)$; *i.e.*, $\max_{G_x}\{\mathbb{E}_{G_x}[\tilde{v}(y)]\} = \tilde{v}^c(x)$ for all 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 $d(\cdot)$.

[13]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 $\tilde{s}^*(y) = \tilde{c}'(a)(y - \underline{y}) + d(y) - \tilde{w}$, where $\tilde{w} = \min\left\{M, \tilde{c}'(a)(a - \underline{y}) - \tilde{c}(a) - u_0\right\}$. Then $\tilde{s}^*(\cdot)$ is optimal.*

As in Theorem 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 $s(\cdot)$ has the desired shape. Intuitively, the profit-maximizing contract is the most convex contract that does not induce the agent to gamble.

Importantly, the slope of $\tilde{v}(\cdot)$, $\tilde{c}'(a)$, does not equal the slope $c'(a)$ of the linear contract from Theorem 1. In fact, one can show that the principal pays strictly less than in the optimal contract from Proposition 1, which is intuitive: if the agents finds risk-taking costly, then the principal can offer some convex incentives without inducing gaming.

6.3 Manipulating the Timing of Output[14]

As noted in the introduction, in many settings the agent can game his contract by manipulating the timing of his output. 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, and the agent has wealth M. 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 $F(\cdot|a) \in \Delta(\mathcal{Y})$.

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

5. The agent chooses a mapping from time t to output at time t, $y_x : [0,1] \to \mathcal{Y}$, subject to $\int_0^1 y_x(t)dt = x$.

6. The agent is paid $\int_0^1 s(y_x(t))dt$.

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

Crucially, 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 practice, and 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 amount of time for which $y_x(t) \leq y$.[15] Then $G_x(\cdot)$ is a distribution: it is increasing, has $G_x(\underline{y}) = 0$, and also has $G_x(\bar{y}) = 1$ since $t \in [0,1]$. 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$_F$); consequently, the results from Sections 3 and 4 apply.

Remark 2. *The optimal contracting problem in this setting coincides with (Obj$_F$)-(LL$_F$) with $u(x) \equiv x$ and $\pi(x) \equiv x$. 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), who find that salespeople facing convex incentives concentrate their sales. Conversely,

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

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.

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 would pursue in our model. 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 show how an agent can blunt convex incentives by engaging in risk-taking. As emphasized in Makarov and Plantin (2015), an agent might also engage in risk-taking in an attempt to manipulate his superiors into thinking that he has high ability. Further research is required to better understand how risk-taking interacts with effort incentives and private information.

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.

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

Kelly, B. and Jiang, H., 2014. Tail risk and asset prices. *Review of Financial Studies*, 27(10), pp.2841-2871.

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 (No. w16367). 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 $v^*(\cdot) \equiv \frac{1}{2}(v(\cdot) + \tilde{v}(\cdot))$ clearly 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

Consider some $s(\cdot)$ that implements a. Define $s^L(y) \equiv s(\underline{y}) + m(y - \underline{y})$, where $m = \frac{1}{a-\underline{y}}\left(\mathbb{E}_{F(\cdot|a)}[s(y)] - s(\underline{y})\right)$.

Suppose $m > 1$. Recall that $s^L_{a^{FB}}(\cdot)$ implements a^{FB}. There are two cases to consider. If $s^L_{a^{FB}}(\underline{y}) = -M$, then since $s^L_{a^{FB}}$ has slope 1, $s^L_{a^{FB}}(y) < s^L(y)$ for all $y > \underline{y}$. Therefore,

$$\mathbb{E}_{F(\cdot|a)}\left[\pi(y - s(y))\right] \leq \pi\left(\mathbb{E}_{F(\cdot|a)}\left[y - s(y)\right]\right) = \pi\left(\mathbb{E}_{F(\cdot|a)}\left[y - s^L(y)\right]\right) <$$
$$\pi\left(\mathbb{E}_{F(\cdot|a)}\left[y - s^L_{a^{FB}}(y)\right]\right) = \mathbb{E}_{F(\cdot|a^{FB})}\left[\pi\left(y - s^L_{a^{FB}}(y)\right)\right],$$

where the first inequality follows from Jensen's inequality, the first equality follows because $\mathbb{E}_{F(\cdot|a)}[s^L(y)] = \mathbb{E}_{F(\cdot|a)}[s(y)]$ by choice of m, the strict inequality follows because $s^L(y) < s^L_{a^{FB}}(y)$ for all $y > \underline{y}$, and the final equality follows because $y - s^L_{a^{FB}}(y)$ is constant in y.

If instead $s^L_{a^{FB}}(\underline{y}) > -M$, then (IR) holds with equality under $s^L_{a^{FB}}(\cdot)$. So

$$\mathbb{E}_{F(\cdot|a)}\left[\pi(y - s(y))\right] \leq \pi\left(\mathbb{E}_{F(\cdot|a)}\left[y - s(y)\right]\right) \leq \pi\left(a^{FB} - c(a^{FB}) + u_0\right) =$$
$$\pi\left(\mathbb{E}_{F(\cdot|a^{FB})}\left[y - s^L_{a^{FB}}(y)\right]\right) = \mathbb{E}_{F(\cdot|a^{FB})}\left[\pi\left(y - s^L_{a^{FB}}(y)\right)\right],$$

where the first inequality follows from Jensen's inequality, the second inequality holds because $E_{F(\cdot|a)}[s(y)] \geq c(a) + u_0$ to satisfy (IR) and $a - c(a) \leq a^{FB} - c(a^{FB})$ and strictly so if $a \neq a^{FB}$, the first equality follows because $\mathbb{E}_{F(\cdot|a^{FB})}\left[s^L_{a^{FB}}(y)\right] = u_0 + c(a^{FB})$, and the final equality holds because $y - s^L_{a^{FB}}(y)$ is constant in y. Therefore, any $s(\cdot)$ corresponding to an $s^L(\cdot)$ with $m > 1$ is dominated by $s^L_{a^{FB}}(\cdot)$, and strictly so if $s(\cdot)$ implements $a \neq a^{FB}$.

Suppose instead that $m \leq 1$. Because $s(\cdot)$ is concave with $s^L(\underline{y}) = s(\underline{y})$ and $\mathbb{E}_{F(\cdot|a)}\left[s(y) - s^L(y)\right] = 0$, there must exist some $\tilde{y} > \underline{y}$ such that $s^L(y) \leq s(y)$ if and only if $y \leq \tilde{y}$. For any $y \in \mathcal{Y}$,

$$\pi(y - s^L(y)) - \pi(y - s(y)) = \int_{y-s(y)}^{y-s^L(y)} \pi'(x)dx.$$

If $y \le \tilde{y}$, then $y - s^L(y) \ge y - s(y)$. Furthermore, $y - s^L(y) \le \tilde{y} - s^L(\tilde{y})$ because $m \le 1$. Therefore, $\pi'(x) \ge \pi'(\tilde{y} - s^L(\tilde{y}))$ for $x \le \tilde{y} - s^L(\tilde{y})$ and so

$$\pi(y - s^L(y)) - \pi(y - s(y)) \ge \pi'(\tilde{y} - s^L(\tilde{y}))(s(y) - s^L(y)).$$

If $y > \tilde{y}$, then $\tilde{y} - s^L(\tilde{y}) \le y - s^L(y) \le y - s(y)$, $\pi'(x) \le \pi'(\tilde{y} - s^L(\tilde{y}))$ for $x \ge \tilde{y} - s^L(\tilde{y})$, and again

$$\pi(y - s^L(y)) - \pi(y - s(y)) \ge \pi'(\tilde{y} - s^L(\tilde{y}))(s(y) - s^L(y)).$$

Consequently,

$$\mathbb{E}_{F(\cdot|a)}\left[\pi(y - s^L(y)) - \pi(y - s(y))\right] \ge \pi'(\tilde{y} - s^L(\tilde{y}))\mathbb{E}_{F(\cdot|a)}\left[s(y) - s^L(y)\right] = 0.$$

Next, we claim $s^L(\cdot)$ induces effort $\tilde{a} > a$. Because $s^L(y) - s(y)$ is negative and then positive, $\mathbb{E}_{F(\cdot|a)}\left[s^L(y) - s(y)\right] = 0$, and $\frac{f_a}{f}$ is monotonically increasing in y, Beesack's Inequality[16] implies that

$$\int_{\underline{y}}^{\bar{y}} (s^L(y) - s(y))\frac{f_a(y|a)}{f(y|a)}f(y|a)dy \ge 0.$$

Hence, (IC) is slack at a under $s^L(\cdot)$. So $s^L(\cdot)$ implements $\tilde{a} > a$. Since $1 - m > 0$, the principal prefers effort $\tilde{a}$ to effort a under contract $s^L(y)$.

Finally, note that $s^L(y) - s^L_{\tilde{a}}(y) > 0$ for all $y \in \mathcal{Y}$. Combining the previous arguments yields

$$\mathbb{E}_{F(\cdot|a)}\left[\pi(y - s(y))\right] \le \mathbb{E}_{F(\cdot|a)}\left[\pi(y - s^L(y))\right] \le \mathbb{E}_{F(\cdot|\tilde{a})}\left[\pi(y - s^L(y))\right] \le \mathbb{E}_{F(\cdot|\tilde{a})}\left[\pi(y - s^L_{\tilde{a}}(y))\right].$$

The preceding shows that any $s(\cdot)$ that implements a is dominated by $s^L_{\tilde{a}}(\cdot)$ for some $\tilde{a} \in [0, a^{FB}]$, and strictly so if $a > a^{FB}$. So $a^* \le a^{FB}$, and $s^L_{a^*}(\cdot)$ implements a^* at maximum profit. ■

[16] Beesack's inequality asserts that if $g(\cdot)$ and $h(\cdot)$ are two integrable functions such that $g(\cdot)$ is first negative then positive, $\int g(z)dz = 0$, and $h(\cdot)$ is non-decreasing, then $\int g(z)h(z)dz \ge 0$. See Beesack (1957) for details.

A.3 Proof of Remark 1

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

Modify (Obj)-(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 $s_a^L(\underline{y}) = -M$, so Proposition 1 implies that $s_a^L(\cdot)$ implements a at maximum profit in this modified problem. So the principal's expected payoff equals $\mathbb{E}_{F(\cdot|a)}\left[\pi^c(x - s_a^L(x))\right]$ in this modified problem.

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

A.4 Proof of Corollary 1

Suppose $a^* > 0$ equals the optimal effort level. Then Proposition 1 implies that the principal's expected payoff equals

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

Note that $c'(a)(a - \underline{y}) - c(a) - u_0 > M$ whenever $\underline{y} < a - \frac{c(a)+u_0+M}{c'(a)}$. So for $\underline{y}$ sufficiently negative, the principal's payoff is bounded above by

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

because $\pi(\cdot)$ is concave. For any $a > 0$, $c'(a) > 0$ and hence $(1-c'(a))+c'(a)\underline{y}+M < -u_0$ for $\underline{y}$ sufficiently negative. But then $a > 0$ is strictly dominated by $a = 0$ and $s(\cdot) = u_0$. The same argument holds for any $a > 0$, so a^* must converge to 0 as $\underline{y} \to -\infty$.

Suppose $c'''(\cdot) \geq 0$ and $\pi(y) \equiv y$. The principal solves

$$a^* \in \arg\max_{a\geq 0} \left\{a - c'(a)(a - \underline{y}) + w(a)\right\} , \tag{8}$$

where $w(a) = \min\left\{M, c'(a)(a - \underline{y}) - c(a) - u_0\right\}$. Define $\hat{a}$ as the smallest $a \geq 0$ such that $c'(a)(a - \underline{y}) - c(a) - u_0 \geq M$ for all $a \geq \hat{a}$, and note that the LHS is strictly increasing in a.

First, we claim that if $w(a^*) < M$ (*i.e.*, if (LL) is slack), then $a^* = a^{FB}$. Towards a contradiction, suppose that $w(a^*) < M$ and $a^* < a^{FB}$. Then the principal's expected profit is equal to $a^* - c(a^*) + u_0$. However, because $a^* < \min\{\hat{a}, a^{FB}\}$ and $a - c(a)$ is increasing for all $a \leq a^{FB}$, the principal can obtain a strictly bigger expected profit by choosing $\min\{\hat{a}, a^{FB}\}$.

Next, we restrict attention to the case in which $w(a^*) \geq M$. The above claim implies that it suffices to consider effort levels $a \geq \hat{a}$. In this case, (8) can be written as $a^* \in \arg\max_{a\geq\hat{a}} \left\{a - c'(a)(a - \underline{y}) + M\right\}$. This problem is strictly concave in a because $c'''(\cdot) \geq 0$, and $a^* = \max\left\{\hat{a}\,, a^{foc}\right\}$, where a^{foc} solves the first-order condition $1 - c'(a^{foc}) - c''(a^{foc})(a^{foc} - \underline{y}) = 0$.

In summary, we have that

$$a^* = \begin{cases} \max\left\{\hat{a}\,, a^{foc}\right\} & \text{if } a^{FB} < \hat{a} \\ a^{FB} & \text{otherwise.} \end{cases}$$

Using the implicit function theorem and that $c'''(\cdot) \geq 0$, it follows that both $\hat{a}$ and a^{foc}, and hence a^* are increasing in $\underline{y}$. ■

B Proofs for Sections 5

Consider an interval $[x_L, x_H]$. The perturbation *raise* increases v parallel to itself by ε_r utils over the interval $[x_L, x_H]$. The initial rate of change of the agent's utility at any given output with respect to this perturbation is thus

$$r_{x_L,x_H}(x) = \begin{cases} 1 & x \in [x_L, x_H] \\ 0 & else \end{cases}.$$

The perturbation *tilt* increases the slope of the contract (in utils per unit of output) by ε_t on $[x_L, x_H]$, and raises the contract by a constant $\varepsilon_t (x_H - x_L)$ utils beyond x_H. Therefore, the initial rate of change of payoffs with respect to *tilt* is

$$t_{x_L,x_H}(x) = \begin{cases} 0 & x \leq x_L \\ x - x_L & x \in (x_L, x_H) \\ x_H - x_L & x \geq x_H \end{cases}.$$

Our first result proves an important properties of any contract that is GHM.

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

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

If $v(x_L) > \underline{u}$, then such an $\tilde{x}$ exists in $(x_L, \hat{x})$ as well. But, somewhere on (x_L, x_H), $n(x) \geq 0$.

Proof. Note that for $x > x_H$, $t_{\hat{x},x_H}(x) = x_H - x_{\hat{x}} = (x_H - x_{\hat{x}}) r_{x_H,\bar{y}}(x)$. 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 $x_H > \underline{y}$, $v(x_H) > \underline{u}$. We thus have $\int n(x) r_{x_H,\bar{y}}(x) f(x|a)\, dx = 0$ by Definition 1.1. Hence, by Definition

1.3, we have

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

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

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

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

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

Finally, since $\int n(x)\, r_{x_L,x_H}(x)\, f(x|a)\, dx \geq 0$, and since we have established that $n(x)$ is weakly negative somewhere on (x_L, x_H), we must also have $n(x)$ weakly positive somewhere on the same interval.

□

Next, we prove some preliminary properties of optimal incentives schemes. Lemma 1 shows that any optimal incentive scheme $v(\cdot)$ must be unique. 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 $\tilde{x} \in \mathcal{Y}$ such that if we replace $v(x)$ by a constant $v(\tilde{x})$ to the right of $\tilde{x}$, 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.

B.1 Proof of Proposition 2

B.1.1 Preliminaries

Definition 1 and Proposition 2 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 x at which there exists a supporting plane L such that $L(x') > v(x')$ for all $x' \neq x$.

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 $\hat{x}$ be any point of normal concavity. Then, for each δ, there is a point in $(\hat{x} - \delta, \hat{x} + \delta) \backslash \hat{x} \cap C_v$. From this, it follows that for each $\varepsilon > 0$, there exists $x_L < x_H$ such that $x_L, x_H \in C_v$, and such that $x_L, x_H \in [\hat{x} - \varepsilon, \hat{x} + \varepsilon]$.*

Proof of Claim. We will show first that for each δ, there is a point in $(\hat{x} - \delta, \hat{x} + \delta) \backslash \hat{x} \cap C_v$. To see that this suffices to show the second part, apply the result first to find a point x_1 in $C_v \cap [\hat{x} - \varepsilon, \hat{x} + \varepsilon]$. Apply the result again to find x_2 in $C_v \cap [\hat{x} - \delta, \hat{x} + \delta]$, where $\delta = (1/2)|x_1 - \hat{x}|$, and finally take x_L and x_H as the smaller and larger of x_1 and x_2.

So, fix $\delta > 0$. Since $\hat{x}$ is not on the interior of a linear segment and not a kink point, there is at least one side of $\hat{x}$, wlog the right side, such that $v(\cdot)$ is not linear on $(\hat{x}, \hat{x}+\delta)$. Let $S(\cdot)$ be the correspondence which for each x assigns the set of slopes of supporting planes at x, and let $s(\cdot)$ be any selection from $S(\cdot)$. Note that since v is concave, for any $x'' > x'$, $\max\{S(x'')\} \leq \min\{S(x')\}$, and hence s is decreasing. Assume first that there is a point $\tilde{x} \in (\hat{x}, \hat{x} + \delta)$

where $s(\cdot)$ jumps discontinuously downward, say from s'' to $s' < s''$. Then, the supporting plane at $\tilde{x}$ with slope $(s' + s'')/2$ qualifies. Assume instead that $s(\cdot)$ is continuous on $(\hat{x}, \hat{x} + \delta)$. It cannot be everywhere constant, since $v(\cdot)$ is not linear on $(\hat{x}, \hat{x} + \delta)$. Hence, since $s(\cdot)$ is continuous, there is a point $\tilde{x}$ at which it is strictly decreasing, so that in specific, $s(\tilde{x}) < s(x)$ for all $x < \tilde{x}$, and $s(\tilde{x}) > s(x)$ for all $x > \tilde{x}$. The supporting plane at $\tilde{x}$ with slope $s(\tilde{x})$ 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 x_L and x_H such that $\int n(x) r_{x_L,x_H} f(x)\, dx < 0$. If either x_L or x_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 2 when each restriction to a free point is tightened to a restriction to C_v.

B.1.2 Properties of the Perturbations

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_{x_L,x_H} or t_{x_L,x_H} in each case with ε_p positive or negative. Second, for some $\hat{x} \in C_v$, we will need to consider some amount ε_t of $t_{\hat{x},\bar{y}}$ and ε_r of $r_{\hat{x},\bar{y}}$. Intuitively, we will use $t_{\hat{x},\bar{y}}$ and $r_{\hat{x},\bar{y}}$ to establish shadow values for (IC-FOC) and (IR), and then, for any particular perturbation p, consider the three deviations together where one uses $t_{\hat{x},\bar{y}}$ and $r_{\hat{x},\bar{y}}$ to undo the effect of p on (IC-FOC) and (IR).

Fix x_L, x_H, and $\hat{x}$. *A priori*, $\hat{x}$ may have arbitrary position relative to x_L and x_H, and moreover, in the case where p is t_{x_L,x_H}, one of x_L or x_H may not be in C_v, depending on whether ε_p is negative or positive. Define $x_0 < x_1 < \cdots < x_K$, $K \leq 4$, as elements of the set $\left\{\underline{y}, x_L, x_H, \hat{x}, \bar{y}\right\} \cap C_v$. For any given $\varepsilon = (\varepsilon_p, \varepsilon_t, \varepsilon_r)$, let $d(\cdot;\varepsilon) : [\underline{y}, \bar{y}] \to \mathbb{R}$ be given by

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

If x_L and x_H are both elements of $\{x_0, \cdots, x_K\}$, as must be true if p is r_{x_L,x_H}, then it follows that d is linear on each interval of the form (x_{k-1}, x_k). Assume that $x_H \notin \{x_0, \cdots, x_K\}$. Then, it must be that p is t_{x_L,x_H} with $\varepsilon_p \geq 0$. In this case, if $x_H \notin (x_{k-1}, x_k)$, then $d(\cdot;\varepsilon)$ is linear on (x_{k-1}, x_k), while if $x_H \in (x_{k-1}, x_k)$, then, since $\varepsilon_p \geq 0$, $d(\cdot;\varepsilon)$ is concave with two linear segments on (x_{k-1}, x_k). Finally, assume $x_L \notin \{x_0, \cdots, x_K\}$. Then, p is t_{x_L,x_H} with $\varepsilon_p \leq 0$, and once again, if $x_L \notin (x_{k-1}, x_k)$, then $d(\cdot;\varepsilon)$ is linear on (x_{k-1}, x_k), while if $x_L \in (x_{k-1}, x_k)$, then since $\varepsilon_p \leq 0$, $d(\cdot;\varepsilon)$ is once again concave with two linear segments on (x_{k-1}, x_k).

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

$$d_k(x;\varepsilon) = \begin{cases} L_k^-(x;\varepsilon) & x \leq x_{k-1} \\ d(x;\varepsilon) & x \in (x_{k-1}, x_k) \\ L_k^+(x;\varepsilon) & x \geq x_k \end{cases}.$$

Note that d_k is concave, and that as $|\varepsilon| = |\varepsilon_d| + |\varepsilon_t| + |\varepsilon_r| \to 0$, 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 x_k, where since $x_k \in C_v$, we can choose L_k such that $L_k(x) > v(x)$ for all $x \neq x_k$, and let

$$v_k(x) = \begin{cases} L_{k-1}(x) & x \leq x_{k-1} \\ v(x) & x \in (x_{k-1}, x_k) \\ L_k(x) & x \geq x_k \end{cases},$$

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

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

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

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

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

But then, except on the zero-measure set of points $\{x_0, \cdots, x_K\}$,

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

B.1.3 Shadow Values

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

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

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),

$$\begin{aligned} Q(\mathbf{0}) &= \left[\begin{array}{cc} \int t_{\hat{x},\bar{y}} f_a(x|a)\,dx & \int r_{\hat{x},\bar{y}} f_a(x|a)\,dx \\ \int t_{\hat{x},\bar{y}} f(x|a)\,dx & \int r_{\hat{x},\bar{y}} f(x|a)\,dx \end{array}\right] \\ &= \left[\begin{array}{cc} \int_{\hat{x}}^{\bar{y}} (x-\hat{x}) f_a(x|a)\,dx & \int_{\hat{x}}^{\bar{y}} f_a(x|a)\,dx \\ \int_{\hat{x}}^{\bar{y}} (x-\hat{x}) f(x|a)\,dx & \int_{\hat{x}}^{\bar{y}} f(x|a)\,dx \end{array}\right], \end{aligned}$$

and so

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

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 $(x-\hat{x})$ 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 ε in some ball around $\mathbf{0}$.) But then by the implicit function theorem, for each $p \in \{t_{x_L,x_H}, r_{x_L,x_H}\}$, we can on the appropriate neighborhood implicitly define $\varepsilon_r(\cdot)$ and $\varepsilon_t(\cdot)$ by

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

which is to say that starting from $\varepsilon = \mathbf{0}$, 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 $t_{\hat{x},\bar{y}}$ and $r_{\hat{x},\bar{y}}$.

Let λ be the rate of change of costs as one increases the agent's utility using $t_{\hat{x},\bar{y}}$ and $r_{\hat{x},\bar{y}}$. 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

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

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 increases the agent's incentives using $t_{\hat{x},\bar{y}}$ and $r_{\hat{x},\bar{y}}$ is

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

Given the shadow values λ and μ, the argument in Section 5 completes the proof of Proposition 2. ■

B.1.4 Proof of Sufficiency

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 concave contract satisfying (IC)-(LL). 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 5 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 $\delta(x) = v^*(x) - v(x)$, and let $\hat{v}(x;\varepsilon) = v(x) + \varepsilon\delta(x)$, 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, $\hat{v}(\cdot;\varepsilon)$ is concave and satisfies (IC)-(LL). Since $u^{-1}(\cdot)$ is convex, and since $\hat{v}(x;\varepsilon)$ is linear in ε, it follows that

$\int u^{-1}(\hat{v}(x;\varepsilon)) f(x|a)\,dx$ is convex in ε. Thus, since

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

it follows that

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

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

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

where the first equality follows by Lemma 3.

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

$$-\mu\int_{\mathcal{Y}\backslash S} \delta(x) f_a(x|a)\,dx = \mu\int_{S} \delta(x) f_a(x|a)\,dx.$$

Similarly, either (IR) is binding at v, in which case $\int \delta(x) f(x|a)\,dx \geq 0$, or

(IR) does not bind at v, in which case $\lambda = 0$, and hence in either case

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

Making these two substitutions thus yields

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

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

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

or equivalently,

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

Fix such an i, and consider δ_1, the restriction of δ to $S_i = [x_L, x_H]$. Since v is linear on S_i, and v^* is concave, δ_1 is concave. For any given K, let $\Delta = (x_H - x_L)/K$, and consider the function δ_K on $[x_L, x_H]$ that agrees with δ_1 on the set of points $\{x_L, x_L + \Delta, \cdots, x_H\}$, and is linear in between these points. Note that δ_K is concave and continuous on $[x_L, x_H]$. Choose K large enough that

$$\int_{S_i} n(x)\, \delta_K(x) f(x|a)\, dx < 0.$$

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

$$\tilde{\delta}(x) = \begin{cases} 0 & x \leq x_L \\ \delta_K(x) & x \in [x_L, x_H] \\ \delta_K(x_H) & x > x_H \end{cases}.$$

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

$[x_H, \bar{y}]$,

$$\int_{x_H}^{\bar{y}} n(x)\,\tilde{\delta}(x)\,f(x|a)\,dx = 0,$$

and hence,

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

Let us next argue that $\tilde{\delta}$ can be expressed as a sum of raises and tilts. For $k \in \{0, \ldots, K\}$, let $x_k = x_L + k\Delta$, and let s_k be the slope of $\tilde{\delta}$ on (x_{k-1}, x_k). Then, we claim that for all x in $[x_L, x_H]$,

$$\tilde{\delta}(x) = \delta(x_0)\,r_{x_0,\bar{y}}(x) + \sum_{k=1}^{K-1}(s_k - s_{k+1})\,t_{x_0,x_k}(x) + s_K t_{x_0,x_K}(x). \qquad (10)$$

To see (10) note first that for $x < x_0 = x_L$, both sides of the equation are 0. At x_0, each side is $\delta(x_0)$, since $r_{x_0,\bar{y}}(x_0) = 1$, and since $t_{x_0,\cdot}(x_0) = 0$. Thus, since both sides are continuous and piecewise linear on $[x_0, \bar{y}]$, it is enough that the two sides have that same derivative where defined. So, fix $\hat{k} \in \{1, \ldots, K\}$, and let $x \in \left(x_{\hat{k}-1}, x_{\hat{k}}\right)$. Note that for $k < \hat{k}$, $t'_{x_0,x_k}(x) = 0$, and for $k \geq \hat{k}$, $t'_{x_0,x_k}(x) = 1$. Hence, the derivative of the right-hand side is

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

as desired, and so, noting that $\tilde{\delta}'(x) = 0$ for $x > x_K = x_H$, we have established (10).

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

1. $\delta(x_0)\int n(x)\,r_{x_0,\bar{y}}(x)\,f(x|a)\,dx < 0$,
2. for some $k < K$, $(s_k - s_{k+1})\int n(x)\,t_{x_0,x_k}(x)\,f(x|a)\,dx < 0$, or
3. $s_K \int n(x)\,t_{x_0,x_K}(x)\,f(x|a)\,dx < 0$.

By Definition 1.1, and since x_0 is free, $\int n(x)\,r_{x_0,\bar{y}}(x)\,f(x|a)\,dx = \int_{x_0}^{\bar{y}} n(x)\,f(x|a)\,dx \geq 0$, and so 1. cannot hold. Since $\tilde{\delta}$ is concave on $[x_L, x_H]$, it follows that

$s_k - s_{k+1} \geq 0$, and so, since x_0 is free, it follows by Definition 1.2 that 2. cannot hold either. Finally, since x_0 and x_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

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

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,

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

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 a $y \in (\underline{y}, \bar{y})$ that is free. Then $n(y) \leq 0$ by Corollary 2. But $v(\cdot)$ is concave and $\rho(\lambda + \mu l(\cdot|a))$ is convex, so $n(y) \leq 0$ for all $y \in \mathcal{Y}$. But then (3) cannot hold between y and $\bar{y}$, so $v(\cdot)$ cannot be GHM, which contradicts that $v(\cdot)$ is optimal. So $v(\cdot)$ cannot have any interior free points, implying that $v(\cdot)$ is linear.

If $\rho(\lambda + \mu l(\cdot|a))$ is concave and (LL) does not bind, then it is GHM and hence optimal by Proposition 2. If (LL) binds, suppose that $v(\cdot)$ has two or more linear segments. Since $v(\cdot)$ is concave, Lemma 3 implies that $\rho(\lambda + \mu l(y|a))$ must be convex at some y; contradiction. So $v(\cdot)$ can have at most one linear segment, on $[y_L, y_H]$. Furthermore, $y_L = \underline{y}$, since otherwise Lemma 3 would imply that $\rho(\lambda + \mu l(y|a))$ is convex somewhere on $[y_L, y_H]$. ■

B.5 Proof of Proposition 3

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 4. *Assume $q > 0$ is not everywhere a constant, is analytic, and has $con(q) = \omega > -\infty$. Assume also that for some $\hat{x}$ on the interior of its domain, $q'(\hat{x}) = 0$. Let $\hat{k} = \min\left\{k | q^{(k)}(\hat{x}) \neq 0\right\} \geq 2$. Then, $q^{(\hat{k})}(\hat{x}) < 0$.*

Proof. 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 x in the domain of q,

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

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

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

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 $\frac{\xi^{(\hat{k}-2)}(\hat{x})}{(\hat{k}-2)!}(x-\hat{x})^{\hat{k}-2}$, where $\xi^{(\hat{k}-2)}(\hat{x}) = q(\hat{x})\, q^{(\hat{k})}(\hat{x})$. Hence, since $(x-\hat{x})^{\hat{k}-2}$ is strictly positive for $x > \hat{x}$, $q^{(\hat{k})}(\hat{x})$ must be strictly negative.

□

Using this lemma, we can prove the following claim, from which Proposition 3 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 $\hat{x}$ at which $\theta = 0$. We will show that immediately to the right of $\hat{x}$, $\theta < 0$. 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 $\hat{x}$ at which $\theta = 0$. Consider first the case that $g''(\hat{x})\, h''(\hat{x}) \neq 0$. Then, since $g' > 0$, it follows by (11) that $g''(\hat{x})\, h''(\hat{x})\, h'(\hat{x}) < 0$, and so, evaluated at $\hat{x}$,

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

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

Assume instead that $g''(\hat{x})\, h''(\hat{x}) = 0$, where, since $\theta(\hat{x}) = 0$, it follows that $g''(\hat{x}) = h''(\hat{x}) = 0$. Thus, since $con\,(g') > -\infty$, it follows from Lemma 4 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 $\hat{x}$, and we are done.

$\square$